

На правах рукописи



Ковалев Владислав Владимирович

**МЕТОД И АЛГОРИТМЫ РАСПОЗНАВАНИЯ
МАЛОРАЗМЕРНЫХ ИЗОБРАЖЕНИЙ ПОДВИЖНЫХ
ОБЪЕКТОВ НА УСТРОЙСТВАХ С ОГРАНИЧЕННЫМ
ВЫЧИСЛИТЕЛЬНЫМ РЕСУРСОМ**

Специальность 1.2.1. — Искусственный интеллект и машинное обучение
(технические науки)

Автореферат
диссертации на соискание учёной степени
кандидата технических наук

Работа выполнена в Федеральном государственном автономном образовательном учреждении высшего образования «Южный федеральный университет» на кафедре вычислительной техники Института компьютерных технологий и информационной безопасности.

Научный руководитель: **Сергеев Николай Евгеньевич**,
доктор технических наук, доцент,
профессор кафедры вычислительной техники
ИКТИБ ФГАОУ ВО «Южный федеральный
университет» (г. Таганрог)

Официальные оппоненты: **Бутакова Мария Александровна**,
доктор технических наук, профессор,
главный научный сотрудник Ростовского филиала
АО «Научно-исследовательский и проектно-
конструкторский институт информатизации,
автоматизации и связи на железнодорожном
транспорте» (г. Ростов-на-Дону),

Пучков Андрей Юрьевич,
кандидат технических наук,
доцент кафедры информационных технологий в
экономике и управлении филиала ФГБОУ ВО
«Национальный исследовательский университет
Московского энергетического института» (г. Смо-
ленск)

Ведущая организация: ФАУ «Государственный научно-исследовательский институт авиационных систем» (г. Москва)

Защита состоится «16» октября 2025 г. в 11:00 часов на заседании объединённого диссертационного совета совета 99.2.107.02 на базе ФГАОУ ВО «Южный федеральный университет» и ФГБОУ ВО «Южно-Российский государственный политехнический университет (НПИ) имени М.И. Платова», по адресу: 347922, г. Таганрог, пер. Некрасовский, 44, корпус Г, ауд. Г-439.

С диссертацией можно ознакомиться в Зональной научной библиотеке ЮФУ по адресу: г. Ростов-на-Дону, ул. Зорге, 21Ж и на сайте ФГАОУ ВО «Южный федеральный университет» по адресу: <https://hub.sfedu.ru/diss/show/1339092>.

Автореферат разослан «___» _____ 2025 г.

Учёный секретарь
объединенного диссертационного
совета 99.2.107.02,
доктор технических наук, доцент



Ю. А. Кравченко

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования и степень её разработанности.

Одним из наиболее важных и активно развиваемым в мире технологических направлений является цифровая обработка изображений. Автоматический поиск и распознавание объектов по их изображению является востребованной задачей этого направления на текущем научно-техническом этапе. В современных системах анализа последовательности видеок кадров поиск и распознавание объектов является неотъемлемой задачей. Чтобы своевременно и безошибочно принимать решение, предъявляются высокие требования к качеству распознавания и скорости обработки данных алгоритмов, для обеспечения которых применяются нейросетевые подходы на базе операции свёртки. Такие алгоритмы получили название свёрточные нейронные сети.

В большинстве случаев системы раннего обнаружения объектов предназначены для установки на автономные носители, у которых ограничен энергетический ресурс. Поэтому одним из главных критериев, влияющих на выбор вычислительных устройств, на которых выполняются свёрточные нейронные сети, является выделенная или разрешённая потребляемая мощность. В свою очередь, потребляемая мощность вычислительного устройства коррелирует с производительностью устройства. Поэтому достижение высоких показателей качества распознавания и быстродействия вычислительно сложными свёрточными нейронными сетями на устройствах с ограниченным вычислительным ресурсом является нетривиальной задачей.

