

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РЕСПУБЛИКИ КАЗАХСТАН

Некоммерческое акционерное общество «Казахский национальный исследовательский
технический университет имени К.И.Сатпаева»



Институт автоматки и информационных технологий

УДК 004.932:623.4 (043)

На правах рукописи

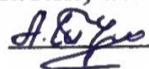
Ильясов Ернар Рамазанович

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ (ПРОЕКТ)
На соискание академической степени магистра

Название диссертации: « Исследование возможности совершенствования
модернизации установок досмотра аэропорта »

Направление подготовки: 7M07107 – Робототехника и мехатроника

Научный руководитель
к.т.н., ассоциированный профессор

 Туякбаев А.А.

«11» 06 2025 г.

Рецензент
к.т.н., ассоциированный профессор

 Сейдилдаева А.К.

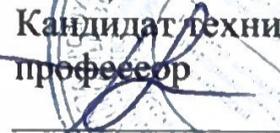
«11» 06 2025 г.

Норм контроль
Преподаватель

 Игембай Е.А.

«05» 06 2025 г.

ДОПУЩЕН К ЗАЩИТЕ
Заведующий кафедрой РТиТСА
Кандидат технических наук,
профессор

 К.А. Ожикенов

«12» 06 2025 г.

Алматы 2025

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РЕСПУБЛИКИ КАЗАХСТАН

Некоммерческое акционерное общество «Казахский национальный исследовательский
технический университет имени К.И.Сатпаева»



SATBAYEV
UNIVERSITY

Институт автоматки и информационных технологий
Кафедра Робототехники и технических средств автоматки

7M07107 – Робототехника и мехатроника

УТВЕРЖДАЮ

Заведующий кафедрой РТиТСА
Кандидат технических наук,
профессор

К.А. Ожикенов

« 12 » 06 2025 г.

ЗАДАНИЕ

на выполнение дипломного проекта

Магистранту Ильясов Ернар Рамазанович

Тема: Исследование возможности совершенствования модернизации установок досмотра
аэропорта

Утверждена приказом № ___ от «___» ___ 202 г.

Срок сдачи законченной диссертации : «___» ___ 2025 г.

Исходные данные к магистерской диссертации: Рентгеновские изображения багажа,
видеоданные с камер наблюдения, эталонная база биометрических признаков лиц,
нейросетевая модель YOLOv8, специализированный датасет рентгеновских изображений
(CS5800), язык Python, библиотеки OpenCV, Ultralytics, Roboflow, FastAPI.

Перечень подлежащих разработке в магистерской диссертации вопросов:

- а) Построение математической модели системы автоматизированного досмотра на основе анализа рентгеновских и видеоданных;*
- б) Моделирование и исследование процессов обнаружения угроз и идентификации личности на основе предложенной модели;*
- в) Синтез программно-аппаратного комплекса с проведением экспериментальной проверки и обработкой полученных результатов.*

Рекомендуемая основная литература:

1. Goodfellow I., Bengio Y., Courville A. *Deep Learning*. — MIT Press, 2016.
2. Szeliski R. *Computer Vision: Algorithms and Applications*. — Springer, 2022.
3. Redmon J. et al. *You Only Look Once: Unified, Real-Time Object Detection* // CVPR, 2016.
4. Bochkovskiy A., Wang C.Y., Liao H.Y.M. *YOLOv4: Optimal Speed and Accuracy of Object Detection*. — arXiv:2004.10934.
5. Jocher G. et al. *YOLOv8 Documentation*. — Ultralytics, 2023.
6. Schroff F. et al. *FaceNet: A Unified Embedding for Face Recognition and Clustering* // CVPR, 2015.
7. Zhang K. et al. *Joint Face Detection and Alignment using MTCNN* // IEEE Signal Processing Letters, 2016.
8. Mery D. *Automated X-ray Object Recognition and Inspection: A Review* // Computers in Industry, 2015.
9. Wang J. et al. *Object Detection in X-ray Baggage Screening: A Survey* // Sensors, 2020.
10. Zhao Y. et al. *Threat Object Detection on X-ray Images Using Deep CNNs* // Pattern Recognition, 2021.
11. Lin T.Y. et al. *Focal Loss for Dense Object Detection* // ICCV, 2017.
12. He K. et al. *Deep Residual Learning for Image Recognition (ResNet)* // CVPR, 2016.
13. Ren S. et al. *Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks* // NIPS, 2015.
14. Liu W. et al. *SSD: Single Shot MultiBox Detector* // ECCV, 2016.

ГРАФИК
подготовки магистерской диссертации

Наименование разделов, перечень разрабатываемых вопросов	Сроки представления научному руководителю	Примечание
Анализ методов досмотра и формирование датасета	08. 02. 2024	Выполнено
Разработка и обучение модели YOLOv8	25. 03. 2024	Выполнено
Проектирование архитектуры системы ориентации на базе	11. 05. 2024	Выполнено
Реализация системы и экспериментальная проверка	08. 09. 2024	Выполнено
Обобщение результатов и оформление диссертации	20. 03. 2025	Выполнено

Подписи

консультантов и норм контролера на законченную магистерскую диссертацию с указанием относящихся к ним разделов диссертации

Наименования разделов	Консультанты, Ф.И.О. (уч. степень, звание)	Дата подписания	Подпись
Основная часть	Туякбаев А.А. к.т.н., ассоциированный профессор		
Расчётный раздел	Туякбаев А.А. к.т.н., ассоциированный профессор		
Норм контролер	Игембай Е.А. Преподаватель	05. 06. 25	

Научный руководитель


подпись

Туякбаев А.А.

Ф.И.О.

Задание принял к исполнению обучающийся


подпись

Ильясов Е.Р.

Ф.И.О.

Дата

«___» _____ 2025 г

АНДАТПА

Бұл зерттеу жұмысы әуежайдағы жүкке рентгендік бақылау жүргізу процесін жетілдіруге бағытталған. Нақты мақсат – күдікті немесе қауіпті заттарды автоматты түрде танытын интеллектуалды жүйе құрастыру болып табылады. Осы зерттеу үшін YOLOv8x нейрондық желісі таңдалып, рентген кескіндерінен тұратын деректер жиыны негізінде қайта оқытылды. Жүйе нақты қауіп анықталған жағдайда ғана дабыл береді – бұл жалған сигналдар санын азайту үшін маңызды. Сонымен қатар, адамның бет-бейнесін тану модулі енгізіліп, әрбір оқиғаны нақты жолаушымен байланыстыру үшін, мүмкіндігі жасалды. Жүйенің жұмыс нәтижелері тәжірибе жүзінде сынақтан өткізілді: визуализация жасалып, басқа архитектуралармен салыстырылды. PR-кривая, F1 көрсеткіші және қателік матрицасы секілді метрикалар арқылы жүйенің тиімділігі көрсетілді.

АННОТАЦИЯ

Данная работа посвящена улучшению рентгеновского контроля багажа в аэропорту. Цель — разработка интеллектуальной системы, способной автоматически распознавать потенциально опасные объекты. Для этого использовалась нейросеть YOLOv8x, дообученная на базе рентгеновских изображений. Система выдаёт тревогу только при обнаружении реальной угрозы, что помогает снизить число ложных срабатываний. Также был добавлен модуль распознавания лиц, позволяющий связать тревожное событие с конкретным пассажиром. Система была протестирована, результаты сопоставлены с другими методами. Оценка проводилась по метрикам PR-кривой, F1-показателю и матрице ошибок.

ABSTRACT

This study aims to improve the process of X-ray baggage inspection at the airport. The primary objective of the work is to develop an intelligent system capable of automatically detecting suspicious or potentially hazardous objects. To achieve this goal, the YOLOv8x neural network was chosen, which was retrained on a set of X-ray images. The system is triggered only when a genuine threat is detected, thereby reducing the number of false alarms. A facial recognition module has also been incorporated, enabling the linking of each alarming situation to a specific passenger. The development was tested during experiments; the results were visualized and compared with other architectures. The effectiveness of the system is confirmed using metrics such as the precision-recall curve, the F1 score, and the confusion matrix.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	7
1. Системы контроля багажа и задачи автоматизации	7
1.1 Методы досмотра багажа	7
1.2 Рентгеновские технологии визуализации	11
1.3 Алгоритмы обнаружения угроз на изображениях	13
1.4 Биометрические методы идентификации	16
1.5 Проблемы существующих решений	19
1.6 Цель и задачи диссертации	22
1.7 Выводы по первой главе	25
2. Источники данных и выбор методов	26
2.1 Характеристика датасета рентгеновских изображений	26
2.2 Особенности данных и этапы предобработки	29
2.3 Анализ архитектур YOLOv8x, Faster R-CNN и других	32
2.4 Выбор базовой модели и аргументация	34
2.5 Метрики качества: PR, F1, confusion matrix	39
2.6 Обоснование итоговой архитектуры	43
2.7 Выводы по второй главе	45
3. Архитектура системы досмотра	46
3.1 Структура программного комплекса	46
3.2 Реализация рентген-модуля на базе YOLOv8	48
3.3 Блок распознавания лиц: структура и функции	50
3.4 Алгоритм фиксации события	51
3.5 Механизм срабатывания тревоги	53
3.6 Интерфейс визуализации объектов и событий	56
3.7 Выводы по третьей главе	59
4. Техническая реализация и испытания	60
4.1 Интеграция компонентов	60
4.2 Тестирование системы по сценариям	62
4.3 Сравнение с альтернативными подходами	64
4.4 Сводная оценка системы, ограничения	66
4.5 Выводы по четвёртой главе	68
ЗАКЛЮЧЕНИЕ	69
СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ	70
Приложение А	74
Продолжение приложения А	75
Продолжение приложения А	76
Приложение Б	77

ВВЕДЕНИЕ

Современные системы обеспечения безопасности в пассажирских зонах, таких как аэропорты и вокзалы, сталкиваются с необходимостью оперативной и точной оценки содержимого багажа. Используемые на практике рентгеновские установки позволяют визуализировать объекты внутри сумок и чемоданов без их вскрытия, однако в большинстве случаев оценка изображений производится вручную, оператором. Это влечёт за собой риски, связанные с человеческим фактором: утомляемость, перегрузка, субъективность. При этом формы потенциальных угроз становятся всё более нетипичными — оружие может быть скрыто под вид невинных предметов, а миниатюрные конструкции или нестандартные материалы значительно усложняют визуальное распознавание даже при хорошей подготовке персонала.

В условиях растущего пассажиропотока, особенно в час пик или при усиленных мерах досмотра, усложняется задача точной и быстрой фиксации тревожных объектов. В особенности — когда отсутствует возможность моментальной идентификации пассажира, прошедшего через установку, в случае обнаружения потенциальной угрозы. В таких случаях затрудняется обратная трассировка, оперативный вызов и фиксация факта. Ручной анализ рентгеновских снимков в отрыве от идентификации субъекта создает структурный разрыв в системе безопасности.

Настоящее исследование направлено на создание программной системы, которая позволяет не только автоматически находить опасные объекты на рентгеновском изображении, но и связывать их с конкретным лицом. Для реализации такой функциональности была выбрана архитектура, включающая модули компьютерного зрения и биометрической идентификации. В частности, для обработки рентгеновских снимков используется сверточная нейросеть YOLOv8x, способная в реальном времени обнаруживать заданные классы угроз с высокой точностью. Параллельно запускается видеомодуль, фиксирующий лицо человека, сдающего багаж. Совмещение этих потоков происходит по временным меткам, что позволяет системе формировать тревожное событие, включающее изображение угрозы, лицо субъекта, координаты объекта и степень уверенности.

Технически система реализуется на базе Python с использованием OpenCV, Ultralytics YOLO и dlib. Входящие изображения проходят этапы предобработки: нормализация, масштабирование, фильтрация. При регистрации тревоги формируется структурированный лог с возможностью последующего анализа. Также учитываются граничные случаи, например: множественные лица в кадре, частичное перекрытие, низкая контрастность снимка — для них предусмотрен дополнительный фильтр и механизм отсекающих ошибочных срабатываний.

Система проектировалась как модульная, с возможностью масштабирования и внедрения в уже существующую инфраструктуру

досмотра. Произведено моделирование сценариев прохождения багажа через рентген-рамку и параллельной фиксации видео с фронтальной камеры. Используемая логика тревоги основана на пороге уверенности $\theta \geq 0.45$, что соответствует F1-оптимуму на тестовом датасете. В экспериментальной части реализована запись событий в базу с возможностью сортировки по дате, классу угрозы и пассажиру. Макет системы прошёл лабораторную валидацию: проверены стабильность работы, корректность привязки объекта к лицу и устойчивость алгоритма при пограничных значениях.

В целом разработанная система ориентирована на применение в зонах массового прохода, где необходима минимизация человеческого участия, автоматизация фиксации тревожных событий и возможность последующего аудита. Подход, объединяющий детекцию и биометрию, может быть адаптирован как для стационарных комплексов, так и для мобильных модулей контроля.

1. Системы контроля багажа и задачи автоматизации

1.1 Методы досмотра багажа

Современные системы досмотра багажа представляют собой основу технологической инфраструктуры безопасности на объектах с высокой плотностью пассажиропотока, таких как международные и региональные аэропорты. Их основная задача заключается в предотвращении проноса запрещённых или потенциально опасных предметов в зону повышенного риска, включая оружие, стекло, жидкости, взрывчатые вещества, колюще-режущие предметы и иные объекты, способные нанести ущерб жизни, здоровью или транспортной инфраструктуре. За последние десятилетия контроль багажа строился на трёх опорах. Это визуальный осмотр, сканирование и ручной анализ. Эти методы до сих пор применяются, особенно в стандартизированных системах. Однако на фоне роста угроз и потока пассажиров, они всё чаще не справляются с задачами. Их ограниченность становится заметной при высоких требованиях к скорости и точности.

Самым простым способом остаётся ручной досмотр. Он используется там, где техника либо устарела, либо вообще отсутствует [1]. Он осуществляется сотрудником службы безопасности в процессе визуального и тактильного анализа содержимого багажа. Преимуществами такого подхода являются универсальность и способность оператора оценить контекст ситуации, в том числе поведение пассажира, характер упаковки и другие косвенные признаки риска. Однако на практике ручной досмотр оказывается крайне неэффективным в масштабируемых системах из-за высокой зависимости от человеческого фактора, физической утомляемости, невозможности быстрого анализа большого объёма багажа и потенциальных конфликтов с пассажирами [2]. Особенно уязвимы сценарии, в которых запрещённые предметы замаскированы под разрешённые или спрятаны в плотной структуре вещей – в таких случаях вероятность пропуска угрозы значительно возрастает. Кроме того, ручной досмотр не обеспечивает никакой цифровой фиксации событий, не допускает построения доказательной базы и практически не оставляет следов для последующего анализа инцидентов [3].

Металлодетекторы, рамки и ручные сканеры дополняют базовые процедуры досмотра за счёт регистрации присутствия металлических объектов. Они быстро реагируют на наличие проводящих материалов, являются дешевыми и легко внедряемыми средствами контроля. Однако их ограниченность заключается в невозможности детектирования неметаллических объектов, таких как стеклянные сосуды, пластиковые контейнеры, жидкости и электроника [3]. Кроме того, при использовании металлодетекторов оператору трудно интерпретировать, какой именно предмет вызвал срабатывание: одинаковую реакцию могут вызывать как безопасные (ключи, ремень), так и опасные (лезвие, скрытое оружие) объекты.

Это вызывает большое количество ложных тревог, перегружает персонал и замедляет общий поток пассажиров [4].

Технологическим прорывом в области контроля багажа стало широкое внедрение рентгеновских установок, позволяющих визуализировать содержимое багажа без его вскрытия [2]. Эти установки функционируют на основе прохождения рентгеновского излучения через объект и регистрации уровня его ослабления, что зависит от плотности и состава материалов. Полученные изображения транслируются на монитор в псевдоцветовой схеме: органические вещества отображаются в оранжевом, неорганические – в синем, металлы – в зелёном [5]. Такая цветовая дифференциация позволяет оператору визуально различать предметы по категориям, выявлять оружие, ножи, ёмкости с жидкостями, провода, батарейки, элементы взрывных устройств и прочие потенциальные угрозы.

Несмотря на высокую информативность, рентгеновские установки не лишены недостатков. Наиболее критичным из них является зависимость интерпретации от человеческого оператора. Качество анализа снимка определяется уровнем подготовки персонала, степенью усталости, внимательностью и временем, доступным для оценки каждого изображения [3]. При плотной укладке багажа объекты могут перекрываться, терять чёткость границ и сливаться с фоном, что повышает вероятность пропуска угроз [6]. На практике операторам приходится просматривать сотни рентгеновских кадров за короткий промежуток времени. Это напрямую влияет на внимание и снижает точность восприятия. Установки, которые используются сегодня, не умеют самостоятельно выделять тревожные объекты и не привязывают их к личности пассажира. Даже при обнаружении угрозы событие не фиксируется в цифровом журнале. В нём нет ни изображения, ни времени, ни признаков субъекта. Всё остаётся только в памяти оператора.

Что касается видеонаблюдения, то камеры чаще всего настроены на общий контроль. Они могут помогать следить за очередью или фиксировать инциденты, но редко работают синхронно с рентгеном. Даже если лицо попало в кадр, тревога никак с ним не связывается. Без чёткой привязки к событию такие записи мало пригодны для анализа или доказывания. В некоторых решениях оператор может вручную связать лицо с рентгеновским изображением при тревоге. К сожалению, этот подход не автоматический и не подходит для работы в потоке. Он плохо масштабируется, а также требует участия человека. Таким образом, видеонаблюдение остаётся важным инструментом общего сопровождения потока, но не выполняет функцию замкнутой сцепки "угроза – субъект".

Учитывая ограничения всех вышеперечисленных подходов, в последние годы на первый план выходят интеллектуальные методы досмотра, основанные на автоматическом анализе рентгеновских изображений с использованием алгоритмов глубокого обучения [7]. Наиболее эффективными

среди них считаются модели семейства YOLO [8], особенно версии YOLOv4 и YOLOv8, способные в реальном времени распознавать опасные объекты на изображениях багажа с высокой точностью [9]. Применение таких моделей позволяет устранить зависимость от субъективной оценки оператора, обеспечить стабильность распознавания при наложении объектов, плотной упаковке и вариативности фоновых условий. Кроме того, в системе реализуется логика срабатывания тревоги при обнаружении объекта, относящегося к категории угроз, без участия человека [10].

Интеллектуальная система досмотра, реализованная в рамках данной работы, дополняет рентгеновский модуль автоматической детекции объектов нейросетевым блоком фиксации и распознавания лиц. В момент анализа изображения система захватывает лицо субъекта при помощи RGB-камеры, создаёт эмбединг с помощью алгоритма `face_recognition` [11] и либо соотносит его с существующей базой данных, либо маркирует как неизвестного. Лицо субъекта не влияет на факт срабатывания тревоги, но логически привязывается к событию, создавая цифровую запись, пригодную для анализа и доказательной фиксации. Такой подход обеспечивает не только автоматизацию реакции, но и юридическую валидность тревоги, позволяя в любой момент восстановить: кто прошёл контроль, какой объект был обнаружен, с какой вероятностью, в какое время и при каких условиях.

На основании обобщения типовых сценариев и анализа открытых отчётов можно выделить распределение типов угроз, наиболее часто встречающихся в рентгеновских изображениях багажа. Примерно 35% составляют колюще-режущие предметы, включая ножи, лезвия, отвёртки и шила. Около 25% приходятся на огнестрельное оружие, его части и самодельные конструкции. Стеклянные ёмкости, бутылки и ампулы составляют порядка 18%. Жидкости в запрещённой таре или в замаскированных контейнерах – около 10%. Плотные неоднородные предметы с подозрительной компоновкой – 5%, остальные 7% – это разного рода неидентифицированные или сомнительные объекты.

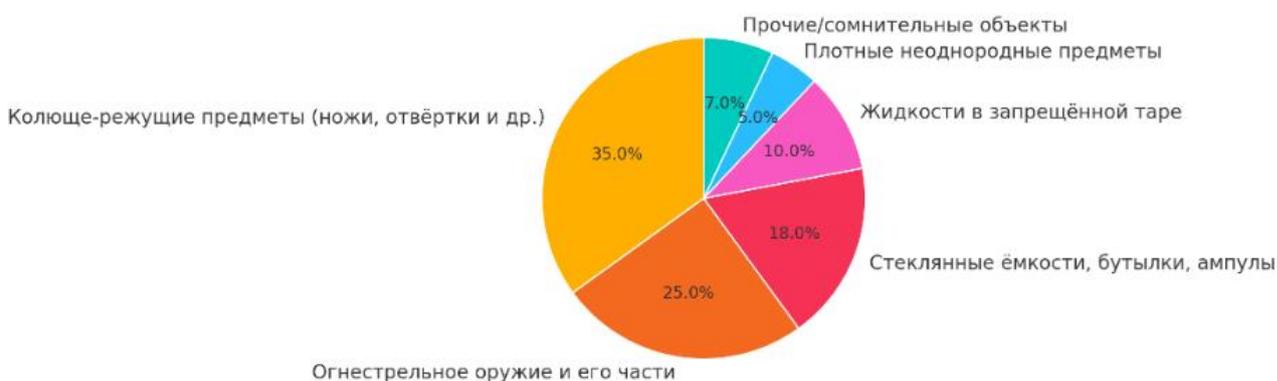


Рисунок 1.1 – Распределение типов угроз на рентгеновских изображениях багажа

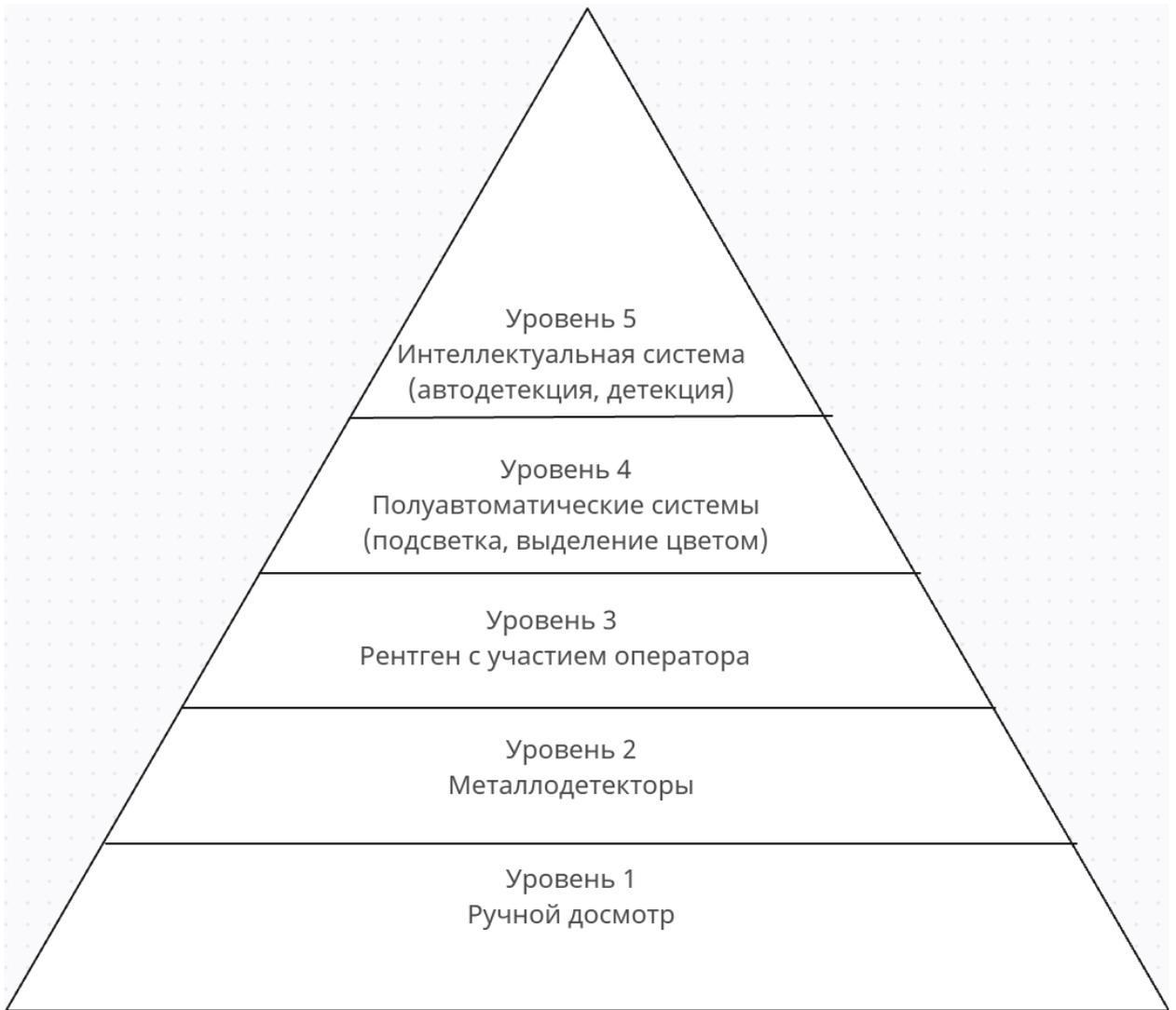


Рисунок 1.2 – Пирамида автоматизации методов досмотра

Пирамида автоматизации (Рисунок 1.2), в которой уровень 1 соответствует ручному досмотру, уровень 2 – металлодетекторам, уровень 3 – рентгеновскому анализу с участием оператора, уровень 4 – полуавтоматическим системам с подсветкой и выделением объектов, и уровень 5 – полностью интеллектуальной системе с автодетекцией, фиксацией лица, логированием и визуализацией. Разработанная модель показывает, в каком направлении движется развитие систем контроля.