Для обеспечения раннего обнаружения объектов распознающим системам приходится анализировать объекты на дальних расстояниях от места их фиксации. Такие объекты представлены малой компактной группой пикселей. К числу целевых объектов, подлежащих распознаванию, относятся легковые автомобили, грузовые автомобили, трактора и др. Чтобы успешно решить задачу распознавания, должны быть учтены характеристики видеокамеры (разрешение матрицы камеры, угол обзора, фокусное расстояние и др.) и соблюдены внешние условия (дальность до объекта, физические размеры объекта и др.). Поиск и распознавание малоразмерных изображений объектов вызывает трудности даже у передовых архитектур свёрточных нейронных сетей из-за недостаточного признакового описания объектов.

В силу всего вышесказанного, распознавание малоразмерных изображений объектов на устройствах с ограниченным вычислительным ресурсом является актуальной научно-технической задачей и носит практически ценный характер.

Наиболее значимые результаты в решении проблемы распознавания малоразмерных изображений объектов достигаются за счёт расширения признакового пространства за счёт комплексирования данных различных типов. Большой интерес вызывают видеок кадры с признаком движения, потому что идентификация признаков движения не усложняет конструкцию распознающей системы. Однако

существующие алгоритмы не учитывают признаки движения в процессе продолжительного наблюдения за объектом. В результате получается парадоксальная ситуация, когда современные алгоритмы на базе искусственного интеллекта не используют весь объём имеющихся признаков, содержащихся в последовательности видеок кадров, для повышения качества распознавания малоразмерных изображений объектов свёрточными нейронными сетями.

В данной работе предложены новые научно-технологические разработки, позволяющие решить научно-техническую задачу, состоящую в создании интеллектуальной системы распознавания малоразмерных изображений объектов на устройствах с ограниченным вычислительным ресурсом, что имеет существенное значение для развития области «Технологии создания высокоскоростных транспортных средств и интеллектуальных систем управления новыми видами транспорта» из Перечня критических технологий РФ и разработок в сфере искусственного интеллекта и машинного обучения.

Объектом исследования являются малоразмерные изображения объектов.

Предметом исследования являются методы и алгоритмы распознавания малоразмерных изображений объектов на устройствах с ограниченным вычислительным ресурсом.

Целью работы является повышение качества распознавания малоразмерных изображений объектов свёрточными нейронными сетями на устройствах с ограниченным вычислительным ресурсом.

Научная задача диссертационного исследования состоит в разработке интеллектуальной системы распознавания малоразмерных изображений объектов на устройствах с ограниченным вычислительным ресурсом.

Для решения общей научной задачи в диссертационной работе необходимо решить следующие частные **основные задачи**:

1. Провести обзор и анализ перспективных методов и алгоритмов распознавания малоразмерных изображений.

2. Разработать метод расширения признакового пространства, который позволит повысить информативность системы признаков изображений для задачи распознавания малоразмерных изображений свёрточными нейронными сетями.

3. Разработать алгоритм предварительной обработки изображений с целью комплексирования признаков движения.

4. Сформировать датасеты разномодальных изображений для задачи распознавания малоразмерных изображений целевых объектов.

5. Провести вычислительный эксперимент с целью исследования разномодальных изображений на информативность системы признаков.

6. Провести вычислительный эксперимент с целью исследования влияния разработанного алгоритма предварительной обработки изображений на качество распознавания малоразмерных изображений свёрточных нейронных сетей.

7. Разработать алгоритм предварительной обработки изображений с целью комплексирования признаков движения для нейронного ускорителя.

8. Провести вычислительный эксперимент с целью исследования временных характеристик алгоритма предварительной обработки изображений на нейронных ускорителях.

9. Разработать алгоритм создания интеллектуальной системы распознавания малоразмерных изображений на устройствах с ограниченным вычислительным ресурсом.

10. Построить интеллектуальную систему распознавания малоразмерных изображений на устройствах с ограниченным вычислительным ресурсом.

Методология и методы диссертационного исследования. Для решения выявленных проблем в диссертационной работе использовались теория и методы машинного обучения, распознавания образов, алгоритмы цифровой обработки изображений, методы и алгоритмы предварительной обработки данных, методы статистического анализа, а также разные технологии программирования.

Научная новизна и соответствие научной специальности: В рамках данного исследования получены новые научные результаты:

1. Разработан метод расширения признакового пространства, повышающий информативность системы признаков, что позволило снизить частоту ошибок распознавания малоразмерных изображений объектов на 11.5 – 14.5 % (п. 4 паспорта специальности; страницы 53-62, 91-100 диссертации).