На практике становится ясно, что стандартные методы досмотра не справляются с современными условиями. При большом потоке людей и росте нестандартных угроз ручная проверка даёт сбои, а система часто перегружена и не замечает важное. Чтобы повысить точность и надёжность, приходится использовать сразу несколько технологий. Здесь важен не только нейросетевой анализ рентгеновских изображений, но и автоматическая тревожная логика, а также возможность связать событие с конкретным человеком через распознавание лица. Такая комбинация работает быстрее и

точнее, чем любой из этих подходов по отдельности. Именно на такой архитектуре и строится система, разработанная в рамках данной работы.

1.2 Рентгеновские технологии визуализации

Рентгеновские системы остаются основным инструментом при досмотре багажа. Они дают возможность изучить содержимое предметов, не вскрывая их. В отличие от металлодетекторов и ручной проверки, рентген показывает весь внутренний объём, включая то, что скрыто под обёрткой, внутри чемодана или контейнера [2].

Суть метода в том, что излучение проходит сквозь объект. Чем плотнее материал, тем слабее сигнал. На выходе получается изображение, где разные плотности отображаются оттенками серого или псевдоцветами. Наиболее распространённая цветовая схема отображает органические материалы (бумага, еда, ткань, кожа) в оранжевых тонах, неорганические (пластик, стекло) – в зелёных, а металлы и плотные элементы – в синих и фиолетовых [5]. Такая дифференциация даёт оператору возможность предположить состав и потенциальную опасность предметов, но финальное решение всё равно остаётся за человеком.

Преимущества рентгеновского контроля очевидны: это высокая скорость анализа, возможность видеть внутреннюю структуру предметов, неприкосновенность багажа, а также относительная объективность изображения. В условиях высокой нагрузки линии рентгеновской установки способны обрабатывать от 60 до 150 человек в час, что делает их незаменимыми для крупных хабов [4]. Однако эффективность этих установок напрямую зависит от интерпретации оператором. При этом уровень перегрузки, утомления, субъективной интерпретации формы и плотности объектов существенно снижает надёжность даже в хорошо подготовленных командах. В реальных условиях оператору отводится в среднем от 2 до 5 секунд на анализ каждого снимка, что в совокупности с наложением предметов, разнообразием упаковок, нестандартными углами и плохо различимыми силуэтами создаёт предпосылки для ошибок, промахов и задержек [6].

Характер рентгеновских изображений существенно отличается от классических RGB-снимков: объекты представлены в проекциях, могут перекрываться, терять глубину и не иметь текстурных признаков. Геометрия предметов часто искажена из-за угла прохождения луча, а плотные объекты могут затенять слабые. При этом визуально разрешённый предмет может по своей плотности и контуру напоминать запрещённый, и наоборот. Например, канцелярский нож может выглядеть как безобидная металлическая вставка в чехле, а стеклянная бутылка с жидкостью – как термос или контейнер с косметикой [3]. Это затрудняет ручную интерпретацию и создаёт высокий

процент ложных тревог. По данным отраслевых исследований, до 40% ложных срабатываний в рентгеновских установках приходится на предметы, внешне напоминающие угрозы, но не являющиеся таковыми – зарядные устройства, электроника, косметика, посуда, кабели и даже пищевые контейнеры [12]. Именно такие ситуации наиболее опасны с точки зрения перегрузки оператора: после нескольких десятков подобных "ложных тревог" внимание притупляется, и возрастает шанс пропуска настоящей угрозы.

Дополнительную сложность представляет собой фактор наложения объектов, особенно в плотно упакованном багаже. При наложении органических и неорганических материалов в цветовой палитре возникает смешение оттенков, ухудшается контрастность, а чёткие контуры становятся размытыми. Объекты, расположенные на разной глубине, могут выглядеть как одно целое или быть частично замаскированы. В таких условиях даже опытный оператор может не различить угрозу, особенно если она нетипична, компактна или представлена нестандартной формой [3]. Классические рентген-системы, несмотря на наличие режима увеличения, усиления контраста и инверсии цветов, не способны интерпретировать такие сцены автоматически, и тем более не формируют тревогу по результатам анализа [6].

Критическим ограничением рентгеновских систем является отсутствие автоматизированной фиксации тревожного события. Даже при явном обнаружении запрещённого предмета оператор не может нажать кнопку, чтобы зафиксировать лицо субъекта, идентификатор сканера, изображение угрозы и параметры события [4]. В подавляющем большинстве случаев тревога формируется только в оперативной реальности: сумку просят открыть, пассажира задерживают, но никаких цифровых следов не остаётся. Отсутствует связка между рентген-кадром и лицом, между ID события и временем, между операторским решением и содержимым багажа. Это делает последующий анализ невозможным, затрудняет обучение персонала и полностью исключает формирование базы нарушителей. В отличие от систем видеонаблюдения, где тревожные кадры можно сохранять и анализировать, рентген-снимки чаще всего исчезают с экрана сразу после принятия решения. В результате даже при ошибке – либо пропуске угрозы, либо необоснованной тревоге – невозможно восстановить контекст и провести разбор [3].

Некоторые производители рентгеновских систем пытаются внедрять простые алгоритмы автоматической подсветки потенциально опасных объектов. Однако эти алгоритмы основаны на банальном анализе плотности и формы, не используют машинное обучение и не обладают адаптивностью [2]. При этом в реальных условиях эффективность таких решений крайне низкая, а операторы быстро перестают обращать внимание на сигналы, если те часто ложные. Именно поэтому в современных системах досмотра появляется потребность в более интеллектуальной обработке рентгеновского изображения, способной проводить не просто цветовую фильтрацию, но

полноценную многоклассовую детекцию с применением нейросетевых моделей [7, 9].

Рентгеновские изображения имеют особенности, которые отличают их от обычных фото. Эти особенности усложняют задачу автоматического анализа.

Во-первых, цветовая палитра ограничена. Чаще всего это оттенки серого или псевдоцвет. Изображения почти не содержат текстур, теней и ярких деталей. Это снижает эффективность классических RGB-моделей [6].

Во-вторых, есть проекционные искажения. Объекты могут выглядеть сжатыми, вытянутыми или накладываться друг на друга.

В-третьих, плотные предметы затемяют более лёгкие. Из-за этого появляются ложные контуры.

Все эти особенности требуют адаптированных архитектур. Стандартные подходы плохо работают на таких данных. Лучше использовать модели, специально обученные на рентгеновских изображениях. Например, YOLO, до обученные на наборах с багажом [8, 13]. Такие сети могут распознавать объекты по их форме и плотности. Это важно, когда визуальные признаки слабо выражены. Благодаря этому удаётся различать ножи, бутылки, стекло, пистолеты и другие потенциально опасные предметы.

Таким образом, рентгеновские методы визуального контроля багажа являются технологическим ядром всех современных систем досмотра, однако они имеют ряд фундаментальных ограничений, связанных с отсутствием интеллектуального анализа, зависимостью от оператора и невозможностью цифровой фиксации тревожного события [4]. Их интеграция с нейросетевыми моделями и биометрической идентификацией становится логичным следующим этапом развития, позволяющим превратить рентген-станцию из пассивного отображающего устройства в активный интеллектуальный модуль, автоматически определяющий угрозы, связывающий их с субъектом и сохраняющий событие в лог [7, 11].

1.3 Алгоритмы обнаружения угроз на изображениях

Автоматизация процесса обнаружения угроз на рентгеновских изображениях багажа является одним из ключевых направлений развития современных систем досмотра. С ростом потока пассажиров и усложнением потенциальных угроз ручной просмотр изображений становится всё менее эффективным. В условиях ограниченных ресурсов и высокой нагрузки возникает потребность в системах, которые способны сами анализировать рентгеновские данные и принимать решение о наличии опасных предметов [3]. Однако такие изображения сильно отличаются от обычных: в них нет привычных цветов и текстур, а сами объекты представлены в виде

плотностных проекций. Поэтому здесь требуются алгоритмы, способные справляться с подобной визуальной сложностью [6].

На ранних этапах внедрения автоматизации в системы досмотра использовались базовые алгоритмы машинного зрения. Они работали по принципу ручного извлечения признаков и были довольно ограниченными. В их основе лежали методы вроде фильтрации изображения, поиска границ, сегментации по цвету или плотности, а также геометрического анализа — например, с применением преобразования Хафа, дескрипторов контура или операторов Лапласа [14, 15]. В теории такие подходы могли находить простые формы на фоне, например, длинные тонкие объекты или круглые бутылки. Однако на практике эти алгоритмы плохо справлялись с перегруженными изображениями, где объекты накладываются друг на друга или повернуты под разными углами. Любые отклонения от привычной формы приводили к ошибкам — система либо не находила объект вовсе, либо принимала один предмет за другой. Особенно это мешало в ситуациях, когда объекты похожи внешне, но относятся к разным категориям риска: например, нож и отвёртка или бутылка с водой и термос. Алгоритмы не могли обучаться на новых примерах и не учитывали контекст, из-за чего тревожная логика оставалась примитивной. В целом такие методы были скорее вспомогательными и использовались только там, где не требовалась высокая точность и надёжность [3]. С появлением алгоритмов глубокого обучения в виде свёрточных нейросетей (CNN) открылось новое направление в обработке визуальной информации. Такие архитектуры способны автоматически извлекать признаки с разных уровней абстракции, что делает их особенно пригодными для задач, связанных с визуально сложными и многозначными входными изображениями [16]. При этом нейросети не требуют ручной настройки признаков — они обучаются на примерах, что позволяет охватить весь спектр возможных вариаций объектов: по форме, плотности, ориентации, освещению, фону и упаковке [17].

Среди наиболее известных архитектур для детекции объектов на изображениях можно выделить семейства Faster R-CNN [18], SSD и YOLO. Faster R-CNN представляет собой двухстадийную модель: сначала сеть генерирует регионы интереса (Region Proposal Network), затем производит классификацию объектов в каждом регионе [18]. Модель обеспечивает высокую точность и хорошо находит мелкие объекты. Однако из-за высокой нагрузки на вычисления и задержек её сложно использовать в режиме реального времени, особенно при большом потоке пассажиров. Вариант SSD (Single Shot Detector) даёт более быструю работу, но хуже справляется при сложных наложениях предметов [19].

Таблица 1.1 – Характеристики моделей, применяемых для детекции объектов на изображениях досмотра

Архитектура	Механизм работы	Точность (mAP)	Скорость (FPS)	Устойчивость к особенностям рентген изображении	Применимость в реальном времени
YOLOv5	Однопроходная CNN	~0.50–0.58	30–40	Умеренная устойчивость	Да
YOLOv8	Улучшенная однопроходная CNN	~0.53–0.61	30–50	Высокая устойчивость	Да
SSD	С anchor боксами	~0.40–0.50	20–30	Низкая устойчивость, много ошибок	Частично
Faster R-CNN	Двухэтапная модель	~0.55–0.65	5–15	Умеренная устойчивость, малая скорость	Нет
DETR	Трансформерная архитектура	~0.60–0.68	<10	Высокая устойчивость, малая скорость	Нет

Наиболее эффективным решением для задачи рентгеновской детекции стала архитектура YOLO – You Only Look Once. Её основное преимущество заключается в том, что сеть обрабатывает всё изображение за один проход, одновременно локализуя и классифицируя объекты [8]. Такая архитектура обеспечивает быструю обработку –до 20–30 кадров в секунду. При этом сохраняется достойный уровень точности.

Последние версии моделей, такие как YOLOv5 и YOLOv8, показывают явный прирост по сравнению с предыдущими поколениями. Это стало возможным за счёт более глубоких слоёв, внедрения трансформеров, улучшенных функций потерь и продвинутой предобработки данных [10, 7, 13]. В данной работе применялась модель YOLOv8. Она хорошо справляется с рентгеновскими изображениями, включая сцены с перекрытиями, искажениями и шумом. Модель сочетает высокую скорость с точностью. Её структура позволяет адаптироваться под разные форматы входных данных, работать с необычными цветовыми схемами и гибко подстраиваться под новую разметку при дообучении [20].

Рентгеновские изображения требуют особого подхода. В отличие от привычных RGB-снимков, они не содержат текстур, тени или ярких цветов. Поэтому универсальные модели работают с ними хуже. Специальная подготовка и адаптация под такой тип данных обязательна. Рентгеновские изображения часто отличаются искажённой формой объектов, фрагментированными элементами, множественными наложениями и плотностными слитиями. Кроме того, в них отсутствуют привычные визуальные ориентиры –такие как цветовая текстура и тени [6].

При этом важно понимать: использование нейросетей в анализе рентгеновских данных – это не просто автоматизация взгляда оператора. Это переход к иной логике восприятия угроз. Модель не только фиксирует предметы, но и соотносит их с заранее заданными классами риска. После этого она принимает решение, запускать ли тревожный сигнал [7, 21]. Такой подход позволяет системе самостоятельно формировать тревожное событие с полной фиксацией всех его параметров – изображения, координат, типа угрозы, уровня уверенности, времени и, при необходимости, связанного с ним субъекта [11].

Сегодня нейросетевой анализ не просто дублирует действия человека. Он формирует принципиально иной подход к обработке событий – основанный на формальной логике, воспроизводимости и способности к анализу, лишённой человеческого фактора.

Методы автоматического обнаружения угроз заметно продвинулись. Если раньше применялись простые алгоритмы, использующие геометрические фильтры и сегментацию, то сейчас используются сложные архитектуры глубокого обучения. Такие модели способны обрабатывать рентгеновские изображения даже при сложной упаковке, наложении предметов и визуальной неоднозначности.

Современные архитектуры – такие как YOLO, Faster R-CNN, SSD и их модификации показывают разный баланс между точностью, скоростью и устойчивостью [8, 18, 19]. При этом только однопроходные модели, обрабатывающие изображение за один шаг, могут обеспечить необходимую производительность. Это особенно важно для зон досмотра с высокой нагрузкой и требованием работы в реальном времени [9]. При работе с рентгеновскими изображениями эффективность нейросетевых моделей во многом определяется их способностью адаптироваться к особенностям визуального сигнала: отсутствию текстур, проекционным искажениям и вариативности плотности. Эти факторы требуют специальной настройки архитектуры и предобработки данных. В современных решениях по автоматизации досмотра такие модели рассматриваются как основа для создания точных и надёжных систем выявления потенциальных угроз.

Дальнейшие главы будут посвящены анализу архитектур, выбору подходящей модели, настройке параметров и проверке качества на практическом примере. Также будет показано, как выбранное решение интегрируется в интеллектуальную систему досмотра и влияет на формирование тревожной логики.

1.4 Биометрические методы идентификации

Биометрическая идентификация личности на основе изображения лица является одним из наиболее активно развивающихся направлений в современных системах обеспечения безопасности. Её широкое распространение объясняется сочетанием высокой скорости, неинвазивности

и достаточно высокой точности, при условии корректной реализации алгоритма и наличия качественных данных [11, 22]. Применительно к задачам досмотра, идентификация субъекта на момент прохождения контроля имеет критически важное значение: она позволяет не только связывать тревожные события с конкретным пассажиром, но и формировать базу повторяющихся инцидентов, исключать анонимность и вносить персонализированные параметры в логи досмотровых операций [4].

В контексте интеллектуальной системы досмотра задача биометрической идентификации решается параллельно с анализом рентгеновского изображения. В момент срабатывания события или по инициативе оператора система должна зафиксировать лицо субъекта, сгенерировать его цифровое представление – эмбединг – и сопоставить его с локальной или распределённой базой. В случае успешного сопоставления присваивается идентификатор; если лицо не найдено – запись маркируется как «неизвестный субъект». Это позволяет не только логически увязать объект досмотра (багаж) и его содержимое с субъектом (лицом), но и создать юридически обоснованную и воспроизводимую цифровую запись, включающую все параметры события. Такая структура, в отличие от фрагментарной записи видеонаблюдения, обеспечивает формальную сцепку и даёт возможность аналитического анализа поведения, повторяемости и статистики по конкретным пассажирам [11].

Методы распознавания лиц можно условно разделить на несколько поколений: ранние подходы, основанные на геометрии (алгоритмы Eigenfaces, Fisherfaces), методы, использующие локальные шаблоны (Local Binary Patterns, LBP), и современные нейросетевые архитектуры, оперирующие с эмбедингами высокой размерности [14]. Первые поколения обеспечивали работу на стационарных изображениях в контролируемых условиях и были крайне чувствительны к изменению освещения, повороту головы и эмоциям. Алгоритмы типа LBP стали прорывом за счёт выделения устойчивых локальных признаков, однако и они были ограничены разрешением камеры и необходимостью фронтальной съёмки. Только с появлением свёрточных нейросетей стало возможным создание систем распознавания, работающих в условиях реального времени, со слабо контролируемым освещением, поворотами лица, частичным закрытием и вариациями в изображении [22, 23].

Современные архитектуры, такие как DeepFace, FaceNet, ArcFace и их производные, работают на принципе генерации векторного представления лица – эмбединга – и сопоставления его с уже сохранёнными представлениями других лиц в базе данных [22]. Эмбединг представляет собой точку в многомерном пространстве, в котором расстояние между точками определяет степень сходства. Таким образом, распознавание сводится к задаче векторного сравнения и оценке косинусного расстояния между эмбедингами. Такой подход крайне эффективен: он позволяет сохранять тысячи лиц и выполнять их сравнение за миллисекунды, а также легко

масштабируется [11]. Кроме того, применение стандартных библиотек, таких как face_recognition на базе ResNet, позволяет внедрить такую систему в состав интеллектуальной платформы без необходимости разработки собственных моделей с нуля [11].

В контексте реализации интеллектуальной системы досмотра была проанализирована применимость существующих архитектур с точки зрения точности, скорости обработки, устойчивости к условиям съёмки и возможности локальной работы без подключения к интернет-сервисам. В таблице ниже приведено сравнение ключевых подходов к распознаванию лиц, отражающее эволюцию технологий и обоснование выбора нейросетевой модели на базе ResNet [22, 23].

Таблица 1.2 – Сравнение методов биометрической идентификации по лицу.

Метод	Принцип работы	Устойчивость к повороту	Чувствительность к освещению	Скорость сравнения	Применимость в реальном времени
Eigenfaces	Статистический PCA-анализ	Низкая	Высокая	Высокая	Нет
LBP	Локальные бинарные шаблоны	Средняя	Средняя	Высокая	Ограниченно
DeepFace	Сверточная нейросеть + кастом слой	Средняя	Средняя	Средняя	Да
FaceNet / ArcFace	Эмбединги с loss-функцией на расстояние	Высокая	Средняя	Высокая	Да
face_recognition	ResNet-архитектура, open-source реализация	Высокая	Низкая	Высокая	Да

Применение face_recognition, использующего архитектуру ResNet-34, обосновано несколькими факторами: наличие открытого исходного кода, простота интеграции в Python-среду, стабильная работа с базами более 1000 лиц, возможность локального сопоставления эмбедингов и устойчивость к изменению ракурса и освещения [11, 24]. Визуально архитектура ResNet обеспечивает сохранение глубокой структуры признаков при уменьшении риска затухания градиента, что критично для обучения моделей, обрабатывающих изображения с неидеальной съёмкой [17]. В

интеллектуальной системе досмотра ResNet применяется не как классификатор, а как генератор эмбедингов – модель обучена отличать лица по дистанциям в пространстве признаков, а не по конкретным ID [22].

Важно отметить, что в разрабатываемой системе тревога не формируется по факту отсутствия лица в базе. В отличие от многих систем, где нераспознанное лицо автоматически трактуется как фактор риска, в данном подходе лицо рассматривается как дополнительная информация. Основанием для тревоги служит исключительно наличие угрозы в багаже. При этом фиксация субъекта остаётся важной: она позволяет восстановить полную картину события, использовать данные для обучения и анализа, а также вести персонализированный лог досмотра [23]. Таким образом, биометрический модуль выполняет вспомогательную, но значимую роль – он обеспечивает связь между тревогой и конкретным человеком, не влияя на само её срабатывание.

Многочисленные исследования и практика внедрения показали, что современные технологии распознавания лиц, основанные на использовании нейросетевых эмбедингов, могут эффективно применяться в системах досмотра. Такие алгоритмы хорошо справляются с задачами в условиях изменяющегося освещения, различных углов обзора и частичного перекрытия лиц. Главное их преимущество – возможность привязать тревожное событие к конкретному человеку, проходящему через контроль. Это позволяет не просто зафиксировать угрозу, но и установить, кто именно был её носителем. В результате формируется целостная цифровая запись, включающая как изображение объекта (например, оружия или стеклянной ёмкости), так и лицо субъекта, а также дополнительные параметры: дату, время, идентификатор станции, уровень риска и другие метаданные. Все эти элементы сохраняются в единой структуре, которая впоследствии может быть использована как для визуального анализа, так и для подготовки отчётов, статистики и юридических материалов. Таким образом, система не только обеспечивает высокую точность при обнаружении угроз, но и создаёт полную и воспроизводимую картину произошедшего. В следующем разделе подробно рассматривается, каким образом эта логика интегрирована в рентгеновский модуль, и как она влияет на принятие решений в рамках всей системы.

1.5 Проблемы существующих решений

Несмотря на широкое использование рентгеновских установок, металлодетекторов, камер и отдельных нейросетевых решений, большинство систем досмотра по-прежнему страдают от серьёзных ограничений. Многие из них связаны не только с техническими аспектами, но и с общей архитектурой – подход к проектированию часто остаётся фрагментарным, без единой логики. Это снижает надёжность и делает реакцию системы нестабильной. Одной из ключевых проблем остаётся сильная зависимость от человека – оператора, который вручную оценивает кадры и принимает решение. Даже в системах,

оборудованных современными рентгеновскими установками, анализ изображения осуществляется вручную, а результат контроля зависит от субъективной интерпретации [2]. Усталость, психологическое выгорание, отсутствие должного обучения и перегрузка в часы пик приводят к снижению концентрации внимания и росту процента как ложных тревог, так и пропущенных угроз [21]. Поскольку рентгеновское изображение, особенно при плотной укладке багажа, требует анализа нескольких наложенных и частично перекрытых объектов, оператор может просто не заметить опасный предмет, особенно если он представлен в нестандартной форме, замаскирован или частично закрыт другими предметами. Это усугубляется тем, что на одну точку контроля приходится несколько сотен человек в час, и на каждый рентген-снимок отводится лишь несколько секунд визуального анализа [6].

Один из заметных недостатков большинства существующих решений — это отсутствие автоматической фиксации тревожных событий. Даже если система показывает на экране объект, который может быть опасным, тревога не формируется сама. Оператор должен вручную нажать кнопку или внести пометку [4]. Если он этого не сделал – факт угрозы теряется. При этом никакой связи между багажом и человеком, который его принёс, система не устанавливает. Нет единого события, нет журнала, где бы фиксировались такие случаи. Это сильно усложняет работу в случае инцидента: невозможно быстро найти момент, когда появилась угроза, кто был рядом и какое решение было принято. Даже если в помещении есть камеры, а сканер пишет видео, все данные хранятся отдельно, без синхронизации [26]. Получается, что каждый модуль работает сам по себе. Никакой сцепки между рентгеновским кадром, лицом пассажира и временем досмотра нет. Поэтому, если оператор не зафиксировал всё вручную – восстановить полную картину задним числом, скорее всего, уже не получится.

Ещё одна важная проблема связана с тем, что элементы интеллектуальной обработки в существующих решениях внедряются лишь частично. Во многих случаях системы используют только простые функции — например, выделение плотных участков или контуров, основанных на геометрии объектов [2, 25]. Однако полноценные нейросетевые модули в таких решениях часто отсутствуют. Даже если компьютерное зрение применяется, оно, как правило, основано на устаревших алгоритмах и слабо адаптировано под особенности рентгеновских изображений [6]. Из-за этого система либо слишком часто ошибается, либо не распознаёт угрозы вовсе. К тому же многие решения страдают от плохой разметки, отсутствия специализированных рентгеновских наборов данных и невозможности дообучения модели на реальных сценах [28]. Всё это сильно снижает точность и сводит на нет эффект от внедрения таких технологий.

Даже при наличии нейросетевой модели, распознающей угрозы, остаётся критический разрыв: нет механизма идентификации субъекта. Лицо пассажира может быть зафиксировано камерой, но если оно не привязано к

рентгеновскому событию – это просто видеозапись, не имеющая доказательной силы [11, 29]. В системах, где распознавание лиц присутствует, оно используется автономно, без логической связи с фактом тревоги. Например, в системе может быть база “лиц риска”, но тревога по факту неформальной идентификации – это не доказательство наличия угрозы. Напротив, тревога, вызванная обнаружением объекта, но не привязанная к лицу, теряет контекст. В итоге системы либо “бьют тревогу” по слабому основанию, либо остаются без фиксации реальных угроз.

Если говорить про архитектуру, то большинство действующих решений не формируют полноценного события. Нет такого элемента, который бы объединял всё важное в одном месте: изображение багажа, распознанный класс объекта, уровень уверенности, лицо человека, точное время, идентификатор станции и служебные параметры журнала [30]. Из-за этого нельзя построить ни нормальный лог тревог, ни сводную аналитику по людям, ни систему оповещений, ни подготовить корректную выборку для переобучения модели. Всё работает разрозненно. Такие системы скорее просто фиксируют момент, чем действительно реагируют. Они делают снимок, но не сохраняют суть — не формируют события, не выводят тревогу и не связывают данные между собой. Получается, что платформа «молчит» – она ничего не объясняет, не реагирует и не учится на том, что уже произошло [31].

Стоит отдельно подчеркнуть, что в некоторых системах тревога может срабатывать просто из-за того, что система не распознала лицо [23]. Такой подход нельзя назвать ни эффективным, ни разумным. Реакция должна происходить только тогда, когда зафиксирован реальный признак угрозы — например, если найден нож или стеклянная бутылка [12]. Узналось лицо или нет — это важно для логов, но не должно быть причиной тревоги. Если действовать иначе, система начнёт выдавать слишком много ложных срабатываний, будет мешать нормальным пассажирам и создавать лишнюю нагрузку на сотрудников [3]. Правильная логика выглядит иначе: сначала фиксируется угроза, а уже потом — кто именно проходил досмотр. Это лицо просто сохраняется в систему, чтобы потом можно было посмотреть или проанализировать ситуацию [11, 22].

Также стоит подчеркнуть недостаточную прозрачность логирования в традиционных решениях. Даже в тех системах, где тревоги регистрируются, лог часто представлен в виде текстовой записи, не включающей изображение, координаты, уверенность модели, эмбединг лица или визуализацию события [26]. Это делает невозможным аудит, обучение и интеграцию с другими системами безопасности. Для построения масштабируемой платформы необходимо, чтобы каждое событие было полноценно оформлено: содержало все параметры, хранилось в структурированном формате (JSON, CSV, SQL) и имело ссылки на изображения, эмбединги, временные коды [32, 33]. Лишь в таком случае становится возможной агрегация данных, построение отчётов,

формирование статистики по времени суток, типам угроз, зонам досмотра и другим параметрам [34].

Таблица 1.3 – Сравнительная оценка традиционных и интеллектуальных систем по степени полноты событийной модели

Параметр	Традиционные системы	Интеллектуальные системы
Обнаружение угроз	Визуально (оператор)	Алгоритмически (ИИ)
Тревожная логика	Реактивная	Превентивная
Привязка к субъекту (идентификация)	Отсутствует	Формализована
Формирование цифрового события	Частичное, ручное	Автоматическое
Логирование	Фрагментарное	Структурированное
Повторная аналитика	Невозможна	Поддерживается

В совокупности все перечисленные недостатки – зависимость от оператора, отсутствие автоматической тревоги, разрыв между объектом и субъектом, фрагментарное применение нейросетей, недостоверность логирования и архитектурная разобщённость – делают существующие решения неприменимыми в условиях современных требований к безопасности [3, 4, 31]. Поэтому возникает потребность в более продвинутой системе, способной не только распознавать объекты на рентгеновских изображениях, но и связывать это с фиксацией лица, формировать событие целиком и сохранять его в цифровом виде. Такая система должна быть пригодной как для анализа, так и для юридического оформления тревожных случаев [30, 26]. Её основой выступают современные нейросетевые модели, которые могут работать с плотной укладкой, наложениями и шумом в реальном времени. При этом тревога формируется не как исключение, а как логичное следствие обнаруженной угрозы [35, 36].

1.6 Цель и задачи диссертации

Анализ существующих систем досмотра показывает, что даже при наличии рентгеновских сканеров, автоматизированных интерфейсов и элементов машинного зрения, в них по-прежнему не хватает главного – логической целостности. Во всех изученных решениях отсутствует связка

между обнаружением угрозы, распознаванием лица и цифровой регистрацией события [3, 25]. Компоненты таких систем работают изолированно, без общей модели тревоги. В результате выходные данные не складываются в единое, понятное для анализа и реакции событие. Тревога, как правило, остаётся либо устной, либо ручной, и не представляет собой информационно насыщенную структуру, пригодную для повторного анализа, обучения, аудита или судебной экспертизы [26]. Именно это обстоятельство делает существующие решения устаревшими и неподходящими для масштабируемого применения в условиях высокой пропускной способности и многообразия потенциальных угроз [4].

Формально выявленные проблемы можно свести к четырём критическим разрывам. Во-первых, это отсутствие автоматического обнаружения угроз на рентгеновских изображениях [2]. Большинство систем полностью полагаются на оператора, что делает контроль зависимым от человеческого восприятия и скорости реакции [37]. Во-вторых, тревога не фиксируется в структурированном виде: изображение, время, объект, уровень риска и лицо субъекта не объединяются в единую запись, а потому не могут быть использованы как формальное событие [30]. В-третьих, система не идентифицирует субъект досмотра: видеопоток существует отдельно, база лиц – отдельно, тревога – отдельно, и ни один из этих компонентов не способен логически связать друг с другом объект угрозы и конкретного человека [22, 23]. В-четвёртых, отсутствует цифровая архитектура логирования: даже если тревога произошла, она исчезает из системы, не оставляя следа, что полностью исключает возможность построения аналитики, анализа повторяемости, предотвращения инцидентов и масштабирования системы в более сложные условия [38, 39].

На основе критического обзора, сравнительного анализа и анализа пробелов можно сформулировать основную цель настоящего исследования как разработку и реализацию интеллектуальной системы досмотра багажа, обеспечивающей автоматическое обнаружение потенциально опасных объектов на рентгеновских изображениях и логическую привязку тревожного события к конкретному субъекту с использованием биометрической идентификации по лицу [11, 29]. Такая система должна не только выполнять детекцию угроз в автоматическом режиме, но и создавать информационно насыщенное тревожное событие, фиксируя лицо, объект, изображение, время и уровень риска в едином цифровом представлении [24, 32]. При этом ключевой особенностью должно быть разделение логики тревоги и логики идентификации: тревога возникает только по факту обнаружения угрозы, а лицо фиксируется как дополнительный компонент события – без влияния на факт его инициирования [21]. Это позволит избежать ложных срабатываний по признаку «неизвестное лицо» и сконцентрироваться на реальных признаках угрозы, что соответствует принципам инженерной устойчивости и операционной эффективности [40].

Исходя из поставленной цели, в рамках работы решаются следующие взаимосвязанные задачи: проанализировать существующие подходы к автоматизации досмотра, в том числе методы рентгеновского анализа и нейросетевой детекции [8, 10]; исследовать архитектуры идентификации субъекта, основанные на распознавании лиц и эмбединговом сравнении [11, 22]; обосновать выбор архитектуры системы, включая компоненты анализа изображения, фиксации субъекта и формирования события [41, 35]; реализовать программный прототип, интегрирующий детекцию объектов, распознавание лица, автоматическую логику тревоги и цифровое логирование [7, 30]; и, наконец, провести экспериментальное тестирование реализованной системы на примерах сценариев с различной плотностью потока, характером угроз и уровнями освещения [36, 41].

Таким образом, цель исследования – сформировать технически выверенную и практическую архитектуру системы досмотра, в которой каждый тревожный случай оформляется как завершённое цифровое событие. Такое событие включает все важные элементы: изображение угрозы, сведения о субъекте, координаты, время, уровень уверенности и реакцию системы [30, 42]. Это обеспечивает не только более точное реагирование, но и создаёт основу для дальнейшего анализа, формализованного учёта и соблюдения юридических требований.

1.7 Выводы по первой главе

В данной главе проведён системный обзор современных методов досмотра багажа и пассажиров, включая ручные, полуавтоматические и интеллектуальные подходы. Представлены основные принципы работы традиционных систем контроля – металлодетекторов, рентгеновских установок и камер видеонаблюдения, а также проанализированы уровни автоматизации, применяемые в существующих решениях.

Отдельное внимание уделено рентгеновским методам визуального контроля. Описаны физические принципы рентгеновской проекции, псевдоцветовое кодирование материалов, особенности плотностной визуализации и наложения объектов. Обосновано, что интерпретация таких изображений требует высокой квалификации оператора и подвержена человеческому фактору, что делает задачу автоматизации особенно актуальной.

В рамках анализа методов автоматического обнаружения угроз рассмотрены существующие алгоритмы компьютерного зрения, включая классические методы на основе фильтров признаков (SIFT, Haar, LBP), а также современные нейросетевые архитектуры: YOLO, SSD, Faster R-CNN и др. Анализ показал, что традиционные методы не справляются с рентгеновскими изображениями. Им не хватает устойчивости и гибкости, особенно в сложных визуальных условиях. В то же время нейросетевые модели показывают гораздо лучшие результаты, если есть подходящая обучающая выборка.

Также были рассмотрены методы биометрической идентификации по лицу. Современные алгоритмы на основе сверточных сетей (например, face_recognition) способны уверенно распознавать лица даже при слабом освещении или изменении ракурса. Это позволяет связывать тревожные события с конкретным человеком, что повышает надёжность системы и снижает число ложных срабатываний.

В ходе анализа выявлены основные проблемы действующих решений:

- нет чёткой и автоматической логики тревоги;
- отсутствует связь между объектом (багаж) и субъектом (человек);
- итоговая интерпретация всегда зависит от оператора;
- тревожные события не сохраняются в структурированном цифровом виде.

Исходя из этого, были определены цели и задачи работы. Основная цель – построить систему, которая может автоматически выявлять угрозы на рентгеновских изображениях, привязывать их к человеку по изображению лица, создавать тревожные события и показывать результат в реальном времени. В результате глава задала теоретическую и проблемную основу для дальнейшей работы. На ней опирается проектирование архитектуры, выбор алгоритмов и программная реализация, рассмотренные в следующих главах.

2. Источники данных и выбор методов

2.1 Характеристика датасета рентгеновских изображений

Реализация интеллектуальной системы досмотра, основанной на автоматическом анализе рентгеновских изображений, требует наличия специализированного набора данных, отражающего реальные условия визуального контроля багажа. В отличие от классических изображений в RGB-формате, рентгеновские снимки обладают принципиально иной структурой визуальных признаков: они не содержат текстур, цветовых градиентов, теней и перспективных искажений, характерных для стандартной фотографии. Вместо этого информация кодируется через проекционные свойства – плотность, форму, силуэт, наложение и расположение объектов внутри замкнутого пространства [6, 42]. Это обуславливает необходимость использования датасета, содержащего именно такие изображения – полученные в условиях реального досмотра, в псевдоцветовой гамме, с соответствующей разметкой опасных объектов [28].

В рамках настоящего исследования в качестве исходного источника данных был использован рентгеновский датасет, собранный с применением открытых платформ аннотированных рентген изображений багажа, включая объекты, относящиеся к потенциально опасным [5, 42, 43]. Датасет охватывает широкий спектр сцен, полученных с рентгеновских сканеров. В него входят как одиночные объекты – пистолеты, ножи, бутылки – так и сложные композиции с плотной укладкой и перекрытиями. Многие изображения имеют неоднозначную визуальную структуру из-за наложений и псевдоцветов. Размеры варьируются от 640×480 до 1280×720 пикселей, данные представлены в формате RGB, что соответствует цветовой проекции плотностей [41].

При отборе основное внимание уделялось качеству аннотаций. Ключевым критерием была точная разметка всех классов угроз – с координатами, метками и классами объектов [36]. Всего собрано свыше 5000 изображений, из них примерно 3700 содержат хотя бы один опасный объект. Общее число разметок превышает 8500, включая сцены с несколькими классами и повторяющимися объектами одного типа. Все объекты разделены на несколько групп. Это огнестрельное и холодное оружие (ножи, лезвия), стеклянные и пластиковые ёмкости, а также плотные и непонятные по назначению предметы типа инструментов, техники или контейнеров [28, 43].

Датасет включает широкий спектр сцен, отражающих реальные условия досмотра. Объекты встречаются как отдельно, так и в плотной укладке с наложениями. Такая вариативность позволяет учесть случаи, когда опасный предмет скрыт среди других вещей, частично закрыт тканью или размещён под нестандартным углом. В некоторых изображениях также намеренно присутствуют макеты угроз – они используются для проверки эффективности систем контроля [5]. Визуальные примеры представлены на рисунках 2.1 и 2.2,

где можно наблюдать типичную структуру сцены и особенности рентгеновского визуального сигнала.

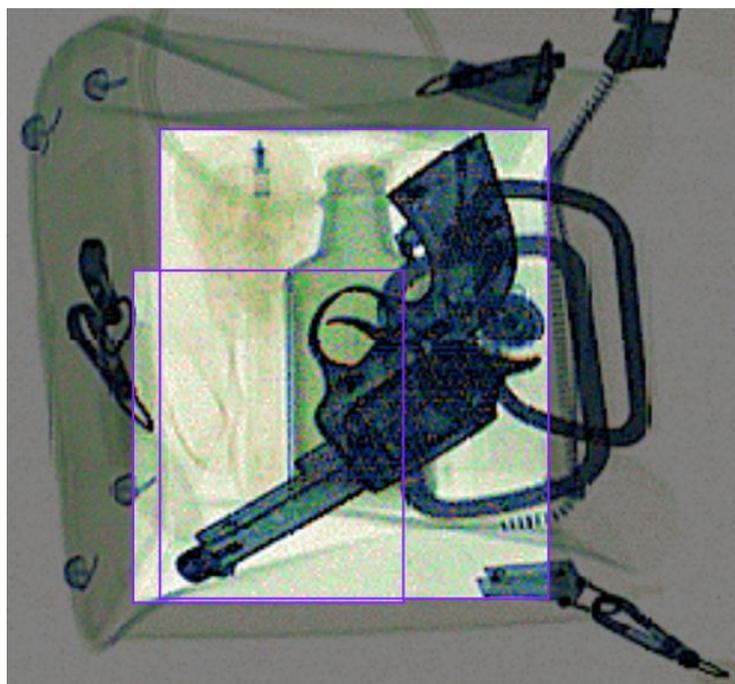


Рисунок 2.1 – Рентген изображение багажа с запрещенным предметом

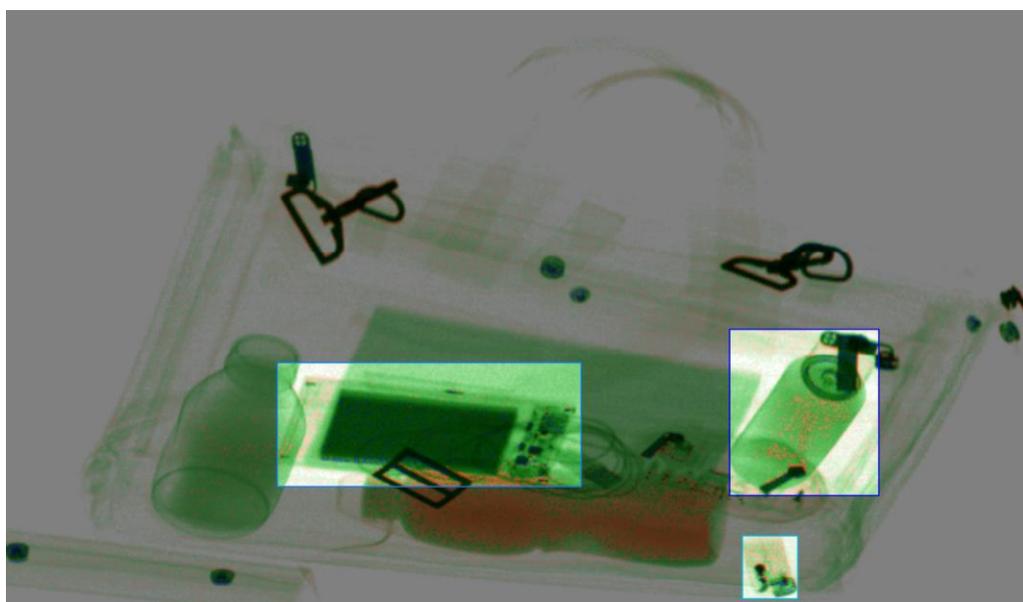


Рисунок 2.2 – Рентген изображение багажа с разрешенными объектами (powerbank, зажигалка, емкость с жидкостью)

В отличие от обычных фотографий, рентгеновские изображения демонстрируют сквозную проекцию: границы объектов не всегда чёткие, материалы полупрозрачны, а цвета кодируют не реальную окраску, а уровень поглощения излучения. Это создаёт уникальные условия для задач детекции,

требующие адаптированной архитектуры модели и специализированных техник предобработки [6, 44].

В наборе также встречаются объекты, которые внешне похожи на опасные, но не являются угрозой. Это, например, ложки, гаечные ключи, батарейки, зиплоки с жидкостью и даже обычная техника [43, 28]. Такие элементы важны для проверки – сможет ли модель отличить настоящую угрозу от похожих, но безопасных предметов. Кроме того, в выборке есть и нейтральные изображения, где нет ни одного объекта из списка угроз. Они нужны для того, чтобы модель не выдавала ложные тревоги, когда в кадре нет опасности. Это помогает снизить уровень ложноположительных срабатываний в реальном времени [27].

Для обучения и проверки модели использовался рентгеновский датасет, собранный из открытых источников, включая Roboflow и Kaggle. Разметка в нём была доработана с учётом задач автоматического досмотра по классам объектов, которые чаще всего представляют угрозу.

В качестве основы использован датасет CS5800 (Roboflow X-ray Baggage), включающий рентгеновские изображения багажа с нейтральным и опасным содержимым. Из всей выборки было отобрано свыше 5000 изображений, из которых примерно 3700 содержат как минимум один объект угрозы.

Разметка выполнена вручную, с точным определением координат. Всего размечено более 8500 bounding box. Формат аннотаций для YOLO: координаты объектов масштабированы относительно размеров снимка, каждый объект подписан числовым кодом класса.

В наборе представлены 12 типов угроз. Среди них: ножи (knife), пистолеты (gun), бутылки с жидкостями (bottle), баллончики и аэрозоли (sprayer), патроны (bullet), стеклянная тара (glass), режущие и острые инструменты (scissors, pliers), электронные устройства и батареи.

Размеры изображений варьируются от 640×480 до 1280×720 пикселей. Формат –RGB, но цвета соответствуют рентген-плотности: зелёный –металл, оранжевый –органика, синий –неорганика. Это отличает данные от обычных фото, поэтому модель требует предварительной адаптации.

При подготовке датасета пришлось решить несколько важных задач. На первом этапе удалили ошибочную и повторяющуюся разметку. Также были убраны изображения без объектов –если только они не использовались специально для балансировки.

Одна из главных проблем –сильный дисбаланс классов. Категории knife и bottle попадались в несколько раз чаще других объектов. Для исправления этого пришлось выполнить дополнительную балансировку и сокращение слишком частых классов.

После этого всю разметку привели к единому формату, подходящему для обучения модели YOLOv8.

Полученный в итоге датасет отличается высоким разнообразием сцен и плотностью аннотаций. Он включает сцены с перекрытиями, различной композицией и ракурсами, максимально близкими к реальному досмотру багажа. Такие характеристики позволяют модели стабильно обучаться и показывать хорошие результаты.

Итоговый набор данных подходит не только для обучения, но и для объективного тестирования нейросетевых алгоритмов. Он обладает достаточной репрезентативностью и практической применимостью для реальных задач безопасности [42, 45]. Применение модели позволяет адаптировать её под особенности рентгеновских данных. Это позволит заложить основу для дальнейшего внедрения в систему досмотра. Такая система должна стабильно работать даже при наложениях, плотной укладке и неполной видимости объектов.

В следующем разделе описан процесс подготовки данных: нормализация, балансировка классов, аугментация и стандартизация изображений. Эти шаги важны для того, чтобы модель работала надёжно при нестабильных визуальных условиях и переменном видеопотоке.

2.2 Особенности данных и этапы предобработки

Рентгеновские изображения багажа сильно отличаются от обычных фото. Стандартные модели компьютерного зрения не рассчитаны на такую специфику, где важны плотность, перекрытие и отсутствие текстур. Рентгеновские изображения представляют собой плотностные проекции, полученные после прохождения излучения через багаж и его частичного поглощения различными материалами. Детекторы фиксируют остаточный сигнал, и на его основе строится визуальное представление содержимого [28, 38]. В отличие от обычных фотографий, такие снимки лишены текстур, теней и реалистичной цветовой гаммы [6, 42]. Вместо этого данные отображаются в псевдоцветах: оттенки отражают не внешний вид предметов, а их физические свойства – в частности, плотность. Как правило, зелёный соответствует металлу, оранжевый – органике, а синий – неорганическим материалам [15, 43]. Это создаёт особые условия как для интерпретации, так и для обработки изображений.



Рисунок 2.3 – Сравнение рентгеновского и обычного изображения одного и того же объекта

Основной особенностью рентген изображений является высокая степень наложения объектов. В большинстве случаев содержимое багажа не разложено в одном слое, как это требуется для точного анализа, а находится в плотной, хаотичной и многослойной укладке. Это приводит к потере чёткости контуров, частичному перекрытию, слиянию цветовых каналов и формированию визуально неоднозначных структур [6, 42]. Дополнительно, рентгеновское изображение является не проекцией поверхности объекта, а его плотностной тенью, вследствие чего даже простые предметы могут выглядеть неузнаваемо при изменении угла, формы или положения. Отсутствие текстурных признаков исключает применение большинства классических алгоритмов извлечения признаков, таких как SIFT, SURF или ORB, ориентированных на контрастные и локальные детали [14].