2. Разработан алгоритм предварительной обработки изображений, который учитывает больше признаков движения, что позволило повысить вероятность распознавания малоразмерных изображений объектов свёрточными нейронными сетями на 14.9 – 20.1 % (п. 4 паспорта специальности; страницы 66-70, 101-108 диссертации).

3. Разработан алгоритм формирования условно-реальных данных, отличающийся возможностью генерировать аннотированные разномодальные изображения, позволяющий заменить ручной процесс разметки данных автоматическим (п. 14 паспорта специальности; страницы 70-75, 92-94 диссертации).

4. Разработан алгоритм предварительной обработки изображений с целью комплексирования признаков движения на нейронном ускорителе, что позволило сократить время выполнения алгоритма в 1.3 – 3.2 раза (п. 4 паспорта специальности; страницы 78-82, 109-119 диссертации).

5. Создана интеллектуальная система распознавания малоразмерных изображений целевых объектов, которая обеспечивает требуемые характеристики в качестве распознавания, скорости обработки данных и допустимой потребляемой мощности вычислительного устройства (п. 4 паспорта специальности; страницы 82-89, 119-123 диссертации).

Новые результаты, полученные в ходе выполнения диссертационного исследования, соответствуют паспорту научной специальности 1.2.1. «Искусственный интеллект и машинное обучение (технические науки)» в части следующих пунктов:

П. 4. Разработка методов, алгоритмов и создание систем искусственного интеллекта и машинного обучения для обработки и анализа текстов на естественном языке, для изображений, речи, биомедицины и других специальных видов данных.

П. 14. Методы и средства формирования массивов условно-реальных данных и прецедентов, необходимых для решения задач искусственного интеллекта и машинного обучения.

Практическая значимость заключается в возможности использования разработанных алгоритмических и программных решений для создания интеллектуальных систем распознавания малоразмерных изображений объектов на устройствах с ограниченным вычислительным ресурсом, удовлетворяющих заданным характеристикам в качестве распознавания, скорости обработки данных и допустимой потребляемой мощности.

Теоретическая значимость заключается в возможности применения подходов увеличения информативности системы признаков, рассмотренных в разработанном методе расширения признакового пространства, для усовершенствования методов распознавания образов.

Положения, выносимые на защиту:

1. Метод расширения исходного признакового пространства, который позволяет повысить информативность системы признаков за счёт добавления признаков движения.

2. Алгоритм предварительной обработки изображений с целью комплексования признаков движения на центральном процессоре, который позволяет повысить качество распознавания малоразмерных изображений свёрточными нейронными сетями.

3. Алгоритм формирования условно-реальных данных, который позволяет генерировать в автоматическом режиме аннотированные разномодальные изображения.

4. Алгоритм предварительной обработки изображений с целью комплексования признаков движения на нейронном ускорителе, который позволяет повысить скорость обработки данных.

5. Алгоритм построения интеллектуальной системы распознавания малоразмерных изображений объектов на устройствах с ограниченным вычислительным ресурсом, которая удовлетворяет требуемым характеристикам.

Достоверность и обоснованность полученных автором теоретических и практических результатов подтверждена корректным проведением вычислительных экспериментов, основанных на валидных данных, репрезентативных критериях и не противоречит выдвинутому теоретическим предположениям.

Реализация и внедрение результатов работы. Основные результаты диссертационного исследования внедрены в учебный процесс кафедры вычислительной техники ФГАОУ ВО «Южный федеральный университет» при подготовке бакалавров основной образовательной программы «Технологии

искусственного интеллекта» по направлению 09.03.01 «Информатика и вычислительная техника» и магистров основной образовательной программы «Высокопроизводительные вычислительные системы и квантовая обработка информации» 09.04.01 по направлению «Информатика и вычислительная техника». Теоретические и практические результаты диссертационной работы внедрены в научно-исследовательской работе «Гепард-Деталь-НКБ ВСБЦВМ-ЭИИ», выполненной в «Акционерное общество Научно-конструкторское бюро вычислительных систем».