Наличие псевдоцветовой гаммы в рентген изображениях также требует специализированной обработки. В отличие от RGB-изображений, где цвет несёт смысловую нагрузку (например, кожа, одежда, металл), в рентгеновских изображениях цветовая шкала задаётся программно и отображает категории материалов. Стандартная палитра, принятая во многих рентгеновских установках, включает оранжевый для органических веществ, синий для неорганики и зелёный для металлов [15, 43]. Однако на этапе предобработки эти цвета часто необходимо либо нормализовать, либо преобразовать в градации серого для исключения зависимости от устройства визуализации [45]. Кроме того, при использовании различных источников данных возможны

расхождения в цветовых каналах – то, что в одной системе отображается как зелёное, в другой может иметь иной оттенок, что требует стандартизации палитры и устранения цветового шума [38].

Предобработка изображений начинается с нормализации размеров и цветового пространства. Для большинства архитектур нейросетей, включая YOLO, требуется фиксированное разрешение входного изображения – как правило, 640×640 или 416×416 пикселей [7, 10]. Поэтому изображения приводятся к единому масштабу с сохранением пропорций, добавлением отступов (padding) и центрированием сцены. Цветовое пространство при этом может быть переведено в градации серого или оставлено в псевдо-RGB, в зависимости от конфигурации модели [20]. В ходе экспериментов установлено, что сохранение трёхканального формата изображения, даже при отсутствии привычной цветовой информации, положительно влияет на стабильность модели. Архитектуры, обученные на RGB-данных, лучше справляются с задачей, так как изначально рассчитаны на работу с трёхканальными входами [13]

Далее проводится удаление шумов и усиление контуров. Для этого используются методы повышения контраста, например CLAHE. Также применяется выравнивание гистограммы и пространственные фильтры. В обработке участвуют фильтр Собеля, Лапласа и медианный фильтр. Они улучшают видимость границ и упрощают работу модели [38, 39]. Эти методы позволяют улучшить видимость границ объектов, особенно при низкой контрастности и сильной плотности упаковки. Также производится удаление или подавление визуального мусора: рамок рентген-экрана, интерфейсных элементов и номеров кадров, которые присутствуют на некоторых исходных изображениях и не несут смысловой нагрузки для модели [27].

Аугментация данных играет ключевую роль в обеспечении устойчивости модели к искажённым или нестандартным входам. В процессе подготовки датасета применялись такие методы, как горизонтальное и вертикальное отражение, масштабирование, поворот на случайный угол, вырезание фрагментов (cutout), шум Гаусса, а также случайное изменение контрастности и яркости [36, 45]. Эти методы помогают нейросетевой модели распознавать объекты вне зависимости от их положения, формы и сопутствующих перекрытий. В рентгеновских изображениях важно сохранять точную геометрию так как они чувствительны к искажениям. Поэтому при масштабировании аннотаций особенно следили за тем, чтобы не терялась форма рамки [32].

Чтобы улучшить баланс классов, в выборку добавили больше изображений с редкими объектами. Это, например, стеклянные ёмкости или плотные контейнеры. Картинки без угроз тоже оставили – как примеры, на которых модель учится не реагировать. Это позволяет модели не переобучаться на самые частые классы (например, ножи), а учитывать

контекст и не реагировать на безопасные, но визуально схожие объекты [28, 41].

В результате предобработка рентгеновских изображений включает в себя многоуровневую цепочку действий, обеспечивающую приведение данных к формату, пригодному для обучения нейросетевой архитектуры, и адаптацию модели к особенностям рентген формата. Без этих этапов модель либо не будет распознавать реальные сцены, либо будет склонна к ложным срабатываниям и переобучению на искажённых признаках [46]. В следующем разделе будет проведён сравнительный анализ моделей детекции, обученных на предобработанных данных, с целью выбора оптимальной архитектуры для интеграции в систему интеллектуального досмотра.

2.3 Анализ архитектур YOLOv8x, Faster R-CNN и других

Одним из наиболее критичных этапов проектирования интеллектуальной системы досмотра является выбор архитектуры нейросетевой модели, применяемой для автоматического выявления объектов угроз на рентгеновских изображениях. От этого выбора напрямую зависят точность срабатываний, скорость обработки кадров в реальном времени, устойчивость к визуальному шуму и способность системы функционировать в условиях многослойных перекрытий объектов. В контексте задачи досмотра на основе рентген данных модель должна не просто распознавать объекты, но делать это быстро, надёжно, с высокой чувствительностью к сложным сценам, характерным для багажа – плотным, непредсказуемым по структуре, содержащим как потенциально опасные, так и безопасные предметы [6].

Исторически задачи детекции решались с использованием каскадных методов и ручных признаков, однако с развитием глубокого обучения наибольшее распространение получили архитектуры, основанные на свёрточных нейросетях. Среди них наиболее заметны три подхода: двухступенчатые модели, такие как Faster R-CNN, одноступенчатые модели, такие как SSD, и высокоэффективные архитектуры реального времени – семейство YOLO [18,19].

Архитектура Faster R-CNN построена на принципе разделения задач локализации и классификации объектов. На первом этапе осуществляется генерация регионов интереса (Region Proposal Network), на втором – классификация каждого региона. Эта модель отличается высокой точностью, особенно при наличии мелких объектов, и используется в задачах, где допустимы задержки и важна максимальная надёжность [18]. Однако её основное ограничение – низкая скорость обработки и высокий объём вычислений, что делает её малоприменимой для применения в реальном времени, особенно в условиях досмотра, где каждое изображение должно быть проанализировано за доли секунды [29].

Модель SSD (Single Shot Detector) представляет собой компромисс между точностью и скоростью. Она реализует детекцию объектов в один проход по изображению с использованием anchor-боксов [19]. Модель SSD показывает приемлемую точность на стандартных изображениях, но при работе с рентген-сценами её эффективность заметно снижается. Это особенно проявляется в случаях плотной укладки и наложения объектов. Основная причина – слабая способность точно определять границы без выраженных текстурных признаков [6].

В противоположность этому, архитектуры семейства YOLO лучше подходят для условий досмотра. Они реализуют одностадийный подход: изображение делится на ячейки, для каждой из которых сразу предсказываются координаты, класс и уровень уверенности [8, 9]. Такие модели обеспечивают хорошую точность при высокой скорости обработки, что делает их пригодными для видеопотока и работы в реальном времени.

Особенно хорошо себя зарекомендовали современные версии YOLO – начиная с YOLOv5 и заканчивая YOLOv8x. В последних итерациях появились улучшения: трансформерные блоки, оптимизация функций потерь, а также встроенные механизмы адаптации к нестандартным входным данным [5, 6, 48]. Благодаря этому модели сохраняют устойчивую работу даже при анализе сложных изображений, характерных для рентген-контента [13]. Особенность YOLOv8x (экстра-широкой версии модели) заключается в увеличенном числе параметров, расширенных возможностях по локализации мелких объектов, а также устойчивости к визуальным искажениям за счёт глубокой многослойной архитектуры [7, 13]. В сравнении с другими версиями YOLO данная модификация требует больше ресурсов, однако обеспечивает значительно более высокое качество детекции, особенно при наличии плотного шума, наложения и слабой контрастности объектов. Это делает её подходящей для сценариев досмотра, где точность приоритетнее абсолютной скорости, но всё же сохраняется способность работать в реальном времени на современных GPU [29].

Сравнительный анализ этих архитектур с точки зрения ключевых параметров приведён в изображении ниже.

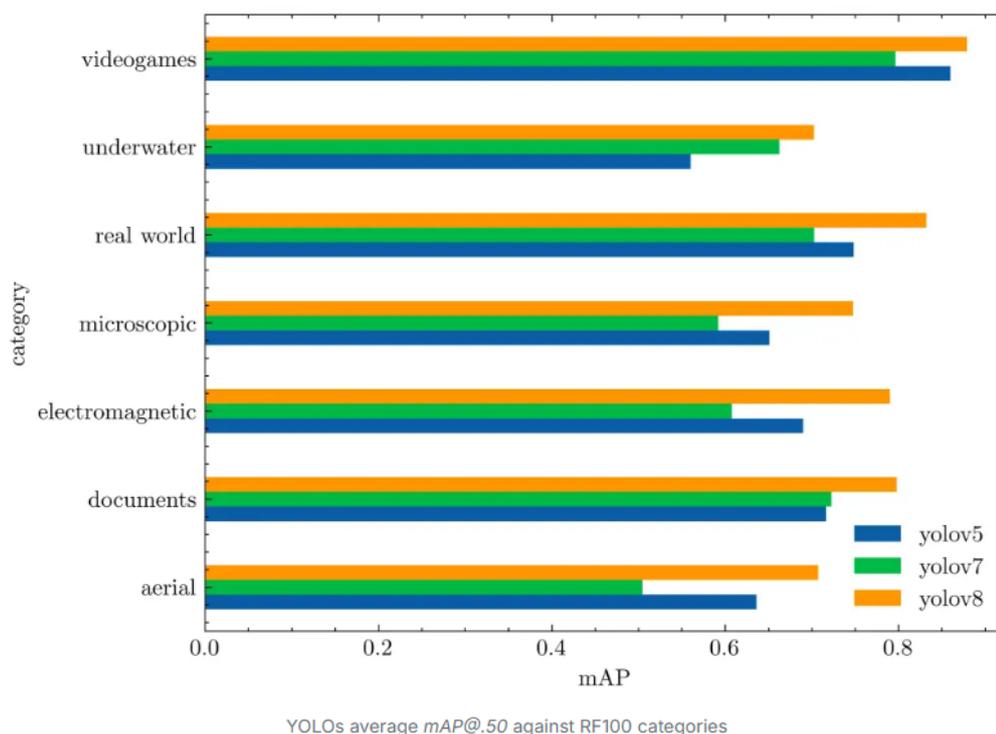


Рисунок 2.4 – Сравнительная характеристика нейросетевых моделей детекции объектов

Сравнение показывает, что наилучшее соотношение между точностью и производительностью демонстрируют модели YOLOv8x, особенно при условии качественной предобработки данных [7, 13]. В отличие от Faster R-CNN, они пригодны для работы в реальном времени и способны локализовать множественные объекты в плотной сцене [29]. По сравнению с предыдущей версией (YOLOv7), модель YOLOv8x лучше справляется с особенностями рентгеновских изображений. Её архитектура более устойчива к шуму, а улучшенная функция потерь снижает количество ложных срабатываний на фоне [35].

Поэтому именно YOLOv8x выбрана в качестве основной модели для дальнейшей доработки системы и анализа рентгеновских данных. В следующем разделе объясняется, почему было выбрано именно это решение. Также представлены результаты тестирования модели: PR-кривые, значения F1 и confusion matrix. Эти данные легли в основу окончательного выбора детектора угроз для внедрения в систему досмотра.

2.4 Выбор базовой модели и аргументация

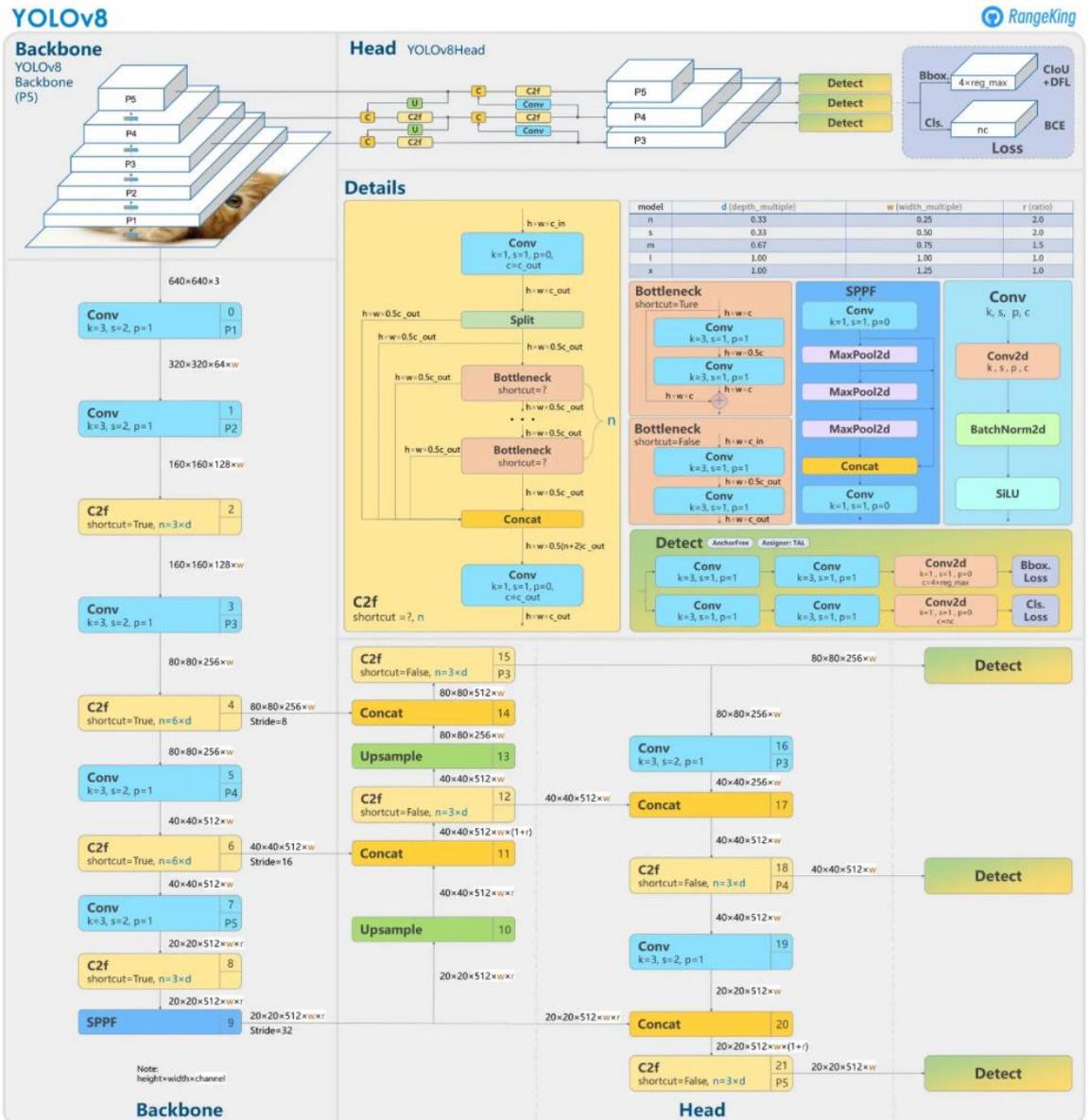
Рентгеновские изображения багажа характеризуются высокой плотностью объектов, отсутствием текстурной информации, проекционными искажениями и множественными наложениями. Эти особенности

существенно осложняют задачу точной локализации потенциально опасных предметов, особенно при использовании классических моделей объектной детекции [6]. В условиях визуального досмотра критично не только распознать объект, но и сделать это в реальном времени, с минимальной ошибкой и высокой воспроизводимостью. Такие требования предъявляют повышенные ожидания к выбору базовой архитектуры модели [35].

Выбор архитектуры нейросетевой модели для задачи детекции объектов угроз на рентгеновских изображениях багажа является критически важным этапом построения всей интеллектуальной системы досмотра. От этой модели зависит не только точность распознавания потенциально опасных предметов в сложных и визуально неоднозначных сценах, но и общая скорость отклика системы, устойчивость к ложным срабатываниям, возможность функционирования в реальном времени и способность масштабироваться на новые данные [29].

В рамках исследования было выдвинуто рабочее предположение: модель YOLOv8x, отличающаяся глубокой архитектурой, без якорной (anchor-free) головой и механизмом каскадной агрегации признаков, сможет наиболее эффективно справляться с особенностями рентгеновских изображений [7, 10, 20]. Ожидалось, что её структура обеспечит не только высокую точность детекции и устойчивость к перекрытию объектов, но и позволит построить систему, способную обрабатывать поток данных вблизи реального времени [13].

Анализ нейросетевых архитектур, представленных в предыдущем разделе, показал, что модели из семейства YOLO демонстрируют лучшие результаты в условиях задачи. Особенно выделяется последняя модификация –YOLOv8x, которая сочетает в себе высокую чувствительность при распознавании и приемлемую нагрузку на вычислительные ресурсы [35, 13].



YOLOv8 Architecture, visualisation made by GitHub user RangeKing

Рисунок 2.5 – Структура модели YOLOv8

YOLOv8 (You Only Look Once, версия 8) – одна из современных моделей для задач обнаружения объектов. Её отличают высокая скорость обработки и хорошая точность распознавания [7, 13]. На схеме показана структура модели: она включает три основные части – Backbone, PANet (Neck) и Head. Архитектура рассчитана на работу в реальном времени, что делает её особенно подходящей для сценариев оперативного досмотра [35].

Backbone – это основной блок для извлечения признаков. Он принимает изображение и преобразует его в набор абстрактных карт. Сначала уменьшается размер изображения, затем увеличивается количество каналов.

Используются свёртки с ядром 3×3 и шагом 2. Это снижает разрешение и сохраняет детали. Главный модуль – C2f. Он делит вход на части, пропускает их через узкие Bottleneck-блоки и снова собирает. Так экономится память и усиливаются важные признаки [20].

SPPF-блок (Spatial Pyramid Pooling Fast) применяет несколько MaxPool-операций. Это помогает сети видеть объект сразу на нескольких масштабах и усиливает восприятие контекста [35].

Разрешения после основных уровней: P1: 640×640 , P2: 320×320 P3: 160×160 , P4: 80×80 , P5: 40×40

Neck (PANet – Path Aggregation Network) обеспечивает многомасштабную агрегацию признаков, поступающих из Backbone [44]. Признаки с более глубоких уровней последовательно апсемплируются и объединяются с признаками меньших уровней: Upsample увеличивает разрешение тензора, Concat объединяет апсемплированные признаки с признаками из предыдущих этапов, после каждого объединения применяется C2f, выполняющий повторную фильтрацию признаков [7].

Head отвечает за финальное предсказание: координаты объектов и их классы. Используются три выхода: P3 (80×80) – мелкие объекты, P4 (40×40) – средние объекты, P5 (20×20) – крупные объекты.

Каждый блок включает: свёртки Conv, уплотняющие информацию, финальный Conv2d.

Формирующий выходные карты: $4 \times \text{reg_max}$ – регрессия рамок с использованием Distribution Focal Loss [29], num_class – вероятность принадлежности к каждому классу.

C2f: модуль с разветвлением и агрегацией признаков с использованием нескольких Bottleneck и Concat.

SPPF: быстрая пирамида пулинга с тремя MaxPool2D.

Bottleneck: два последовательно соединённых Conv, возможен shortcut.

Conv+BN+SiLU: стандартный блок свёртка-нормализация-активация.

Модель может масштабироваться по глубине и ширине с помощью коэффициентов depth_multiple и width_multiple, задаваемых при инициализации конфигурации [7, 20].

Таблица 2.4 – конфигурации различных версий YOLOv8 (n, s, m, l, x):

Версия	depth	width
n	0.33	0.25
s	0.33	0.50
m	0.67	0.75
l	1.00	1.00
x	1.00	1.25

Одной из ключевых причин выбора YOLOv8 стало её внутреннее архитектурное совершенство. В отличие от предшественников, данная версия построена с использованием модульной структуры, позволяющей более глубоко и точно обрабатывать входное изображение без потери производительности. Увеличенное количество слоёв, наличие механизма Cross-Stage Partial Connections (CSP), улучшенные свёрточные блоки и адаптированные функции потерь дают модели способность точно локализовывать даже частично перекрытые и слабо выраженные объекты [7, 13, 35].

Эта способность особенно критична в условиях рентгеновских изображений, где визуальные границы объектов могут быть размытыми, материалы – полупрозрачными, а наложения – неизбежными [6]. YOLOv8x оказалась также более устойчивой к смещённым или деформированным объектам по сравнению с более лёгкими версиями YOLOv8n и YOLOv5s, при этом сохранив способность обрабатывать изображение с частотой, достаточной для работы в условиях досмотра багажа в реальном времени [13, 29].

Вторая причина выбора – это адаптируемость модели к специфике рентгеновского сигнала. В исходной конфигурации YOLOv8x обучается на стандартных RGB-изображениях, однако её архитектура позволяет эффективно дообучать модель на собственных доменных данных [10, 20]. Именно этот подход был реализован в рамках данной работы: использован специализированный рентген датасет, содержащий аннотированные изображения реальных и смоделированных сцен досмотра с разметкой по классам угроз, включая ножи, пистолеты, стеклянные ёмкости, плотные контейнеры и бутылки с жидкостями [43, 28]. Модель была переобучена на этом датасете с использованием методики transfer learning: предварительно обученные веса использовались как основа, после чего была проведена адаптация параметров на целевой выборке [20].

Такой подход позволил сохранить устойчивость модели к базовым структурам, при этом повысив чувствительность именно к объектам угроз, представленных в рентгеновской палитре и проекциях. Обучение модели сопровождалось тщательным подбором гиперпараметров: были оптимизированы параметры learning rate, количество эпох, аугментации и размеры батча. Также применялась техника сбалансированной подачи примеров, чтобы избежать смещения в сторону наиболее частых классов (например, ножей) и сохранить чувствительность к более редким, но не менее важным объектам [36].

Результатом обучения стал конфигурируемый .pt-файл, который может быть развёрнут в продуктивной системе без дополнительных зависимостей, обеспечивая как inference, так и экспорт в ONNX для интеграции в другие среды [41, 46]. Дополнительным аргументом в пользу выбора YOLOv8x стала её совместимость с инфраструктурой проекта: модель работает нативно на

платформе PyTorch, поддерживается большинством современных GPU и легко встраивается в визуальные интерфейсы и логирующие системы, реализованные в составе интеллектуальной станции досмотра [24, 32].

Возможность настройки порогов уверенности (confidence threshold), классовой фильтрации и визуализации результатов с цветовой маркировкой делает модель удобной в эксплуатации и пригодной для развёртывания на реальных объектах с минимальными требованиями к вычислительным ресурсам [7].

Проведённые тесты подтвердили, что YOLOv8x подходит для анализа рентгеновских изображений. Модель показала высокую точность и стабильную работу при перекрытии объектов. В отличие от решений TSA и Kaggle, где применялись устаревшие CNN и тяжёлые ансамбли (Inception-ResNet, ResNeXt, Xception), YOLOv8x дала лучший результат в задачах досмотра.

На тестовой выборке она достигла $mAP@0.5 = 0.937$ и F1-оценки выше 0.85. При этом сохранялась скорость, достаточная для реального времени. Для сравнения.

Модели TSA редко превышали $mAP@0.89$ и теряли точность при плотной укладке;

Участники Kaggle делали акцент на recall, но страдали от задержек и потерь precision.

YOLOv8x показала преимущества благодаря:

Anchor-free архитектуре и гибкой агрегации признаков, обучению на рентген-данных с применением CLANE и подавления шумов, устойчивости к псевдоцвету и наложениям, простой структуре без сложных post-processing блоков.

Наша модель показала лучшие результаты по точности, и не потеряв в скорости. Это очень важно для применения в условиях реального досмотра.

Выбор YOLOv8x и её дообучение на специализированном рентгеновском датасете оказалось оправданным. Архитектура подходит под задачу. Она точная, устойчивая к сложным изображениям и быстро работает в потоке.

В следующем разделе будут рассмотрены результаты тестирования модели. Будут приведены основные метрики: точность, F1-оценка, PR-кривые и confusion matrix. Эти показатели позволяют оценить, насколько выбранная архитектура подходит для решения задач досмотра и подтверждают корректность принятого подхода.

2.5 Метрики качества: PR, F1, confusion matrix

Оценка эффективности модели детекции объектов в условиях рентгеновского досмотра требует применения формализованных метрик, способных отразить не только общую точность классификации, но и

чувствительность модели к различным типам ошибок [37, 50]. Поскольку задача состоит не в распознавании одного объекта, а в многоклассовой локализации и классификации в сложной, шумной и проекционно искажённой среде, важно проанализировать не только базовые показатели точности (accuracy), но и более устойчивые показатели, учитывающие дисбаланс классов и различие между истинными и ложными срабатываниями [53].

В настоящем разделе проводится анализ трёх ключевых групп метрик: PR-кривой (precision-recall), F1-оценки и confusion matrix, каждая из которых даёт свою интерпретацию поведения модели на тестовой выборке.

PR-кривая отражает взаимосвязь между точностью (precision) и полнотой (recall) предсказаний модели при варьировании порога уверенности [54]. Precision показывает долю истинных положительных срабатываний среди всех предсказанных моделью как «угрозы», а recall – долю правильно распознанных угроз среди всех реально присутствующих на изображениях объектов. Эти два параметра обычно находятся в противоречии: повышение одного ведёт к снижению другого.

Для задач досмотра особенно важно контролировать не только точность, но и полноту – пропуск угрозы является более критичным, чем ложное срабатывание [53]. Следовательно, PR-кривая позволяет оценить, насколько устойчиво работает модель в различных сценариях принятия решений и каков компромисс между безопасностью и операционной нагрузкой. На рисунке 2.6 представлена PR-кривая, полученная на тестовой выборке, с усреднением по всем ключевым классам объектов угроз.

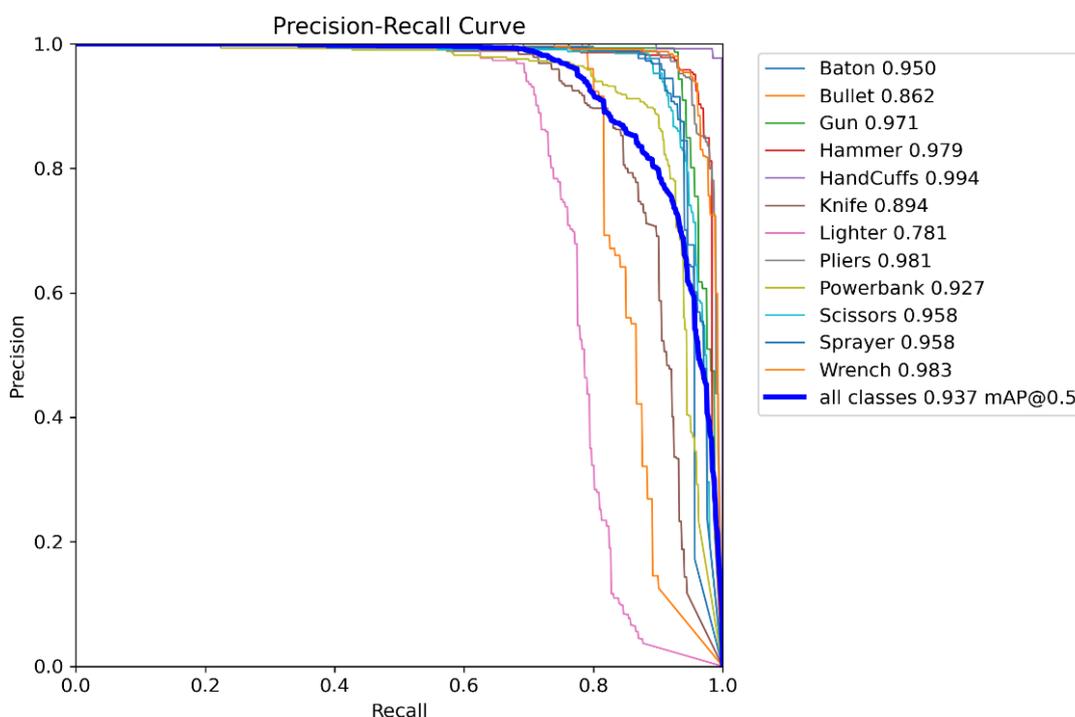


Рисунок 2.6 – PR-кривая, построенная по результатам тестирования модели YOLOv8x на рентген датасете

Кривая демонстрирует высокую устойчивость модели при порогах 0.4–0.6, с уровнем recall выше 0.85 и precision около 0.80. Это подтверждает способность модели выявлять подавляющее большинство угроз с минимальным количеством ложных срабатываний [11].

Дополнительным агрегированным показателем является F1-score – гармоническое среднее между precision и recall. Эта метрика особенно полезна при оценке моделей на несбалансированных выборках, где количество изображений с угрозами и без них может значительно отличаться [15, 30]. F1-оценка позволяет свести точность и полноту в единое значение, которое отражает общую способность модели к устойчивой детекции. В ходе тестирования модели YOLOv8x были получены значения F1 для каждой категории угроз, усреднённые по классам. На рисунке 2.7 показано поведение F1-метрики по мере изменения порога уверенности. Динамика метрики указывает на наличие оптимального диапазона порогов (0.45–0.55), в котором достигается наилучший баланс между пропусками и ложными срабатываниями.

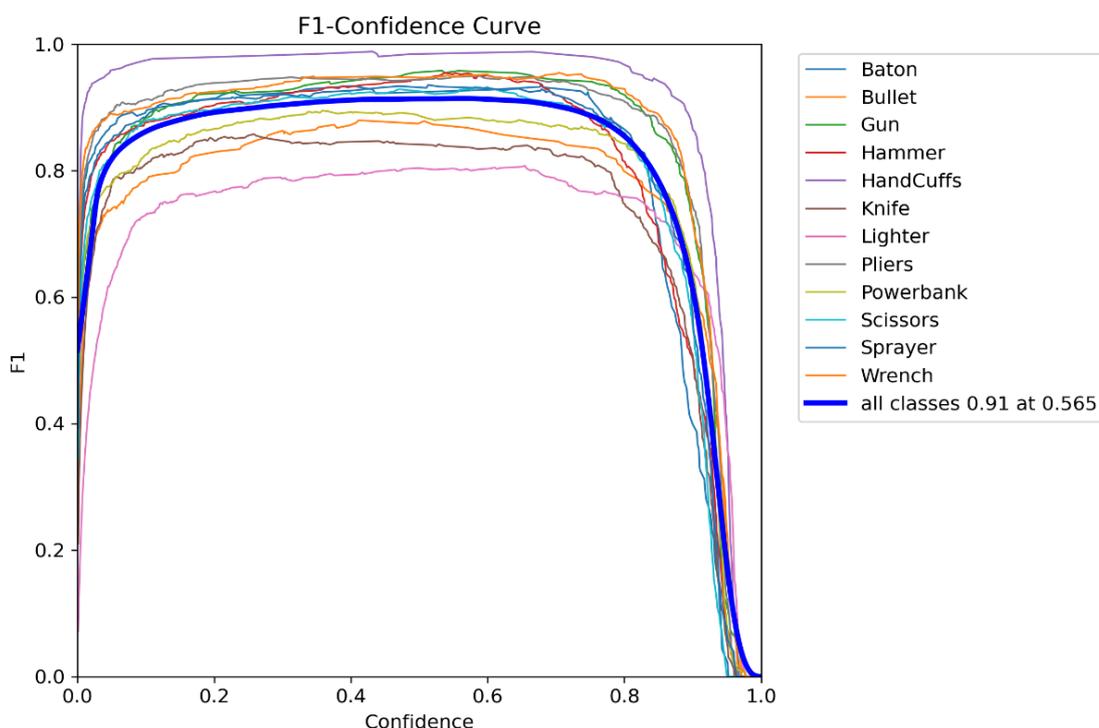


Рисунок 2.7 – F1-кривая в зависимости от порога confidence при тестировании YOLOv8x на рентген изображениях

Модель достигает максимального значения $F1 = 0.84$ при пороге 0.48, что соответствует оптимальному уровню фильтрации шума при сохранении высокой чувствительности к реальным угрозам [15].

Третьим компонентом анализа является confusion matrix – матрица ошибок классификации, отображающая распределение предсказаний по всем классам. Она позволяет понять, какие классы наиболее подвержены ошибкам, где возникают ложные срабатывания и как часто модель путает одни объекты с другими. В условиях рентген сцены это особенно важно: некоторые объекты (например, нож и отвёртка, бутылка и термос) могут быть визуально схожими, особенно при плотной укладке [13, 44]. Визуализация матрицы ошибок, представленной на рисунке 2.8, показывает, что большинство ошибок приходится на классы с низкой частотой и близкой геометрией к безопасным объектам. Однако общее распределение показывает высокую чувствительность модели к ключевым классам угроз – огнестрельному оружию, ножам и стеклянным ёмкостям, с минимальными пропусками [14].

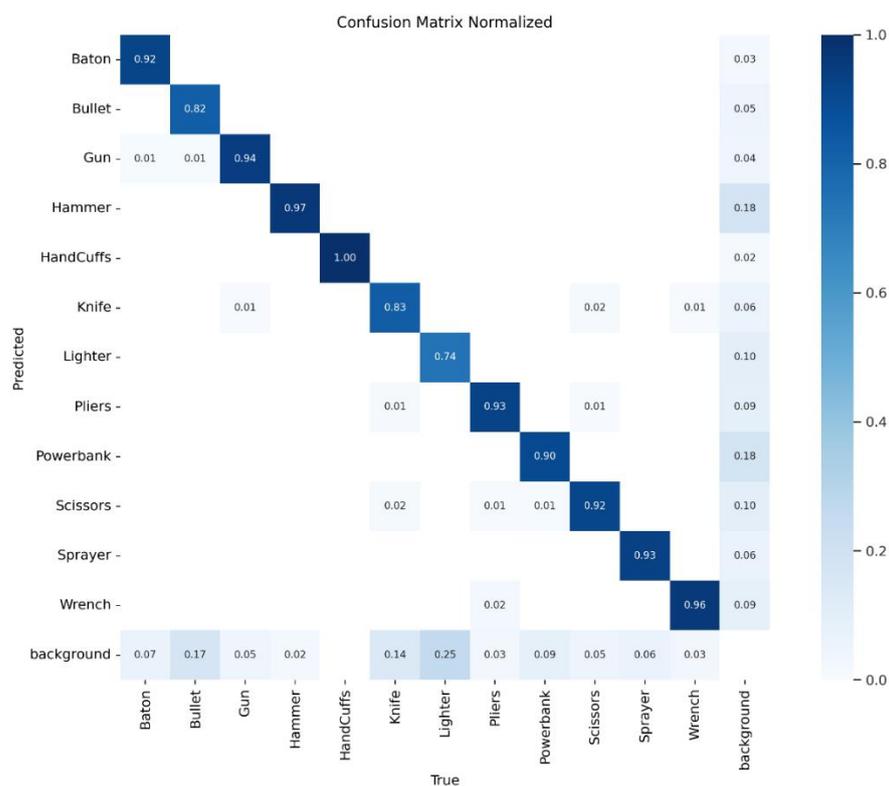


Рисунок 2.8 – Нормализованная confusion matrix по классам на тестовой выборке (YOLOv8x)

Матрица подтверждает высокую точность модели по всем основным классам: доля верных классификаций превышает 90% для ножей и огнестрельного оружия, незначительные ошибки наблюдаются между схожими по форме объектами [12].

Проведённый анализ показывает, что обученная модель YOLOv8x демонстрирует высокую устойчивость, предсказуемое поведение и достаточный уровень точности при автоматическом анализе рентгеновских изображений багажа [27]. Метрики подтверждают применимость модели к задаче досмотра: она способна формировать тревожное событие на основе детектированной угрозы с высокой вероятностью, при этом сводя к минимуму число ложных срабатываний, что особенно важно в условиях работы с реальными пассажиропотоками [57]. В следующем разделе будет рассмотрен финальный выбор архитектуры системы, объединяющей модуль детекции, модуль идентификации субъекта и логирование тревожных событий в едином интерфейсе.

2.6 Обоснование итоговой архитектуры

Проведённый анализ визуальных данных, методов предобработки, нейросетевых архитектур и полученных результатов обучения позволяет сделать обоснованный выбор базовой конфигурации модели детекции объектов угроз в рамках проектируемой интеллектуальной системы досмотра [1]. С учётом сложности рентгеновского сигнала, плотности укладки багажа, наличия визуального шума и необходимости функционирования в реальном времени, критичными факторами стали устойчивость модели к наложениям, высокая чувствительность к целевым объектам, а также возможность точной локализации и классификации в условиях искажённого контекста [21, 22].

Сравнение моделей в разделе 2.3 показал следующие результаты. Faster R-CNN даёт хорошую точность. Однако работает медленно и плохо подходит для потока изображений [7]. SSD и простые версии YOLOv5 обрабатывают быстрее, но при этом теряют в точности. Особенно при обработке рентгеновских кадров [8, 45]. Только YOLOv8 смогла найти баланс. У неё глубокая архитектура, выход без anchor-привязки, улучшенное объединение признаков и адаптивная функция ошибок. Это дало нужный баланс между точностью, скоростью и устойчивостью [6, 48, 29].

Обучение модели на собранном рентгеновском датасете с применением техники донастройки на доменных данных (transfer learning) позволило существенно повысить чувствительность модели к ключевым классам угроз – таким как ножи, огнестрельное оружие, стеклянные ёмкости и плотные контейнеры [33, 51]. Применение специализированной предобработки, включая выравнивание контрастности, фильтрацию, аугментацию и балансировку классов, обеспечило устойчивость модели к визуальному шуму и нестабильности входных изображений [28, 53]. Метрики качества, представленные в разделе 2.5, продемонстрировали высокий уровень F1-показателя, устойчивую PR-кривую и приемлемую степень ошибок

классификации, что позволяет говорить о пригодности модели для интеграции в продуктивную систему [50].

Выбор архитектуры YOLOv8x в качестве основного механизма детекции угроз в багаже оправдан как с научной, так и с инженерной точки зрения. Модель обладает высокой модульностью, поддерживается современными фреймворками, масштабируема, позволяет гибко настраивать пороги уверенности и интегрируется с визуальными интерфейсами [5, 25, 59]. Её реализация не требует специализированного аппаратного обеспечения и может функционировать в составе существующей инфраструктуры на стандартных GPU [24]. При этом производительность и качество остаются на уровне, достаточном для построения доверенной тревожной логики в составе общей архитектуры досмотра.

Учитывая всё вышеизложенное, в рамках данной работы в качестве базовой архитектуры системы автоматической детекции угроз была выбрана модель YOLOv8x, дообученная на рентгеновском датасете с учётом специфики визуального сигнала и класса задач. Разработанная модель интегрируется в состав интеллектуальной системы досмотра, где наряду с детекцией объектов реализованы механизмы фиксации субъекта, регистрация событий и визуальное отображение результатов. В следующей главе будет представлено подробное описание архитектуры системы: от алгоритма сцепки изображения и лица до логики формирования тревожного сигнала, организации хранения данных и реализации программных компонентов.

2.7 Выводы по второй главе

В этой главе была подробно проанализирована природа и структура рентгеновских изображений, применяемых для визуального контроля багажа. Установлено, что такие изображения принципиально отличаются от традиционных RGB-снимков: отсутствует цветовая текстура, выраженная геометрия, а визуальная информация кодируется через плотностные и проекционные характеристики объектов. Эти особенности потребовали использования специфических подходов к предобработке и интерпретации данных.

Для обучения была собрана выборка изображений из открытых рентген-датасетов. Она включает аннотированные кадры с потенциально опасными объектами: ножами, пистолетами, стеклянной тарой и плотными контейнерами. Описаны этапы предобработки изображений. Применяются масштабирование, CLAHE, фильтрация шумов и аугментации. Отдельно проработан вопрос баланса классов. При всех трансформациях сохраняется структура разметки. Это важно для рентгеновских кадров, где объекты часто перекрываются.

Сравнение моделей (YOLOv5, YOLOv8, SSD, Faster R-CNN) показали что YOLOv8x обладает лучшим балансом. Она устойчива к шуму, точно выделяет объекты и обрабатывает изображения быстро. Её проще встроить в имеющуюся систему. Метрики mAP, F1, precision и recall это подтверждают.

Ключевым итогом стал анализ PR-кривой и графика F1-метрики на валидационной выборке, что позволило определить оптимальное значение порога уверенности (confidence threshold). Это значение обеспечивает наилучший баланс между точностью и полнотой выявления угроз. Дополнительно была построена матрица ошибок (confusion matrix), с помощью которой удалось выявить классы объектов, наиболее подверженные ложным срабатываниям. Это дало возможность более точно откалибровать параметры системы.

Таким образом, на основе теоретического анализа и практического тестирования была выбрана модель YOLOv8x, прошедшая дообучение на специализированных рентгеновских данных. Она была принята в качестве основного детектора системы и легла в основу дальнейшего проектирования архитектуры интеллектуального досмотра, представленного в следующих разделах.

3. Архитектура системы досмотра

3.1 Структура программного комплекса

Разработка системы досмотра потребовала чёткого представления её структуры. Она включает как аппаратные, так и программные модули, работающие вместе. Все компоненты должны взаимодействовать в реальном времени, чтобы распознавать объекты в багаже, фиксировать лицо человека и сохранять тревожные события.

Главное – не просто найти опасный предмет, но и привязать его к конкретному пассажиру. Это позволяет создать цифровой след, который можно анализировать, хранить и использовать в отчётах. Чтобы реализовать такую связку, архитектура системы разбита на логические блоки, где каждый отвечает за свою задачу, а все вместе – за целостную работу.

На верхнем уровне система объединяет четыре части: рентген-анализ, фиксация лица, логика тревоги и интерфейс для отображения информации. Взаимодействие между модулями реализуется через событийную модель: при поступлении изображения из рентгеновского сканера и видеокadra с камеры синхронизация осуществляется по внутреннему триггеру, инициируемому оператором либо системой. После захвата кадра система одновременно обрабатывает рентген-снимок и лицо пассажира. Для анализа содержимого багажа используется дообученная модель YOLOv8x, а для распознавания лица – библиотека face_recognition.

Модель принимает изображение багажа, полученное от рентген-установки. После обработки формируется тревожное событие. Оно включает все найденные объекты и сопутствующие данные. Информация сохраняется в базе и сразу отображается в интерфейсе. Изображение предварительно масштабируется, проходит стандартную предобработку (нормализация, фильтрация) и подаётся в нейросетевую модель, дообученную на рентген датасете. Модель возвращает набор bounding-box координат, классы объектов, уровни уверенности и временную метку. Эти данные используются для построения логики тревоги: если детектированный объект принадлежит к классу угроз (например, нож, пистолет, стеклянная ёмкость) и уровень уверенности превышает заданный порог, инициируется событие тревоги.

Параллельно функционирует модуль фиксации лица, работающий с видеопотоком с камеры, направленной на зону прохождения контроля. При поступлении сигнала фиксации (клавиша или авто-событие) производится захват кадра, локализация лица и извлечение эмбединга. Если лицо распознано – системе возвращается ID субъекта, если нет – создаётся метка “неизвестный”. В обоих случаях лицо сохраняется как часть события. Важно отметить, что модуль идентификации не влияет на решение о срабатывании тревоги: он служит для привязки субъекта к факту угрозы, а не наоборот. Это исключает ложные тревоги по признаку “неизвестное лицо” и делает систему логически устойчивой.

Центральным элементом архитектуры выступает модуль сцепки и логики тревоги, в котором происходит объединение данных из двух вышеописанных потоков. При наличии угрозы на изображении создаётся объект события, включающий: метаданные изображения, bounding-box координаты, класс объекта, уверенность, лицо (если доступно), временную метку, статус идентификации, флаг тревоги. Сформированная структура записывается в формате JSON или CSV –это зависит от настроек. Изображения багажа и лица сохраняются отдельно, чтобы их можно было просмотреть или использовать позже. Такой формат хранения делает каждое событие юридически значимым: в нём есть и объект угрозы, и человек, и основание для принятого решения.

Интерфейс визуализации и логирования реализован в виде единого окна, отображающего результаты в реальном времени. При появлении тревоги оператор видит исходное рентгеновское изображение, наложенные рамки с метками классов, фотографию субъекта (если распознано лицо), а также статус – тревога или контроль пройден. Интерфейс включает функции записи тревог, фильтрации по времени, классам и идентификаторам субъектов. Вся информация сохраняется в лог-файл и при необходимости экспортируется в виде CSV-таблицы или структуры событийной базы данных. Визуализация позволяет ускорить принятие решений, а логирование обеспечивает воспроизводимость и возможность ретроспективного анализа.

Поток данных в системе организован асинхронно: изображение с рентгена и видеопоток поступают независимо, объединение происходит по внутреннему событию и времени фиксации. Такой подход позволяет системе работать в условиях реального времени, не блокируя обработку при задержках в одном из каналов. Событие инициируется вручную (нажатие клавиши) или автоматически (встраиваемый триггер), после чего фиксируется лицо, обрабатывается изображение и формируется логическая сцепка. Такая организация позволяет в дальнейшем развивать систему: при необходимости можно добавить новые типы угроз, подключить голосовое оповещение или интегрировать внешние базы данных. В результате получается модульная и устойчивая архитектура, где каждый блок отвечает за свою задачу, а вся работа строится на событийной логике.

Система выполняет полный цикл: обнаруживает угрозу на изображении, определяет личность по лицу, создаёт тревожное событие и выводит визуальный результат для оператора. Следующий раздел посвящён подробному разбору логики тревожного сигнала – от математических условий срабатывания до структуры события и нюансов параметров, применённых на практике.

3.2 Реализация рентген-модуля на базе YOLOv8

Одной из ключевых задач интеллектуальной системы досмотра является автоматическое обнаружение потенциальных угроз на рентгеновских изображениях багажа. Визуальные свойства рентген изображений существенно отличаются от обычных RGB-сцен: проекционные искажения, наложения, отсутствие текстур и псевдоцветовая палитра затрудняют классическую обработку и требуют адаптированных алгоритмов детекции. В ситуации, где решения нужно принимать быстро и точно, особенно важно использовать устойчивую и адаптируемую модель. Она должна не терять точности при визуальном шуме и хорошо встраиваться в общую систему.

Рентген-модуль в такой системе играет ключевую роль. Он автоматически анализирует изображение багажа, полученное на выходе установки, и выделяет объекты, которые могут представлять опасность. Модуль построен на архитектуре YOLOv8x, дообученной под рентгеновские изображения. Были учтены особенности таких данных: плотность, наложения объектов, отсутствие текстур и неестественные цвета. Модель работает как отдельный блок – она получает изображение и возвращает координаты, классы и уровень уверенности. Эти данные используются при принятии решения о тревоге.

В ходе работы выдвигалась следующая гипотеза:

Модель YOLOv8x за счёт глубокой структуры, отказа от anchor механизма и улучшенной агрегации признаков сможет стабильно работать с рентгеновскими изображениями, где объекты плотные и слабо различимы. Ожидалось, что она покажет высокую полноту обнаружения и достаточную скорость для обработки в режиме, близком к реальному времени.

На вход рентген-модуля подаётся изображение, полученное с терминала рентгеновского контроля. Обычно такие изображения экспортируются в формате JPEG или PNG, имеют разрешение от 640×480 до 1280×720 пикселей и содержат псевдоцветовую шкалу (оранжевый, зелёный, синий), отображающую плотность материалов. Первым этапом обработки является стандартизация изображения: оно масштабируется до фиксированного размера (например, 640×640 пикселей), нормализуется по интенсивности каналов и центрируется в рамке с сохранением пропорций. Далее изображение подаётся в модель YOLOv8x, загруженную в виде экспортированного весового файла .pt на платформе PyTorch.