Апробация работы. Основные положения и результаты диссертационной работы докладывались и обсуждались: на XVIII Всероссийской научной конференции молодых учёных, аспирантов и студентов «Информационные технологии, системный анализ и управление» (Таганрог, 2020); VII, IX Всероссийской научно-технической конференции «Фундаментальные и прикладные аспекты компьютерных технологий и информационной безопасности» (Таганрог, 2021, 2023); Международном научно-техническом конгрессе «Интеллектуальные системы и информационные технологии – 2023».

Публикации. Основные результаты по теме диссертации изложены в 13 печатных изданиях, 6 из которых опубликованы в журналах, рекомендованных ВАК РФ, 7 статей в источниках, индексированных в РИНЦ.

Личное участие соискателя в получении результатов. Все основные результаты диссертации получены непосредственно и лично автором.

Объём и структура работы. Диссертация состоит из введения, 4 глав, заключения и 3 приложений. Полный объём диссертации составляет 162 страницы, включая 55 рисунков и 17 таблиц. Список литературы содержит 139 наименований.

Краткое содержание работы

Во введении обоснована актуальность тематики диссертационного исследования, приведены: объект, предмет, цель исследования, сформулирована научная задача, указаны методы исследования, сформулированы положения, выносимые на защиту с сопутствующей научной новизной и соответствием паспорту специальности.

В первой главе рассмотрены основные понятия темы исследования. Сформулированы основные типы задач распознавания образов. Определен критерий малоразмерности изображений объектов с точки зрения способности наблюдателя распознавать изображения объектов через устройство наблюдения по методу Джона Джонсона. С учётом метода Джонсона выбран минимальный размер изображения объекта, который находится в диапазоне 8 – 14 пикселей, что соответствует распознаванию с вероятностью 30 – 50%. Сам объект должен иметь большие физические размеры и находиться на дальних расстояниях.

Основные критерии качества распознавания изображений объектов, на основе которых проведены исследования. Ошибка обнаружения оценивается с помощью коэффициента Джаккара или *IntersectionOverUnion (IOU)*, который характеризует меру степени пересечения между двумя обрамляющими прямоугольниками.

Критерии качества классификации в задаче распознавания рассчитывается исключительно на основе четырёх возможных результатов сопоставления экспертной оценки с решением алгоритма (предиктора) на тестовом наборе данных: *TruePositive (TP)* – количество истинно позитивных срабатываний, *FalsePositive (FP)* – количество ложно позитивных срабатываний, *FalseNegative (FN)* – количество ложно негативных срабатываний, *TrueNegative (TN)* – количество истинно негативных срабатываний. Их производными являются вероятность правильного распознавания P_{pac} и частота ложной тревоги $F_{\text{лт}}$:

$$F_{\text{лт}}|_{FP \leq K_{\text{кадр}}} = \frac{FP}{K_{\text{кадр}}}, \quad (1)$$

где $K_{\text{кадр}}$ – количество изображений в выборке. Для фиксированной частоты ложной тревоги оценивается вероятность правильного распознавания $P_{\text{pac}}|_{F_{\text{лт}}=\text{const}}$ как долю истинно позитивных решений среди всех позитивных образцов:

$$P_{\text{pac}}|_{F_{\text{лт}}=\text{const}} = \frac{TP}{TP + FN}. \quad (2)$$

Классическим критерием, отражающим долю ошибок алгоритма среди всех решений:

$$\text{ErrorRate} = \frac{FP + FN}{TP + FP + FN}. \quad (3)$$

Еще одними производными показателей ошибок классификации являются количественные характеристики качества распознавания Полнота (*Recall*) и Точность (*Precision*). *Precision* характеризует долю правильно предсказанных позитивных решений среди всех позитивных решений алгоритма для одного класса:

$$\text{Precision} = \frac{TP}{TP + FP}. \quad (4)$$

Recall отражает долю истинно позитивных решений среди всех образцов одного класса:

$$\text{Recall} = \frac{TP}{TP + FN}. \quad (5)$$

Варьированием порога уверенности решения алгоритма формируется зависимость *Precision – Recall*. На основе зависимости функции *Precision –*

Recall можно вычислить критерий *Average Precision* (*AP*), который характеризует качество распознавания одного класса *AP* и равен площади под кривой. Критерием качества распознавания всех классов является интегральная характеристика *Mean Average Precision* (*mAP*):

$$mAP = \frac{\sum_{i=1}^k AP_i}{k}, \quad (6)$$

где k – количество классов.