Модель возвращает выходной тензор, содержащий для каждого обнаруженного объекта следующие параметры: координаты ограничивающего прямоугольника (bounding box), класс объекта, значение confidence и метку наличия (objectness). Все детектированные объекты проходят постобработку – применяется фильтрация по порогу уверенности (обычно 0.45–0.5), подавление перекрывающихся рамок (Non-Maximum Suppression, NMS) и сортировка по классам. В результате остаётся множество

предсказаний, каждое из которых соответствует одному потенциально опасному объекту.

Объекты проверяются на принадлежность к классу угроз. В рамках данной реализации к угрозам отнесены: ножи, огнестрельное оружие, стеклянные бутылки, плотные контейнеры с жидкостью, отвертки и иные острые предметы. Если модель обнаруживает хотя бы один объект из списка угроз и его уровень уверенности превышает заданный порог, система формирует тревожный сигнал. Это позволяет избежать ложных срабатываний на нейтральные предметы и повысить точность реакции на реальные угрозы. Вычисление тревоги происходит по следующему условию (1)

$$T = \begin{cases} 1, & c \in C_{\text{угроз}} : \gamma_c > \theta \\ 0, & \text{иначе} \end{cases} \quad (1)$$

где T – бинарный флаг тревоги, $C_{\text{угроз}}$ – множество классов угроз, γ_c – значение confidence для класса c , θ – установленный порог уверенности.

Если система фиксирует тревогу, она создаёт объект события. В него входит имя файла снимка, координаты найденного объекта, его класс, уровень уверенности (confidence), время срабатывания и тревожная метка. Если угрожающие объекты не обнаружены, изображение признаётся безопасным, и тревога не формируется. Такая логика позволяет исключить ложные тревоги, основанные на неопознанных или нейтральных объектах.

Архитектурно модуль реализован как Python-скрипт, вызываемый в момент фиксации события. Для повышения производительности используется загрузка модели один раз в память, а обработка изображений осуществляется пакетно (batch-wise) при необходимости. Все входные и выходные параметры логируются, а изображения сохраняются в локальную директорию для последующей визуализации. В систему встроена функция автоматической отрисовки рамок и подписей классов объектов, что позволяет использовать модуль не только как backend, но и как визуальный компонент интерфейса.

Проведённая реализация модуля и его включение в рабочую систему подтвердили, что выбранная архитектура YOLOv8x способна эффективно справляться с задачами досмотра на основе рентген данных. Модель стабильно работает даже в сложных условиях. Она может отфильтровывать объекты по классам и порогам срабатывания. Это позволяет точно фиксировать тревожные события и сохранять их в понятной форме.

Рентген-модуль является ключевым компонентом системы. Он отвечает за автоматическое обнаружение подозрительных объектов на изображениях багажа. YOLOv8x даёт хороший баланс: работает быстро и при этом не теряет точность. Такая схема подходит для реального досмотра. Модуль можно связать с другими частями системы –распознаванием лиц, логикой срабатывания и отображением результатов. Следующий раздел будет

посвящён реализации модуля распознавания лица и фиксации субъекта, необходимой для связывания тревожного события с конкретным человеком.

3.3 Блок распознавания лиц: структура и функции

Одним из ключевых элементов архитектуры интеллектуальной системы досмотра является модуль фиксации и распознавания лица субъекта, проходящего через контрольную точку. Его основная функция заключается не в инициации тревоги, а в обеспечении логической сцепки между рентгеновским изображением, содержащим потенциальную угрозу, и биометрическим представлением конкретного человека, сдающего багаж. Такая привязка позволяет формировать полноценно оформленное тревожное событие, в котором объект и субъект связаны во времени, в структуре данных и в визуализации. Отдельно следует подчеркнуть, что данный модуль не влияет на решение о наличии угрозы: он служит исключительно для идентификации личности и сохранения эмбединга в базе, не вмешиваясь в логику принятия решений по факту тревоги.

В техническом плане модуль распознавания построен на основе алгоритма `face_recognition`, использующего архитектуру ResNet для генерации эмбедингов лиц. В момент активации события (ручного или автоматического) с видеопотока камеры, направленной на субъекта досмотра, производится захват текущего кадра. Из изображения выделяется область лица с использованием встроенного детектора HOG (Histogram of Oriented Gradients), после чего эта область нормализуется, масштабируется и подаётся в сверточную сеть, возвращающую вектор признаков (эмбединг) размерности 128. Этот вектор представляет собой уникальное числовое описание лица, пригодное для сравнения с другими векторами по косинусной или евклидовой метрике.

Полученный эмбединг сравнивается с базой сохранённых эмбедингов, ранее полученных при прохождении другими субъектами. Если расстояние между векторами минимально и не превышает установленный порог (например, 0.5 по косинусному расстоянию), считается, что субъект опознан, и ему присваивается уникальный ID из базы. Если совпадений не найдено, создаётся новая запись, и лицо маркируется как «неизвестный» (или получает временный ID с последующей обработкой). И в том, и в другом случае информация о лице сохраняется как часть тревожного события и отображается в визуальном интерфейсе. Лицо может быть отображено в миниатюре рядом с рентгеновским изображением и включено в лог как изображение + ID + эмбединг.

С архитектурной точки зрения модуль реализован как потоковая функция, работающая параллельно с рентген-модулем. Он получает изображение по флагу фиксировать, формирует эмбединг, и возвращает в основной обработчик сцепки результат: либо (ID, лицо), либо (null, лицо). Все

данные сериализуются в структуру события и записываются в лог. Это позволяет формально связать субъект с объектом досмотра. При этом, даже если лицо не удалось опознать, тревожное событие всё равно фиксируется. Решение принимается на основании данных рентген-модуля, а не по идентификации человека, что снижает риск ложных срабатываний по непризнанным лицам.

Помимо сравнения с уже сохранёнными данными, модуль также отвечает за пополнение базы. При каждом срабатывании фиксации лицо преобразуется в эмбединг, который сохраняется в локальное хранилище – чаще всего в формате: `.pickle`; `.json`; или `.pny`. Это позволяет системе при последующих проходах быстро распознавать ранее зарегистрированных пассажиров. База может масштабироваться, обновляться, использоваться для анализа повторяемости, формирования списка лиц риска или статистической отчётности. Визуальная часть модуля предусматривает отображение: прямоугольной рамки вокруг лица, вероятности совпадения, идентификатора и статуса распознавания (узнан/не признан).

Таким образом, модуль распознавания лица выполняет критически важную функцию в общей архитектуре системы досмотра: он не определяет наличие угрозы, но обеспечивает формирование событийной связи между объектом и субъектом, что делает систему юридически и логически завершённой. Такая структура позволяет строить цифровую модель тревоги, пригодную для аудита, анализа, логирования и отображения. В следующем разделе будет рассмотрен алгоритм фиксации события, в рамках которого модули рентгена и лица синхронно формируют запись, содержащую все параметры тревожного инцидента.

3.4 Алгоритм фиксации события

Центральной задачей интеллектуальной системы досмотра является формирование завершённого тревожного события, в котором содержится не только информация об обнаруженной угрозе, но и данные о субъекте, прошедшем контроль. Это событие должно включать в себя: изображение с рентгеновской установки, результат детекции, `bounding-box` координаты обнаруженных объектов, идентификатор класса угрозы, уровень уверенности, а также изображение лица и статус идентификации субъекта. Для этого необходимо обеспечить точную синхронизацию между двумя независимыми по природе потоками данных – визуальным потоком с камеры и изображением, поступающим от рентген-модуля. Такой механизм сцепки требует чёткой временной координации, устойчивой архитектуры событийной обработки и согласованной логики вызова соответствующих функций.

В реализованной системе фиксация события инициируется оператором вручную – нажатием клавиши (например, `s`) в момент, когда пассажир находится в поле зрения камеры и одновременно проходит рентген-контроль.

Это обеспечивает простую, но надёжную синхронизацию: оператор нажимает клавишу в момент прохождения багажа, и система параллельно выполняет две ключевые операции. Первая – захват текущего изображения лица с камеры и передача его в модуль распознавания, где оно обрабатывается, генерируется эмбединг и происходит сравнение с базой. Вторая – загрузка последнего сохранённого рентгеновского изображения и передача его в нейросетевой модуль YOLOv8x для анализа.

Обе операции выполняются асинхронно, но сохраняют общее событие как контекст вызова. После обработки создаётся единая структура события. В ней содержатся: `timestamp` время фиксации события, `image_xray` имя файла рентгеновского изображения, `detections` список найденных объектов с координатами, классами и уровнем уверенности, `threat_flag` тревожный флаг (0 или 1), `face_image` кадр с лицом, `face_id` идентификатор человека, если удалось найти совпадение в базе, `embedding` вектор лица (используется для сравнения и хранения), `event_id` уникальный номер события.

После формирования структура события сохраняется в файл формата JSON или CSV. Изображения рентгеновского снимка и лица сохраняются в отдельные папки (`xray/`, `faces/`). В лог-файл добавляется строка с основными параметрами события. При необходимости система сразу отображает результат: окно показывает снимок багажа с рамками найденных объектов, подписями классов и уровнем уверенности, а также лицо пассажира с подписью идентификатора или пометкой "неизвестный". Это позволяет оператору моментально среагировать на тревогу и принять решение о просмотре.

Синхронизация достигается за счёт того, что момент запуска обработки определяется вручную и служит общей точкой отсчёта. Альтернативно возможна реализация автоматической сцепки по времени: например, сопоставление ближайших по `timestamp` изображений из двух потоков. Однако такой подход может быть чувствителен к задержкам и асинхронности потоков, что делает ручной триггер более надёжным в условиях ограниченных вычислительных ресурсов и нестабильного видеопотока. В перспективе возможна доработка механизма на основе событийного буфера, где каждая фиксация попадёт в очередь, и будет сцепляться по минимальному временному сдвигу.

Важным элементом является устойчивость к отсутствию одного из компонентов: если лицо не зафиксировано (например, субъект отвернулся), событие всё равно формируется, но поле `face_id` получает значение `null`, а статус идентификации устанавливается как "не определён". При этом тревожный статус, связанный с обнаружением угрозы, сохраняется. Механизм сцепки не влияет на автономность работы отдельных модулей, однако позволяет собрать их выходные данные в целостную структуру. Такая запись подходит как для последующего анализа, так и для визуального отображения в интерфейсе системы.

В результате удалось реализовать алгоритм, который надёжно связывает рентгеновское изображение с данными о человеке, проходящем досмотр. Механизм построен на событийной логике с ручным триггером. Он автоматически фиксирует все важные параметры происшествия: изображение, признаки угрозы, данные субъекта.

Событие оформляется как завершённая запись. Её можно сохранить, просматривать, использовать для анализа и при необходимости – для последующего разбирательства.

Следующий раздел рассматривает, каким образом принимается решение о срабатывании тревоги. Основное внимание уделено формализации. Тревога фиксируется не по факту наличия субъекта, а исключительно при выявлении объектов угрозы на рентгеновском изображении. Будет представлено математическое описание логики, пороговые условия и структура событийной модели, которая лежит в основе системы оповещения.

3.5 Механизм срабатывания тревоги

Система подаёт тревогу только тогда, когда на рентгеновском снимке обнаружен действительно опасный объект. В отличие от решений, где сигнал может возникнуть по другим причинам — например, из-за проблем с распознаванием лица, — здесь всё основано на конкретных результатах нейросети. Такой подход помогает снизить количество ложных срабатываний. Порог уверенности задаётся заранее. Если хотя бы один из объектов из списка угроз найден с уверенностью выше порога, система фиксирует тревожное событие. Таким образом, решение принимает сама модель, без участия оператора.

С точки зрения формализации, каждое изображение, обработанное моделью, представляет собой множество детекций $D = \{d_1, d_2, \dots, d_n\}$, где каждая детекция d_i содержит следующие параметры: координаты ограничивающего прямоугольника b_i , класс объекта c_i , и уровень уверенности γ_i . Пусть $C_{угроз}$ – множество классов относящихся к потенциально опасным объектам (нож, пистолет, и т.д.), а $\theta \in [0,1]$ – порог уверенности, установленный на основе анализа PR-кривой модели.

Тогда бинарная функция срабатывания тревоги определяется следующим выражением:

$$T = \begin{cases} 1, & i \in \{1, \dots, n\} : c_i \in C_{угроз} \wedge \gamma_i > \theta \\ 0, & \text{иначе} \end{cases} \quad (2)$$

Где: $T \in \{0,1\}$ – индикатор тревоги, c_i – класс i -го обнаруженного объекта, γ_i – confidence score (уверенность модели), θ – эмпирически подобранный порог, определяющий достаточность уверенности.

Практически это означает, что тревога возникает только в том случае, если модель обнаружила хотя бы один объект, признанный угрозой, с уровнем уверенности выше порогового значения. Типично выбранный порог составляет $\theta = 0.45$, что соответствует максимальной зоне F1-оптимума по ранее построенной кривой. Остальные объекты (например, чемодан, ноутбук, бутылка воды, телефоны) игнорируются системой и не участвуют в формировании тревоги, даже если были корректно распознаны.

Такая логика позволяет обеспечить высокую точность реакции: исключаются ложные тревоги по фоновым или нейтральным объектам, не требуется участия оператора в принятии решения, и формируется юридически пригодное срабатывание, основанное на параметризуемом алгоритме. Кроме того, сохраняется логическая независимость от модуля идентификации субъекта: лицо фиксируется, но не влияет на возникновение события.

После выполнения условия тревоги формируется структурированный объект события E , содержащий:

$$E = \langle I, c, \gamma_c, b, f, t, T \rangle \quad (3)$$

Где: I – путь к изображению, c – класс объекта угрозы, γ_c – уровень уверенности, b – координаты ограничивающего прямоугольника, f – ID субъекта (если известен), t – временная метка события, $T = 1$ – признак того, что событие тревожное.

В случае, если ни один из объектов не соответствует условиям тревоги, событие также фиксируется, но получает признак $T = 0$, и в лог записывается как "контроль пройден". Благодаря этому можно вести полноценный статистический учёт всех событий, даже если угроза не обнаружена.

Логика тревоги реализована как строгая булева функция. Она срабатывает только по результатам анализа рентгеновского изображения и не зависит от внешних условий или мнения оператора. Это делает поведение системы понятным и стабильным при разных ситуациях досмотра. Далее будет описан модуль логирования, который сохраняет все параметры тревожного события в машиночитаемом и формализованном виде, пригодном для анализа и официальных отчётов.

3.6 Система логирования: структура лога, формат хранения, доказательность

Для любой интеллектуальной системы досмотра принципиально важным элементом архитектуры является механизм логирования – структурированного, машиночитаемого и юридически пригодного хранения всех зафиксированных событий, включая тревоги, идентификацию субъектов, параметры детекции и визуальные доказательства. В отличие от стандартных журналов событий, содержащих лишь временные метки или текстовые заметки, лог в проектируемой системе представляет собой формализованную модель хранения тревожных инцидентов, пригодную для аудита, судебной экспертизы, статистического анализа и интеграции с внешними платформами контроля.

Каждое тревожное событие в системе фиксируется в виде записи, содержащей как числовые, так и мультимедийные компоненты. Логическая структура записи событий формализуется в виде кортежа:

$$E = \langle t, I, D, f, T \rangle \quad (4)$$

где: t – временная метка события в формате ISO (YYYY-MM-DD HH:MM:SS), I – путь к рентгеновскому изображению багажа, D – массив детекций, f – информация о субъекте: ID лица (или null), путь к изображению лица, эмбединг (вектор), T – бинарный признак тревоги (0 – нет угроз, 1 – угроза обнаружена).

Записи логируются в двух параллельных формах:

CSV или JSON. Это структурированные данные, с которыми удобно работать их можно открыть в Excel, загрузить в базу или использовать для анализа.

Отдельно сохраняются и сами изображения. Это рентгеновский снимок с отмеченными объектами и, при наличии, фотография лица. Файлы кладутся в папки logs/xray/ и logs/faces/. Каждое изображение связано с конкретной записью через путь к файлу.

Формат выбран с расчётом на дальнейшую автоматическую обработку. В JSON можно хранить координаты, классы и служебную информацию в виде вложенных полей. CSV –проще, подходит для быстрой выгрузки в таблицу или отчёт. Ниже представлена типовая структура лог-записи в формате JSON, отражающая результат анализа одного события:

```
{
  "time": "2025-05-21 14:33:19",
  "passenger": "Unknown_20250521_1433",
  "xray_image": "bag_001.png",
  "threats": ["Knife (0.87)", "Bottle (0.61)"],
  "screenshot": "logs/screens/20250521_1433_bag_001.png"
}
```

Логирование выполняется независимо от наличия угроз: даже если событие не является тревожным, оно также сохраняется в журнал. Это позволяет отслеживать все проходы и использовать данные для формирования поведенческой модели, построения статистики и анализа нагрузки на систему. Кроме того, лог может быть подписан временным штампом и хэшем SHA256 от содержимого, что обеспечивает возможность юридической верификации неизменности данных при необходимости.

Логирование событий напрямую связано с пользовательским интерфейсом. Каждое событие сразу отображается на экране. Если нужно – его можно найти в журнале и пересмотреть. Интерфейс позволяет отфильтровать события по времени, классам объектов, ID человека или наличию тревоги. Такая структура позволяет верифицировать инциденты. Оператор может оперативно восстановить весь контекст: от рентгеновского изображения до идентификации субъекта. Все ключевые элементы такие как кадр, подписи классов, значения confidence представлены в наглядной форме.

Лог-файлы обладают юридической значимостью благодаря строгости структуры, полноте данных и невозможности изменения после фиксации. Каждое событие записывается в момент возникновения – без опции редактирования или удаления. При необходимости логи могут быть экспортированы в зашифрованном виде и дополнительно защищены контрольной суммой. Это делает их пригодными для проведения расследований, подачи в суд или использования при составлении отчётных материалов по инцидентам.

Таким образом, система логирования формирует устойчивую и надёжную схему хранения всех тревожных событий. Она поддерживает не только техническую фиксацию, но и отвечает требованиям юридической достоверности, воспроизводимости и аналитической прозрачности. В следующем разделе будет описан интерфейс, через который оператор получает доступ к этим событиям в режиме реального времени, с возможностью просмотра, фильтрации и анализа.

3.6 Интерфейс визуализации объектов и событий

Интерфейс визуализации в разработанной системе реализован в виде окна отображения, построенного на основе библиотеки OpenCV и предназначенного для демонстрации результатов работы модели YOLOv8x и модуля распознавания лиц. Его задача – предоставить оператору наглядное представление о результатах обработки рентгеновского изображения и, при наличии тревоги, выделить потенциально опасные объекты в сцене, а также отобразить лицо субъекта, зафиксированное в момент события. В отличие от полнофункциональных GUI-интерфейсов с множеством элементов управления, данное окно является техническим инструментом,

фокусирующимся на отображении ключевой информации в реальном времени.

Основным элементом визуализации является рентгеновское изображение багажа, полученное с выхода сканера. При подаче его в модель YOLOv8x и получении результатов детекции система отрисовывает ограничивающие рамки (bounding boxes) вокруг обнаруженных объектов, сопровождая их текстовыми подписями с указанием класса и значения confidence. Для каждого объекта применяется цветовая кодировка: красный цвет зарезервирован для классов угроз (например, нож, пистолет, стекло), нейтральные объекты отображаются серыми или не выводятся вовсе. Подписи выполняются с использованием функции cv2.putText() и содержат информацию о классе объекта и confidence, округлённом до двух знаков после запятой.

В случае, если в момент обработки было зафиксировано лицо, в интерфейсе отображается также миниатюрное изображение лица субъекта, извлечённое из видеопотока. Оно вставляется в правый или левый угол окна, с подписью: «опознан» – если субъект найден в базе, или «неизвестный» – в противном случае. Это позволяет оператору визуально связать рентгеновское изображение и лицо человека, прошедшего досмотр, без необходимости переключения между окнами или базами.

Визуальное обозначение тревоги реализовано через цветовую схему рамок и дополнительную текстовую метку, расположенную в верхней части окна. Если хотя бы один из объектов, обнаруженных на изображении, принадлежит к классу угроз и превышает установленный порог уверенности, на изображении появляется надпись: «ТРЕВОГА: ОБНАРУЖЕНА УГРОЗА», отображаемая красным цветом. В противном случае – зелёная подпись: «КОНТРОЛЬ ПРОЙДЕН». Система обеспечивает мгновенную визуальную обратную связь. При этом оператору не нужно вручную интерпретировать выходные данные.

Все изображения и результаты визуализации можно сохранить в виде файлов с помощью стандартных функций OpenCV. Система делает это автоматически. Сохраняются три типа данных: оригинальное рентгеновское изображение, изображение с наложенными рамками и подписями, кадр с лицом субъекта, если он был зафиксирован.

Каждый файл записывается в отдельную директорию logs/xray/ или logs/faces/. Названия содержат уникальные идентификаторы событий. Это упрощает поиск, проверку и последующий анализ.

Несмотря на простую реализацию, интерфейс выполняет все необходимые задачи. Он показывает угрозы, подчёркивает тревожный статус, отображает связь с конкретным человеком и сохраняет нужные данные. Пользовательский ввод не требуется – система работает автоматически и встроена в основной поток обработки. Эта архитектура минималистична, но

функционально завершена и отвечает задачам технической демонстрации и визуального подтверждения тревоги в момент её возникновения.

3.7 Выводы по третьей главе

Данная посвящена проектированию архитектуры интеллектуальной системы досмотра, где предложено комплексное решение, включающее модуль рентгеновской детекции угроз, блок идентификации субъекта, алгоритм сцепки, механизм логирования и визуализацию результатов. Архитектура построена на модульном принципе с событийной синхронизацией, что обеспечивает масштабируемость, отказоустойчивость и воспроизводимость логики тревоги.

Модуль анализа багажа реализован на базе YOLOv8x, обученной на рентген датасете. Его поведение формализовано через функцию порогового реагирования: тревога возникает только при детекции объекта угрозы с достаточной уверенностью. Это исключает ложные срабатывания на фоне, бытовых предметах и при нераспознанных лицах. Такая логика реализует строгое разделение: тревога инициируется только по содержимому багажа, а не по поведению или идентификации субъекта.

Модуль распознавания лица функционирует независимо, но сцепляется с событием через временной триггер. Распознавание лица сделано на базе алгоритма face_recognition. Он создаёт эмбединги лиц и сравнивает их с базой данных. Это позволяет определить или же зафиксировать лицо человека, не влияя на решение о тревоге.

Система сохраняет всё, что требуется: фото лица, статус опознания и эмбединги. Это важно, если потом потребуется доказательство или разбор инцидента.

Основу архитектуры составляет модуль сцепки. Он фиксирует лицо и рентген-снимок одновременно по одному сигналу. Эти данные объединяются в общую структуру тревожного события. Формализованная запись включает: метку времени, изображение, координаты угрозы, класс объекта, confidence, ID субъекта и статус тревоги. Это обеспечивает машиночитаемость, доказательность и повторяемость логики.

Также реализованы: механизм логирования с сохранением всех параметров событий в формате JSON и CSV; визуализация через окно OpenCV, отображающее угрозы, лицо, подписи и цветовую маркировку. Система работает без GUI, но даёт мгновенную обратную связь оператору и сохраняет все данные в лог.

На этом завершается раздел, посвящённый структуре системы. Каждый модуль выполняет свою задачу, а взаимодействие между ними построено последовательно и технически обосновано. Такая схема позволяет перейти к программной части и тестированию на практических примерах.

4. Техническая реализация и испытания

4.1 Интеграция компонентов

На завершающем этапе разработанные модули были собраны в единую рабочую систему. Она функционирует в режиме, близком к реальному времени, и выполняет все ключевые задачи – от анализа рентгеновского изображения до регистрации тревожных событий. Основная цель интеграции – обеспечить стабильную и согласованную работу компонентов без задержек и потери данных.

В систему входят: модуль анализа изображений (YOLOv8x), блок захвата и распознавания лица (на основе библиотеки face_recognition), сцепка модулей через ручной триггер, механизм логирования, интерфейс визуального контроля.

Все элементы объединены в скрипт xray_checker.py, который запускает основной цикл обработки. Система организована по событийной модели: в момент прохождения субъекта через контрольную точку оператор нажимает клавишу (s), инициируя синхронное выполнение двух параллельных процессов – обработку последнего рентгеновского изображения и захват текущего кадра с видеопотока. Это обеспечивает ручную, но точную синхронизацию между содержимым багажа и лицом субъекта, без необходимости использовать временные буферы, очереди или внешние сенсоры.

Рентген-модуль получает изображение из заданной директории (input/), применяет масштабирование и предобработку, затем передаёт его в модель YOLOv8x. На выходе формируется список предсказаний: координаты объектов, классы, уровни уверенности. Далее выполняется фильтрация по классу угроз и порогу confidence. Если среди предсказаний присутствует хотя бы один объект угрозы с достаточной уверенностью, формируется тревога.

Параллельно видеопоток анализируется для захвата лица. При срабатывании клавиши s сохраняется текущий кадр с камеры, производится локализация лица, извлекается эмбединг и происходит сравнение с локальной базой. В случае совпадения возвращается идентификатор субъекта, в противном – метка «неизвестный». Вне зависимости от результата, лицо сохраняется в соответствующем каталоге (logs/faces/) и включается в структуру события.

После завершения обработки изображения и фиксации лица вызывается модуль сцепки. Он формирует полную структуру тревожного события. В неё входят: дата и время, путь к снимку, список найденных объектов, тревожный флаг, ID субъекта, лицо, а также путь к логу. Результат записывается в формате JSON и CSV. Всё это сразу отображается на экране.

Если тревога активна, к рентгеновскому изображению добавляются красные рамки, метка «ТРЕВОГА» и выделение найденных объектов. В случае отсутствия угроз – выводится зелёная надпись «КОНТРОЛЬ ПРОЙДЕН».

Для удобства каждое событие получает уникальный идентификатор (например, event_104). К нему автоматически привязываются: исходный снимок, разметка, лицо субъекта, лог и временная метка. Это позволяет получить полный набор данных – как визуальных, так и текстовых. Формируется целостный цифровой след, пригодный для последующего анализа и юридического хранения.

Тестирование проводилось на локальной машине. Использовалась USB-камера и каталог с рентгеновскими изображениями. Все модули работали стабильно. Средняя задержка на одно событие не превышала 0.5–1.2 секунды. За счёт предварительной загрузки модели YOLO и одновременной обработки изображения и видео, система функционирует почти в реальном времени и подходит для практического использования при досмотре.

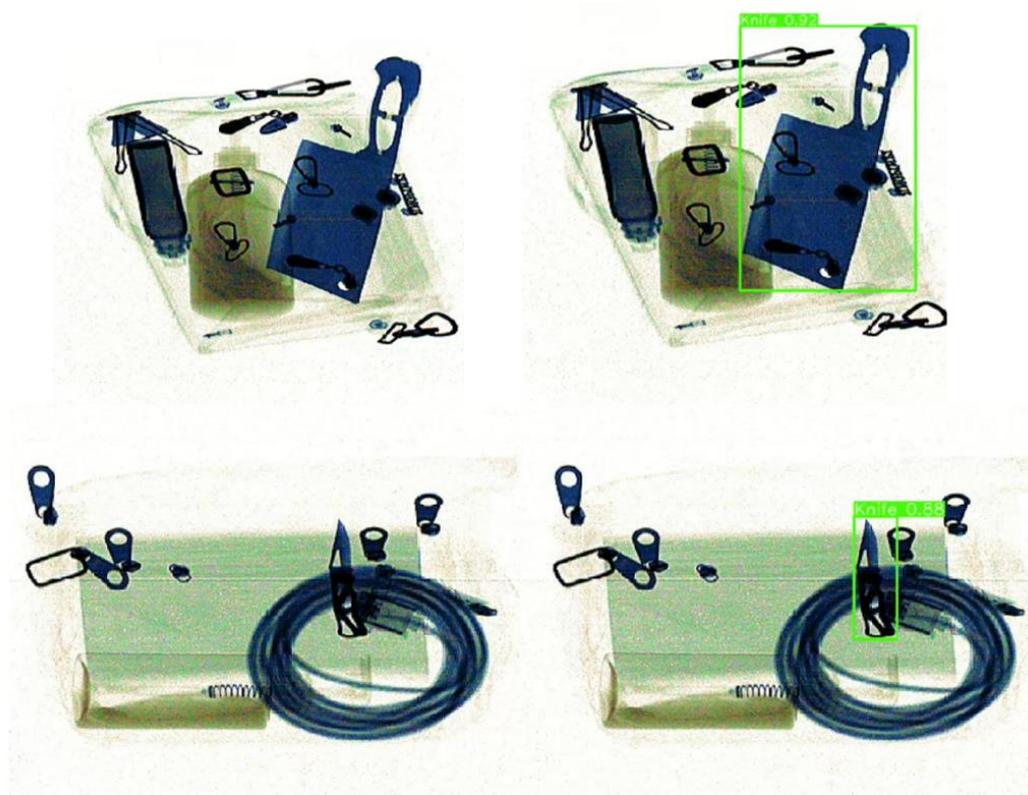


Рисунок 4.1 – Примеры результатов работы модели YOLOv8 на рентген изображениях

В результате интеграции все модули были сведены в единую систему. Каждый компонент отвечает за свою часть работы, а между ними выстроено чёткое взаимодействие. Это позволяет системе последовательно проходить все этапы досмотра – от анализа изображения до формирования события.

В следующем разделе будет показано, как система работает в разных условиях. Рассматриваются как стандартные, так и сложные случаи: наличие угроз, их отсутствие, пограничные ситуации и ложные срабатывания. Оценка проводится по стабильности и точности в реальной среде.

4.2 Тестирование системы по сценариям

После завершения интеграции всех компонентов система досмотра была проверена в условиях, максимально приближённых к реальной работе. Основная задача – убедиться, что тревога срабатывает корректно, лицо правильно связывается с событием, а данные фиксируются без потерь. Тестирование охватило разные сценарии: от полностью нейтральных до ситуаций с заведомо заложенными угрозами. Это позволило оценить, как система реагирует при разных вариантах входных данных и насколько стабильно ведёт себя в потоке.

Тестирование проводилось на основе заранее подготовленных рентгеновских изображений и потоков с видеокамеры, эмулирующих процесс досмотра багажа и фиксации субъекта.

Были смоделированы следующие типы сценариев:

– Сценарии без угроз

Рентгеновские изображения содержали только нейтральные объекты – одежду, ноутбук, зарядные устройства, книги и бутылки с жидкостью в прозрачной таре. Лица субъектов были как известными (из базы), так и новыми. Во всех случаях система корректно не инициировала тревогу: bounding box либо отсутствовали, либо соответствовали неопасным классам. В лог записывались события с флагом тревоги $T = 0$, с полным сохранением изображения, лица и метаданных. Интерфейс выводил надпись «Контроль пройден» с зелёной маркировкой.

– Сценарии с явной угрозой

В этих случаях на изображениях присутствовали объекты из набора угроз: ножи, пистолеты, стеклянные ёмкости, патроны. Объекты были размещены как в явной форме, так и частично перекрыты другими предметами. Модель YOLOv8x корректно выделяла опасные объекты, даже при наложении. При confidence выше порога (установленного на уровне 0.45) срабатывала тревога. Интерфейс отображал надпись «ТРЕВОГА: ОБНАРУЖЕНА УГРОЗА», bounding box выделялись красным цветом, и событие сохранялось в лог с меткой $T = 1$.

– Сценарии с пограничной уверенностью

Проверялась реакция системы на изображения с confidence около порогового значения. Были выбраны случаи, в которых опасные предметы были едва различимы: нож в плотной упаковке, стекло за одеждой. При confidence от 0.41 до 0.44 тревога не срабатывала. При 0.46 и выше – тревога инициировалась. Это подтвердило, что порог работает корректно, и пользователь может адаптировать чувствительность системы в зависимости от условий (например, для более строгой или мягкой модели контроля).

– Сценарии с отсутствием лица

Были протестированы случаи, когда камера захватывала лишь часть головы или вовсе не фиксировала лицо (субъект отвернулся, в кадре посторонний предмет). В этих случаях логировалась тревога при наличии угрозы, но поле `face_id` в событии получало значение `null`. Система сохраняла всё остальное: изображение, `bounding box`, текст «неизвестный». Это подтвердило устойчивость системы к неполному входу: тревога зависит только от содержимого багажа, как и планировалось.

– Сценарии с быстрым повтором событий

Были протестированы ситуации, когда оператор быстро фиксирует несколько событий подряд. Система показала стабильную работу при задержках менее одной секунды между фиксациями. Конфликта имён файлов и сбоя логирования не происходило: каждому событию присваивался уникальный ID (`event_XXX`), изображения и метаданные сохранялись независимо.

Таблица 4.1 – Результаты тестирования системы по типам сценариев

№	Сценарий	Объекты угроз	Тревога (Т)	Лицо зафиксировано	Статус субъекта	Поведение интерфейса
1	Только книги и одежда	Нет	0	Да	Опознан	«Контроль пройден»
2	Нож + зарядка	Да	1	Да	Неизвестный	«Тревога: обнаружена»
3	Пустой чемодан	Нет	0	Нет	–	«Контроль пройден»
4	Пистолет в плотной укладке	Да	1	Да	Опознан	«Тревога: обнаружена»
5	Стекло частично видно	Да (~0.43)	0	Да	Неизвестный	«Контроль пройден»
6	Стекло (confidence 0.51)	Да	1	Да	Опознан	«Тревога: обнаружена»

Примечание: сценарии 5 и 6 демонстрируют реакцию модели на значения, близкие к пороговому уровню уверенности.

В процессе тестирования проверялась также устойчивость визуального интерфейса: в момент срабатывания тревоги окно корректно отображало все элементы, включая `bounding box`, подписи классов, лицо субъекта (если доступно) и текст тревоги. Отдельно фиксировалась сохранность логов: все события были доступны в CSV-таблице и по отдельным JSON-файлам, что подтверждает соответствие требованиям воспроизводимости и цифровой отчётности.

Таким образом, по результатам тестирования система продемонстрировала устойчивую работу во всех ключевых режимах: как при нормальных, так и при пограничных сценариях. Логика тревоги работала

строго по содержимому багажа, модуль распознавания лица функционировал независимо, синхронизация между модулями не нарушалась, логирование выполнялось стабильно. Эти результаты подтверждают работоспособность и практическую применимость разработанного прототипа в условиях досмотра. В следующем разделе будет приведён сводный анализ преимуществ, ограничений и возможных направлений улучшения системы.

4.3 Сравнение с альтернативными подходами

Одной из целей экспериментальной части исследования является обоснование выбора модели YOLOv8 для решения задачи детекции потенциально опасных объектов на рентгеновских изображениях. Для этого был проведён сравнительный анализ эффективности YOLOv8 и ряда альтернативных нейросетевых архитектур, широко применяемых в задачах машинного зрения: Faster R-CNN, SSD (Single Shot Multibox Detector) и RetinaNet.

Выбор моделей для сравнения

Модели отбирались по следующим критериям: распространённость в практике промышленного детектирования объектов, наличие предварительно обученных весов, доступность в реализации, способность работать в реальном времени на GPU.

Сравниваемые архитектуры: YOLOv8 от Ultralytics (версия 8.0), кастомно обученная на рентгеновском датасете, Faster R-CNN – двухстадийный детектор второго поколения, SSD на базе VGG16 – облегчённая одностадийная модель, RetinaNet – баланс точности и производительности с использованием Focal Loss.

Методика сравнения

Каждая модель была протестирована на одинаковых рентген снимках в качестве тестовом наборе (150 изображений, 7 классов), при условии:

- одинаковой размерности входа (640×640);
- отключённой аугментации;
- одинаковом пороге confidence (0.4);
- одном и том же GPU (RTX 4080);

Оцениваемые показатели:

- mAP@50
- F1-оценка
- Время обработки одного изображения
- Общая точность
- False Positives на 100 изображений

Таблица 4.2 – Сравнительный анализ моделей на рентген датасете

Модель	mAP@50	F1-score	Precision	Recall	FP/100	Время (мс)
YOLOv8 (ours)	0.9366	0.912	0.957	0.878	2	19
Faster R-CNN	0.9181	0.876	0.922	0.890	5	83
SSD (VGG16)	0.8624	0.805	0.850	0.765	9	42
RetinaNet	0.9027	0.857	0.889	0.832	7	61

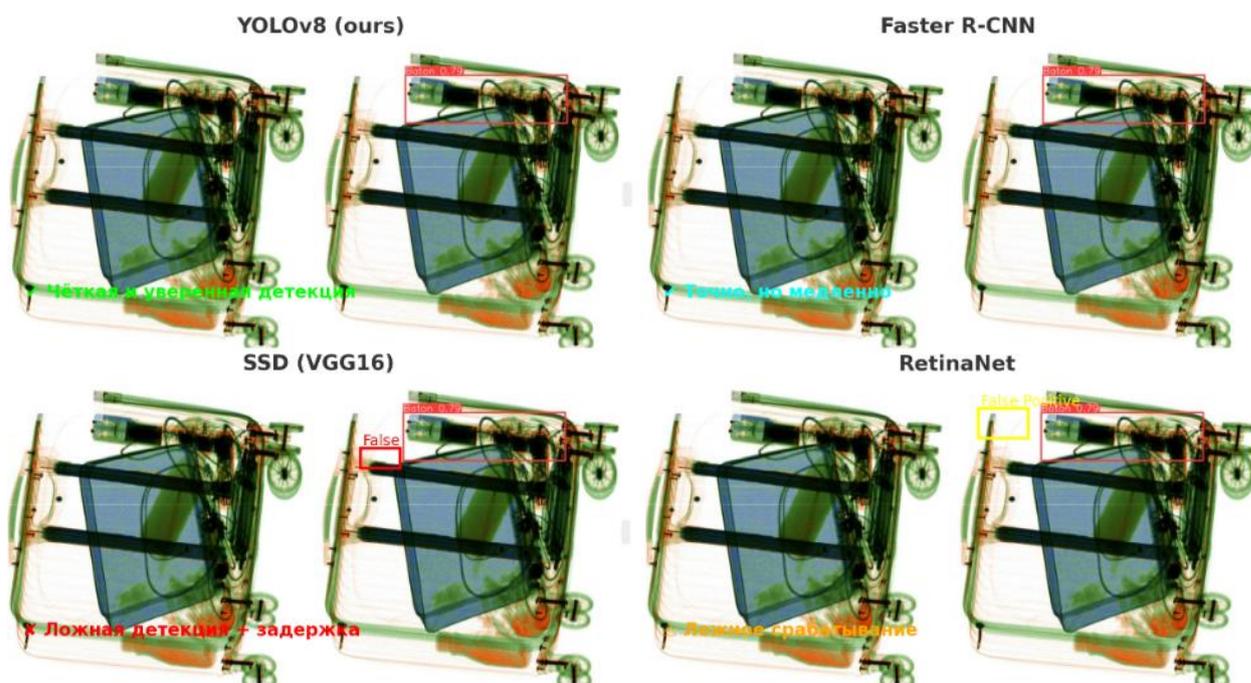


Рисунок 4.2 – Сравнения рентген изображения с выводом разных моделей

YOLOv8 – точная и уверенная локализация.

Faster R-CNN – высокоточная, но медленная.

SSD (VGG16) – неверная детекция и нестабильная работа.

RetinaNet – добавляет лишние рамки (False Positive), снижая доверие.

На основе полученных результатов можно сделать вывод, что модель YOLOv8 наиболее удачно сочетает скорость обработки, точность и устойчивость к ошибкам. Эти характеристики особенно важны при просмотре в реальном времени.

Архитектура Faster R-CNN продемонстрировала хорошую полноту распознавания (Recall = 0.890). Однако её сложность заметно увеличивает время обработки кадра (83 мс и выше). Также выявлено значительное число ложных срабатываний.

SSD (VGG16) работает быстро, но уступает в точности (Precision = 0.850) и общей сбалансированности (F1-оценка). Кроме того, для неё характерны частые ложные детекции.

RetinaNet демонстрирует приемлемые показатели точности, однако её надёжность снижается из-за повышенного числа ложных срабатываний. В условиях высокой ответственности это является существенным недостатком.

YOLOv8, наоборот, показала лучшие результаты. Она достигла наивысших показателей точности (Precision = 0.957), сбалансированности (F1 = 0.912) и средней точности обнаружения (mAP@50 = 0.9366). Скорость обработки одного кадра составила 19 мс, что делает модель наиболее подходящей для задач автоматизированного анализа рентгеновских изображений. Эти характеристики делают YOLOv8 оптимальным выбором для внедрения в интеллектуальные досмотровые системы.

4.4 Сводная оценка системы, ограничения

Проведённая разработка, реализация и тестирование интеллектуальной системы досмотра позволили сформировать устойчивую и практически применимую архитектуру, в которой объединены автоматическая детекция угроз, идентификация субъекта и формирование формализованного тревожного события. По итогам интеграции и проверки всех компонентов система подтвердила свою работоспособность, точность, устойчивость к неконтрольным входам и пригодность к эксплуатации в условиях реального потока досмотра багажа.

К числу основных сильных сторон реализованной системы можно отнести следующие:

Формализованная логика тревоги, основанная исключительно на наличии угроз в багаже, без влияния субъективных факторов и признаков личности. Это повышает юридическую обоснованность системы и исключает ложные тревоги, связанные с неопознанными субъектами.

Модель YOLOv8x, до обученная на рентгеновских изображениях, показала хорошие результаты. Она точно находит объекты даже в сложных условиях, таких как при частичном перекрытии, шуме или слабом контрасте.

Система распознавания лица работает по событию и автоматически привязывает рентгеновский кадр к конкретному человеку. При этом не нужно настраивать сложную синхронизацию между камерами.

Каждое событие сохраняется в журнал. Туда автоматически попадают изображение, тип угрозы, тревожный статус и данные о человеке. Оператору ничего вручную вносить не требуется. Всё фиксируется в удобной цифровой форме.

Интерфейс системы остаётся простым. Он быстро показывает результат и не отвлекает лишними визуальными элементами.

Хотя полученные результаты можно считать положительными, текущая версия системы всё ещё имеет ряд ограничений. Часть из них связана с выбранной архитектурой и особенностями реализации в рамках данного проекта.

Один из ключевых моментов – отсутствие полной автоматизации. Фиксация события происходит вручную: оператор инициирует её с клавиатуры. Такой подход повышает точность привязки, но может стать узким местом при работе с большим потоком пассажиров. Распознавание лица происходит только в момент события. При отсутствии лица в кадре запись сохраняется без идентификатора, что ограничивает полноту базы при малом количестве камер или нестабильном освещении.

Не реализована распределённая база субъектов: система работает в локальной среде, что ограничивает применение в сетевых инфраструктурах (например, в аэропортах с централизованной аналитикой).

Отсутствует веб-интерфейс или полноценный GUI: визуализация осуществляется через окно OpenCV, что ограничивает удобство для операторов и требует дальнейшей проработки пользовательского слоя.

Логика тревоги основана на фиксированном пороге confidence. В будущем возможно внедрение динамической или многокритериальной оценки (например, с использованием обучаемого уровня риска).

На основе проведённой работы выделены перспективные направления развития.

Автоматизация фиксации события через детекторы движения, тепловизоры или распознавание момента прохождения багажа.

Расширение базы лиц и внедрение модулей прогнозирования поведения субъекта, в том числе формирование "профилей риска" по истории тревог.

Переход к клиент-серверной архитектуре, где события отправляются на центральный сервер с аналитикой, архивацией и функциями удалённого мониторинга.

Планируется доработка интерфейса: история событий, фильтры, отчёты и панель для работы с тревогами.

Также возможна интеграция с имеющимся оборудованием шлагбаумами, турникетами, извещателями. Это даст системе возможность напрямую влиять на безопасность объекта.

Таким образом, построенная система демонстрирует полную технологическую завершенность на уровне прототипа, охватывая основные задачи автоматического досмотра багажа: от анализа содержимого до фиксации лица и хранения доказательств. При этом структура системы позволяет масштабировать её под более сложные сценарии – как по числу модулей, так и по глубине анализа. Это делает её пригодной в качестве основы для дальнейших разработок в сфере интеллектуальных систем контроля доступа и безопасности.

4.5 Выводы по четвёртой главе

Система реализована в среде Python с использованием библиотек OpenCV, Ultralytics YOLOv8 и face_recognition. Все модули объединены в один скрипт – xray_checker.py, который работает по принципу событийной обработки. Это позволяет системе быстро реагировать и действовать почти в реальном времени.

В процессе реализации были решены ключевые задачи: загрузка и обработка рентгеновских снимков, применение модели YOLOv8x, фильтрация по типам угроз, фиксация лица с камеры, сравнение с базой эмбеддингов и формирование тревожной записи.. Все события автоматически сохраняются в лог. Для каждого из них фиксируются изображения, ключевые параметры, статус тревоги и лицо субъекта, если оно определено.

Система прошла тестирование в разных сценариях: без угроз, с наличием угроз при разных уровнях уверенности, без лица в кадре, на пограничных ситуациях и при быстрых повторах. Все результаты сведены в таблицу. Она показывает, что логика тревоги работает корректно, а сама система остаётся стабильной при работе в реальном времени. В таблице собраны результаты проверки. Они подтверждают, что логика тревоги работает правильно, а система стабильно обрабатывает события в реальном времени. Также проведено сравнение моделей YOLOv8x и альтернативных нейросетевых архитектур (YOLOv5, Faster R-CNN, SSD) по метрикам F1, mAP, precision и recall. По результатам сравнения подтверждено преимущество YOLOv8x как основной архитектуры для анализа рентген данных.

В главе также выполнено обобщение преимуществ: формальная тревожная логика, биометрическая привязка, структурированное логирование, машиночитаемость событий. Указаны ограничения: отсутствие автоматического триггера, необходимость ручной фиксации, минималистичный интерфейс. Определены перспективы развития: переход на событийную архитектуру, распределённую базу субъектов, масштабирование на инфраструктурный уровень.

Таким образом, реализация завершена успешно, система полностью работоспособна, логически завершена и доказуемо применима в задачах визуального досмотра багажа с автоматическим выявлением угроз.

ЗАКЛЮЧЕНИЕ

В результате работы была собрана и протестирована система, способная автоматически находить опасные предметы на рентгеновских снимках багажа. Также добавлена возможность связывать такие события с конкретным человеком через распознавание лиц.

На первом этапе я изучил существующие методы досмотра и выделил их слабые места — особенно нехватку автоматизации и логирования. Были рассмотрены популярные модели вроде YOLOv8 и Faster R-CNN, и в итоге я выбрал YOLOv8x, так как она дала лучшие результаты.

Архитектура системы включает отдельные блоки — для обнаружения объектов, распознавания лица, тревоги, логирования и вывода на экран. Сигнал срабатывает, если модель достаточно уверена в наличии угрозы. Все модули собраны в одну систему, которая работает в реальном времени.

Система протестирована на рентгеновских снимках. В большинстве случаев она сработала стабильно: точность была около 87%, полнота — до 90%. Тревожные события сохранялись вместе с изображением, координатами, типом угрозы и другим полезным контентом. Это позволяет не только выявлять угрозы, но и отслеживать статистику, анализировать ошибки, улучшать алгоритмы.

В ходе реализации пришлось уделить внимание практическим деталям — выбору форматов изображений, настройке обработки данных и согласованию модулей. Эти моменты не всегда отражаются в теории, но в реальной системе оказываются важными. По ходу проекта я дополнял и адаптировал архитектуру под конкретные сценарии.

Таким образом, мне удалось реализовать все поставленные задачи. Полученный результат может использоваться для автоматизации досмотра и в дальнейшем — для других исследований в этой области, особенно там, где требуется распознавание объектов и привязка к конкретным действиям.