Выявлены проблемы, связанные с высокими затратами на создание интеллектуальной системы распознавания малоразмерных изображений объектов на устройствах с ограниченным вычислительным ресурсом. В результате анализа причин, определяющих высокие затраты разработки распознающей системы, выявлен ряд существенных проблем, связанных с неудовлетворительным качеством распознавания, высокой стоимостью процедуры обучения свёрточных нейронных сетей, неприемлемой временной задержкой обработки данных на устройствах с ограниченным вычислительным ресурсом.

Во второй главе предложен метод расширения признакового пространства $X = \{x_i\}$ с целью повышения информативности системы признаков, который опирается на априорные знания о характере признаков, протекающих во времени. Парадигма метода заключается в формировании нового признакового пространства $Y = \{y_i\}$ на основе предварительной обработки изображений, определяющих признаковое пространство X , в котором минимизированы зоны неоднозначности посредством увеличения расстояние между границами образов. Это достигается за счёт расширения исходного признакового пространства новыми различительными признаками движения.

Формирование решающей функции D модели свёрточной нейронной сети (СНС) осуществляется с помощью обучения на тренировочной выборке, которая определяет признаковое пространство Y . В процессе обучения СНС вычисляются весовые коэффициенты $W = \{w_i\}$ искомого решения, которые определяют разделяющую гиперплоскость. Решающая функция D , полученная на данных с расширенным признаковым пространством, сформировала весовые коэффициенты, которые определили границу между образами для классификации объектов с меньшим количеством ошибок.

На первом этапе в блок формирования динамических признаков поступают временная последовательность изображений, на основе которой выделяются признаки движения, характеризующие различную скорость движения объектов Q . В результате чего формируются новые признаки, характеризующие движение объектов в процессе их изменения.

На втором этапе происходит формирование результирующих изображений, которые определяют новое признаковое пространство Y с помощью объединения статических признаков X со всей совокупностью признаков движения. Таким образом, результирующее изображение содержит исходные признаки изображения и новые признаки движения $Q = \{q_i\}$.

С учётом разработанного метода расширения признакового пространства с целью повышения информативности системы признаков распознавание одноэтапными свёрточными нейронными сетями примет вид (рисунок 1).

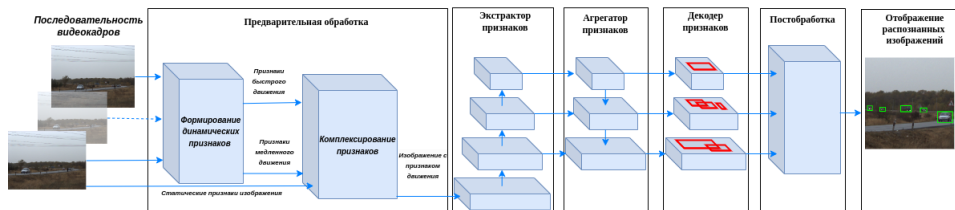


Рис. 1 – Структурная схема распознавания малоразмерных изображений объектов одноэтапными свёрточными нейронными сетями с применением предварительной обработки изображений

Для повышения качества распознавания малоразмерных изображений предложен алгоритм предварительной обработки изображений с целью комплексирования признаков движения. Предполагается, что на вход алгоритма поступает последовательные во времени кадры в градациях серого со статичным фоном. На первом этапе последовательные во времени изображения накапливаются в линии задержки с отводами (рисунок 2). Период выдачи кадров камерой определяется временной задержкой T_0 . Накопленная последовательность видеок кадров поступает в блок идентификации движения.