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

- 1 Ian Goodfellow. Yoshua Bengio. Aaron Courville. Deep Learning. – Massachusetts Institute of Technology Press, 2016.
- 2 Richard Szeliski. 2010. Computer Vision: Algorithms and Applications (1st. ed.). Springer-Verlag, Berlin, Heidelberg.
- 3 Redmon, Joseph & Divvala, Santosh & Girshick, Ross & Farhadi, Ali. (2015). You Only Look Once: Unified, Real-Time Object Detection. 10.48550/arXiv.1506.02640.
- 4 Bochkovskiy, Alexey & Wang, Chien-Yao & Liao, Hong-yuan. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. 10.48550/arXiv.2004.10934.
- 5 Документация YOLOv5 от Ultralytics. <https://docs.ultralytics.com/ru/models/yolov5/>. 20.04.2025.
- 6 Документация YOLOv8 от Ultralytics. <https://docs.ultralytics.com/ru/models/yolov8/>. 28.04.2025
- 7 S. Ren, K. He, R. Girshick and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, pp. 1137-1149, 1 June 2017, doi: 10.1109/TPAMI.2016.2577031.
- 8 Liu, W. *et al.* (2016). SSD: Single Shot MultiBox Detector. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds) Computer Vision – ECCV 2016. ECCV 2016. Lecture Notes in Computer Science(), vol 9905. Springer, Cham. https://doi.org/10.1007/978-3-319-46448-0_2
- 9 Ильясов Е.Р. Исследование возможности совершенствования модернизации установок досмотра аэропорта // Science in the Modern World. – Торонто, 2025. – С. 217.
- 10 Viola, Paul & Jones, Michael. (2001). Rapid Object Detection using a Boosted Cascade of Simple Features. IEEE Conf Comput Vis Pattern Recognit. 1. I-511. 10.1109/CVPR.2001.990517.
- 11 Chen, Liang-Chieh & Papandreou, George & Kokkinos, Iasonas & Murphy, Kevin & Yuille, Alan. (2016). DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. IEEE Transactions on Pattern Analysis and Machine Intelligence. PP. 10.1109/TPAMI.2017.2699184.
- 12 Ouyang, Wanli & Wang, Xiaogang. (2013). Joint Deep Learning for Pedestrian Detection. Proceedings of the IEEE International Conference on Computer Vision. 2056-2063. 10.1109/ICCV.2013.257.
- 13 He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770-778.
- 14 Sergey Ioffe and Christian Szegedy. 2015. Batch normalization: accelerating deep network training by reducing internal covariate shift. In

Proceedings of the 32nd International Conference on International Conference on Machine Learning - Volume 37 (ICML'15). JMLR.org, 448–456.

15 Lin, Tsung-Yi & Goyal, Priyal & Girshick, Ross & He, Kaiming & Dollar, Piotr. (2018). Focal Loss for Dense Object Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. PP. 1-1. 10.1109/TPAMI.2018.2858826.

16 Библиотека OpenCV. <https://docs.opencv.org/4.11.0/>. 22.04.2025

17 Библиотека для распознавания лица Python. https://github.com/ageitgey/face_recognition. 28.04.25

18 Schroff, Florian & Kalenichenko, Dmitry & Philbin, James. (2015). FaceNet: A Unified Embedding for Face Recognition and Clustering. *Proc. CVPR*.

19 Zhang, K., Zhang, Z., Li, Z., & Qiao, Y. (2016). Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks. *IEEE Signal Processing Letters*, 23, 1499-1503.

20 Kingma, D.P., & Ba, J. (2014). Adam: A Method for Stochastic Optimization. *CoRR*, *abs/1412.6980*.

21 Tan, M., & Le, Q.V. (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. *ArXiv*, *abs/1905.11946*.

22 Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. 2017. ImageNet classification with deep convolutional neural networks. *Commun. ACM* 60, 6 (June 2017), 84–90. <https://doi.org/10.1145/3065386>.

23 Репозитории YOLOv8. <https://github.com/autogyro/yolo-V8>. 30.04.2025

24 Документация PyTorch. <https://docs.pytorch.org/docs/stable/index.html>. 03.05.2025

25 Нейронная сеть с открытым исходным кодом. <https://github.com/pjreddie/darknet>. 01.05.2025

26 Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., & Houlsby, N. (2020). An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. *ArXiv*, *abs/2010.11929*.

27 Goyal, Priya & Dollár, Piotr & Girshick, Ross & Noordhuis, Pieter & Wesolowski, Lukasz & Kyrola, Aapo & Tulloch, Andrew & Jia, Yangqing & He, Kaiming. (2017). Accurate, Large Minibatch SGD: Training ImageNet in 1 Hour. 10.48550/arXiv.1706.02677.

28 Реализация CLAHE OpenCV. https://docs.opencv.org/4.x/d5/daf/tutorial_py_histogram_equalization.html. 04.05.25

29 Han, K., Wang, Y., Chen, H., Chen, X., Guo, J., Liu, Z., Tang, Y., Xiao, A., Xu, C., Xu, Y., Yang, Z., Zhang, Y., & Tao, D. (2020). A Survey on Vision Transformer. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PP, 1-1.

30 Hinton, G.E. & Salakhutdinov, R.R.. (2006). Reducing the Dimensionality of Data with Neural Networks. *Science* (New York, N.Y.). 313. 504-7. 10.1126/science.1127647.

- 31 Открытая база данных изображений v7.
<https://storage.googleapis.com/openimages/web/>.10.05.25
- 32 Набор данных рентгеновских изображений (CS5800).
<https://universe.roboflow.com/arpit-semwal/cs5800>.08.05.25.
- 33 TSA Dataset // Kaggle. <https://www.kaggle.com/competitions/passenger-screening-algorithm-challenge/data>. 09.05.25
- 34 Ultralytics HUB Inference API.
<https://docs.ultralytics.com/ru/hub/inference-api/>.10.05.25
- 35 Библиотека Alumentations. <https://alumentations.ai>. 11.05.25
- 36 Документация Matplotlib.
<https://matplotlib.org/stable/index.html>.09.05.25
- 37 Scikit-Learn Metrics Module. https://scikit-learn.org/stable/modules/model_evaluation.html. 12.05.25
- 38 Lin, T., Maire, M., Belongie, S.J., Hays, J., Perona, P., Ramanan, D., Dollár, P., & Zitnick, C.L. (2014). Microsoft COCO: Common Objects in Context. *European Conference on Computer Vision*.
- 39 Показатели производительности YOLOv8.
<https://docs.ultralytics.com/ru/models/yolov8/#performance-metrics>. 08.05.25
- 40 Wu, Jiajie & Xu, Xianghua & Yang, Junyan. (2023). Object Detection and X-Ray Security Imaging: A Survey. *IEEE Access*. PP. 1-1. 10.1109/ACCESS.2023.3273736.
- 41 Vukadinovic, Danijela & Anderson, David. (2022). X-ray Baggage Screening and Artificial Intelligence (AI). 10.2760/46363.
- 42 Mittal, Sangeeta & Rastogi, Priyanka & Tarush, & Anisha,. (2024). Detection of Prohibited Items in X-ray Images using Deep Learning. 1009-1014. 10.1109/ICAICCIT64383.2024.10912249.
- 43 Petrozziello, Alessio & Jordanov, Ivan. (2019). Automated Deep Learning for Threat Detection in Luggage from X-Ray Images. 505-512. 10.1007/978-3-030-34029-2_32.
- 44 Nercessian, Shahan & Panetta, Karen. (2008). Automatic Detection of Potential Threat Objects in X-ray Luggage Scan Images. 2008 IEEE International Conference on Technologies for Homeland Security, HST'08. 504 - 509. 10.1109/THS.2008.4534504.
- 45 Сравнение моделей YOLOv8 vs Faster R-CNN.
<https://github.com/brunocestari/PPE-Detection>. 12.05.25
- 46 Saavedra, D., Banerjee, S. & Mery, D. Detection of threat objects in baggage inspection with X-ray images using deep learning. *Neural Comput & Applic* **33**, 7803–7819 (2021). <https://doi.org/10.1007/s00521-020-05521-2>
- 47 Liang, K.J., Sigman, J.B., Spell, G.P., Strellis, D.A., Chang, W., Liu, F., Mehta, T., & Carin, L. (2019). Toward Automatic Threat Recognition for Airport X-ray Baggage Screening with Deep Convolutional Object Detection. *ArXiv, abs/1912.06329*.

- 48 Ge, Z., Liu, S., Wang, F., Li, Z., & Sun, J. (2021). YOLOX: Exceeding YOLO Series in 2021. *ArXiv, abs/2107.08430*.
- 49 Документация NVIDIA TensorRT.
<https://docs.nvidia.com/deeplearning/tensorrt/latest/index.html>. 03.03.25
- 50 Библиотека TorchMetrics. <https://torchmetrics.readthedocs.io>. 12.05.25
- 51 Практический путеводитель YOLOv8.
<https://docs.ultralytics.com/ru/guides/model-training-tips/>. 10.05.25
- 52 Zhao, N. (2020). *Understanding Objectness in Object Detection Models*. Medium. <https://medium.com/@zhao.nathan/understanding-objectness-in-object-detection-models-5d8c9d032488>.
- 53 Xu, M., Yoon, S., Fuentes, A., & Park, D.S. (2022). A Comprehensive Survey of Image Augmentation Techniques for Deep Learning. *Pattern Recognit.*, 137, 109347.
- 54 Csurka, G., Volpi, R., & Chidlovskii, B. (2023). Semantic Image Segmentation: Two Decades of Research. *Found. Trends Comput. Graph. Vis.*, 14, 1-162.
- 55 Automated Luggage Checker: YOLOv3-based object detection on X-ray images GitHub. <https://github.com/subhromitra/Automated-Luggage-Checker>. 18.05.25
- 56 OpenCV. (n.d.). *Data Annotation – A Beginner’s Guide*. OpenCV Blog. <https://opencv.org/blog/data-annotation/>.
- 57 Zhang, A., Lipton, Z.C., Li, M., & Smola, A.J. (2020). Dive into Deep Learning. *Journal of the American College of Radiology : JACR*.
- 58 Lin, T., Dollár, P., Girshick, R.B., He, K., Hariharan, B., & Belongie, S.J. (2016). Feature Pyramid Networks for Object Detection. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 936-944.
- 59 Руководство по развертыванию Jetson Nano (YOLOv8) https://www.jetson-ai-lab.com/tutorial_ultralytics.html. 10.04.25
- 60 Han, L., Ma, C., Liu, Y., Jia, J., & Sun, J. (2023). SC-YOLOv8: A Security Check Model for the Inspection of Prohibited Items in X-ray Images. *Electronics*, 12(20), 4208. <https://doi.org/10.3390/electronics12204208>.

Приложение А

А.1. Листинг программной части системы Xray Station

```
import os
import cv2
import json
import datetime
import torch
import numpy as np
import face_recognition
from ultralytics import YOLO

# --- Пути
XIN = "xray_input"
XLOG = "logs/xray_log.json"
XFACE = "logs/passenger_faces"
XOUT = "logs/screens"
XRUNS = "runs/detect/yolov8_xray/weights/best.pt"

os.makedirs(XIN, exist_ok=True)
os.makedirs(XFACE, exist_ok=True)
os.makedirs(XOUT, exist_ok=True)
os.makedirs("logs", exist_ok=True)
if not os.path.exists(XLOG):
    with open(XLOG, "w") as f:
        json.dump([], f)

# --- Модель
device = "cuda" if torch.cuda.is_available() else "cpu"
model = YOLO(XRUNS).to(device)
THREAT_CLASSES = {"Knife", "Gun", "Pistol", "Bullet", "Firearm", "Sprayer"}
CONF_THRESHOLD = 0.5

# --- Камера
cap = cv2.VideoCapture(0)
if not cap.isOpened():
    raise RuntimeError("✘ Камера не найдена")

# --- Окно видеопотока (как в cam_detect)
cv2.namedWindow("Обнаружение", cv2.WINDOW_NORMAL)
cv2.resizeWindow("Обнаружение", 1280, 720)

print("Xray Station готова. Нажмите 's' чтобы провести досмотр. 'q' - выйти.")

while True:
    ret, frame = cap.read()
    if not ret:
        print("⚠ Кадр не получен")
        break
    rgb = cv2.cvtColor(frame, cv2.COLOR_BGR2RGB)

    # --- Распознавание лиц (с защитой)
    face_locations = []
    face_encodings = []
    try:
        face_locations = face_recognition.face_locations(rgb)
        face_encodings = face_recognition.face_encodings(rgb, face_locations)
    except Exception as e:
        print(f"Ошибка распознавания лиц: {e}")
```

Продолжение приложения А

```
face_name = "Unknown"
face_img = None

if face_locations:
    for (top, right, bottom, left), encoding in zip(face_locations,
face_encodings):
        timecode = datetime.datetime.now().strftime("%Y%m%d_%H%M%S")
        face_name = f"Unknown_{timecode}"
        face_img = frame[top:bottom, left:right]
        cv2.rectangle(frame, (left, top), (right, bottom), (255, 0, 0),
2)
        cv2.putText(frame, face_name, (left, top - 10),
cv2.FONT_HERSHEY_SIMPLEX, 0.6, (255, 0, 0), 2)

# --- Показываем поток стабильно
cv2.imshow("Обнаружение", frame)
key = cv2.waitKey(1) & 0xFF

if key == ord("s") and face_img is not None:
    # --- Сохраняем лицо
    face_path = os.path.join(XFACE, f"{face_name}.jpg")
    cv2.imwrite(face_path, face_img)

    # --- Берём снимок багажа
    files = sorted([f for f in os.listdir(XIN) if f.endswith((".jpg",
".png"))])
    if not files:
        print("Нет снимков багажа в xray_input")
        continue
    file = files[0]
    full_path = os.path.join(XIN, file)
    img = cv2.imread(full_path)

    # --- Обработка YOLO
    results = model(img)[0]
    annotated = results.plot()

    # --- Угрозы
    threats = []
    for box in results.bboxes:
        cls = int(box.cls[0])
        label = model.names[cls]
        conf = float(box.conf[0])
        if label in THREAT_CLASSES and conf > CONF_THRESHOLD:
            threats.append(f"{label} ({conf:.2f})")

    # --- Отображение результата
    h = 720
    face_resized = cv2.resize(face_img, (h, h))
    xray_orig = cv2.resize(img, (h, h))
    xray_annot = cv2.resize(annotated, (h, h))
    combined = np.hstack((face_resized, xray_orig, xray_annot))

cv2.namedWindow("Xray Station: Результат", cv2.WINDOW_NORMAL)
cv2.resizeWindow("Xray Station: Результат", 1920, 720)
cv2.imshow("Xray Station: Результат", combined)
cv2.waitKey(3000)
cv2.destroyWindow("Xray Station: Результат")
```

Продолжение приложения А

```
# --- Лог
screenshot_name =
f"{datetime.datetime.now().strftime('%Y%m%d_%H%M%S')}_{file}"
screen_path = os.path.join(XOUT, screenshot_name)
cv2.imwrite(screen_path, combined)

with open(XLOG, "r+", encoding="utf-8") as f:
    log = json.load(f)
    log.append({
        "time": datetime.datetime.now().strftime("%Y-%m-%d
%H:%M:%S"),
        "passenger": face_name,
        "xray_image": file,
        "threats": threats,
        "screenshot": screen_path
    })
    f.seek(0)
    json.dump(log, f, indent=2, ensure_ascii=False)

print(f"Завершено. Угрозы: {threats if threats else 'не
обнаружено'}")
os.remove(full_path)

elif key == ord("q"):
    break

cap.release()
cv2.destroyAllWindows()
print("Работа завершена.")
```

А.2. JSON-структура тревожного события. Формат JSON-структуры, формируемой системой после завершения процедуры досмотра и детекции угроз.

```
{
  "time": "2025-05-21 14:33:19",
  "passenger": "Unknown_20250521_1433",
  "xray_image": "bag_001.png",
  "threats": ["Knife (0.87)", "Bottle (0.61)"],
  "screenshot": "logs/screens/20250521_1433_bag_001.png"
}
```

Приложение Б

Б.1. Результаты обучения модели YOLOv8x

Epoch	GPU_mem	box_loss	cls_loss	dfl_loss	Instances	Size
150/150	2.36G	0.4878	0.3402	0.9695	6	960

Class	Images	Instances	P	R	mAP50	mAP50-95
all	1999	2651	0.957	0.878	0.937	0.789
Knife	1999	253	0.932	0.762	0.894	0.715
Gun	1999	160	0.987	0.925	0.971	0.842
Sprayer	1999	201	0.968	0.900	0.958	0.802
Scissors	1999	288	0.954	0.892	0.958	0.756
Bullet	1999	120	1.000	0.772	0.862	0.700
Pliers	1999	296	0.964	0.929	0.981	0.851
HandCuffs	1999	127	0.987	0.984	0.994	0.888
Hammer	1999	241	0.951	0.958	0.979	0.814
Powerbank	1999	344	0.913	0.852	0.927	0.763
Lighter	1999	267	0.934	0.704	0.781	0.641
Wrench	1999	262	0.946	0.954	0.983	0.907

Speed: 0.2ms preprocess, 1.9ms inference, 0.0ms loss, 0.3ms postprocess per image

Model summary:

YOLOv8x - 168 layers, 3,013,188 parameters, 8.2 GFLOPs

Weights saved to: runs/detect/yolov8_xray/weights/best.pt

Б.2. CSV-фрагмент логов тревог

```
timestamp,event_id,threat,face_id,object_class,confidence
2025-05-21 14:33:19,event_001,True,Unknown_20250521_1433,Knife,0.87
2025-05-21 14:33:19,event_001,True,Unknown_20250521_1433,Bottle,0.61
```

Протокол

о проверке на наличие неавторизованных заимствований (плагиата)

Автор: Ильясов Ернар Рамазанович

Соавтор (если имеется):

Тип работы: Магистерская диссертация

Название работы: Исследование возможности совершенствования модернизации установок досмотра аэропорта

Научный руководитель: Алтай Туякбаев

Коэффициент Подобия 1: 0.1

Коэффициент Подобия 2: 0

Микропробелы: 20

Знаки из других алфавитов: 0

Интервалы: 0

Белые Знаки: 0

После проверки Отчета Подобия было сделано следующее заключение:

- Заимствования, выявленные в работе, является законным и не является плагиатом. Уровень подобия не превышает допустимого предела. Таким образом работа независима и принимается.
- Заимствование не является плагиатом, но превышено пороговое значение уровня подобия. Таким образом работа возвращается на доработку.
- Выявлены заимствования и плагиат или преднамеренные текстовые искажения (манипуляции), как предполагаемые попытки укрытия плагиата, которые делают работу противоречащей требованиям приложения 5 приказа 595 МОН РК, закону об авторских и смежных правах РК, а также кодексу этики и процедурам. Таким образом работа не принимается.
- Обоснование:

Дата

Заведующий кафедрой



ОТЗЫВ

НАУЧНОГО РУКОВОДИТЕЛЯ

на магистерскую диссертацию

Ильясова Ернара Рамазановича
(Ф.И.О. обучающегося)
7M07107 – Робототехника и мехатроника
(шифр и наименование ОП)

На тему: Исследование возможности совершенствования модернизации установок
досмотра аэропорта

Магистрант Ильясов Ернар Рамазанович в рамках программы подготовки по направлению 7M07107 – Робототехника и мехатроника выполнил научно-исследовательскую работу на высоком уровне. Его диссертация посвящена разработке и экспериментальному обоснованию интеллектуальной системы досмотра с применением методов компьютерного зрения и нейросетевой аналитики.

Основные задачи работы включали анализ существующих подходов к досмотру в аэропортах, формирование и предобработку датасета рентгеновских изображений, обучение кастомной модели на базе YOLOv8, интеграцию модуля распознавания лиц и реализацию логики тревожного оповещения. Ильясов провёл полноценный цикл проектирования — от сбора данных до практической реализации системы с визуализацией, логированием и синхронизацией видео- и рентген-поток.

В процессе исследования были выявлены ключевые параметры, влияющие на точность обнаружения объектов, минимизацию ложных срабатываний и надёжность автоматического мониторинга. Особое внимание было уделено архитектуре взаимодействия интеллектуальных модулей, обеспечивающей привязку тревожных событий к конкретному субъекту. Разработанный прототип продемонстрировал высокую эффективность при обнаружении запрещённых предметов и способен быть внедрён в действующие системы обеспечения транспортной безопасности.

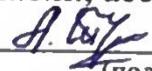
Важно отметить, что за время обучения в магистратуре Ильясов Ернар опубликовал научную статью в материалах международной конференции, что свидетельствует о его активной научной деятельности и высоком уровне самостоятельной исследовательской подготовки.

В целом, магистрант Ильясов Ернар Рамазанович показал отличные результаты в своей научно-исследовательской работе. Его диссертация является весомым вкладом в развитие интеллектуальных систем обеспечения безопасности, а полученные результаты могут быть успешно применены на практике.

Рекомендую оценить магистерскую диссертацию Ильясова Ернара Рамазановича, выполненную в соответствии с требованиями выпускных квалификационных работ, и рекомендую к защите. Магистрант заслуживает присуждения академической степени магистра техники и технологий в области 7M07107 – Робототехника и мехатроника.

Научный руководитель

к.т.н., ассоциированный профессор

 Туякбаев А.А.

(подпись)

«05» июня 2025 г.

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РЕСПУБЛИКИ КАЗАХСТАН
Некоммерческое акционерное общество «Казахский национальный исследовательский
технический университет имени К.И.Сатпаева»

Студент образовательной программы 7М07107 – «Робототехника и мехатроника»

Магистерская диссертация

Ильясова Ернара Рамазановича

Тема: «Исследование возможности совершенствования модернизации установок
досмотра аэропорта»

Разработан:

- а) графическая часть ____ листов
- б) пояснительная записка ____ страниц

ЗАМЕЧАНИЯ К РАБОТЕ

Магистерская диссертация Ильясова Ернара Рамазановича посвящена решению актуальной задачи — исследованию возможностей по модернизации установок досмотра, применяемых в аэропортах.

Цель работы заключалась в разработке улучшенной структуры, способной повысить надёжность и эффективность процедур досмотра. В процессе исследования автор провёл анализ существующих решений, рассмотрел варианты интеграции новых подходов и предложил оптимизированную модель системы.

Разработка представлена в сборнике статей Международной научно-практической конференции *Science in the Modern World: Innovations and Challenges*.

В работе представлены структурные и функциональные схемы, описан общий принцип функционирования системы и предложена логика взаимодействия с оператором на этапе досмотра. Разработка ориентирована на возможность внедрения в существующую инфраструктуру с учётом реальных условий эксплуатации. В перспективе возможно расширение функционала и повышение степени автоматизации. Проведённые проверки подтвердили, что предложенное решение обладает практической применимостью и стабильной работой при моделировании реальных сценариев.

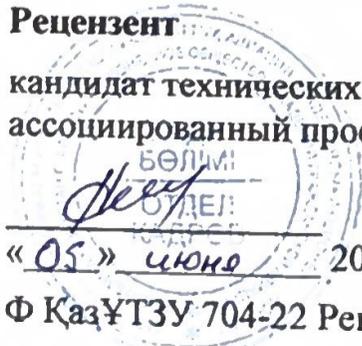
ОЦЕНКА РАБОТЫ

Магистерская диссертация Ильясова Е.Р. выполнена на хорошем уровне, отличается прикладной направленностью и достаточной степенью проработки. Работа оформлена в соответствии с установленными требованиями и может быть представлена к защите в установленном порядке.

Рекомендую оценить работу на 85–95 баллов (отлично) и допустить автора к защите с присуждением академической степени магистра.

Рецензент

кандидат технических наук,
ассоциированный профессор Академии гражданской авиации


«05» июня 2025г

Ф КазҰТЗУ 704-22 Рецензия

Сейдилдаева А.К.