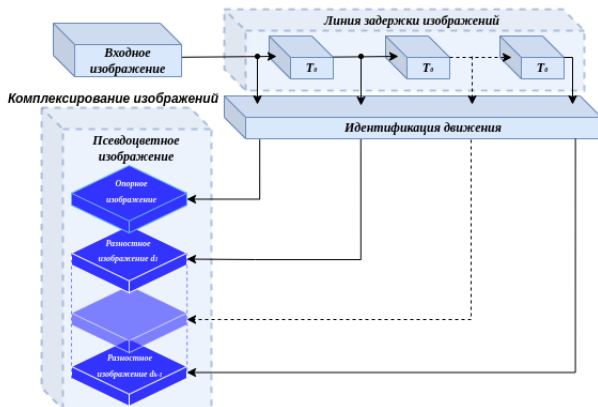


Рис. 2 – Структурная схема алгоритма предварительной обработки изображений с целью комплексирования признаков движения

С учётом компромисса между качеством идентификации и вычислительной сложностью, алгоритм идентификации движения основан на построении

разностных изображений между опорным изображением и задержанными изображениями:

$$d_i(x,y) = |f(x,y,t_0) - f(x,y,t_{i+1})|, \quad (7)$$

где i – индекс разностного изображения, который находится в диапазоне от 0 до $k - 1$, где k – количество задержанных кадров; $f(x,y,t_0)$ и $f(x,y,t_{i+1})$ – изображения, полученные в моменты времени t_0, t_{i+1} ; x и y – пространственные координаты. Изображение, полученное в момент времени t_0 , является опорным, а изображения, полученные в моменты времени t_i , где $i = 1, 2, \dots, k - 1$, являются задержанными. Разностные изображения d_i формируются путем сравнения опорного изображения с задержанными изображениями.

Предложен алгоритм автоматического формирования аннотированных условно-реальных разномодальных изображений, что позволяет уменьшить стоимость разработки распознающей системы. В алгоритме используется средство моделирования 3D-графики Unreal Engine 5, на основе которого можно получить информацию о синтезируемой сцене. При формировании данного изображения используются вспомогательные каналы, такие как: карта глубины, карта нормалей, карта масок. С учётом всех возможностей средств разработки 3D-графики существует возможность формирования разномодальных изображений. К числу таких изображений относятся синтетические инфракрасные изображения, радиолокационные изображения, видеокадры с признаком движения и др.

Далее необходимо провести комплексирование сформированных разномодальных данных изображений выбранным способом. К числу наиболее распространенных методов комплексирования изображений можно отнести поканальное слияние на уровне пикселей, метод главных компонент, дискретное вейвлет-преобразование и др. Формирование разметки осуществляется на основе изображения с масками объектов, где интенсивности масок имеют уникальные значения, определяющие конкретные классы объектов.

В третьей главе рассмотрено решение проблемы, связанной с высокой временной задержкой выполнения разработанного алгоритма предварительной обработки изображений на центральном процессоре. Чтобы повысить скорость обработки данных алгоритма, предложена его реализация на нейронном ускорителе. Это достигается за счет объединения вычислений предварительной обработки изображений с экстрактором признаков СНС и переносом на вычислительные мощности нейронного ускорителя.

Алгоритм формирует комплексированное изображение с двумя разностными каналами на основе опорного изображения и двух изображений, задержанных на T_0 и $4T_0$. С учётом выбранных задержек, для получения комплексированного изображения необходимо накопить последовательность из пяти изображений. На основе накопленной последовательности формируется тензор из изображений, полученных в настоящий момент времени и задержанных на T_0 и $4T_0$. Формирование опорного изображения и двух разностных изображений реализовано с помощью операции «двумерная свёртка» с размерностью фильтров $3 \times$

3 x 1 x 1 и значениями 0, 1, -1. Очередность значений фильтра 1, -1, 0 формирует разностное изображение, характеризующее наличие медленного движения, а фильтр со значениями 1, 0, -1 формирует разностное изображение, характеризующее наличие быстрого движения. Выполнение алгоритма можно представить в виде матричных преобразований (свёртка):

$$Y = \sigma(X * W) = |X * W|, \quad (8)$$

где Y – трёхмерная матрица выходных значений; X – трёхмерная матрица входных значений; σ – функция активации абсолютных величин; W – трёхмерная матрица весовых коэффициентов.

Для решения проблемы, связанной с созданием интеллектуальной системы распознавания малоразмерных изображений объектов на устройствах с ограниченным вычислительным ресурсом, предложен алгоритм, который позволяет построить аппаратно-программный комплекс, удовлетворяющий заданным характеристикам. Порядок действий алгоритма формирования набора для создания распознающей системы:

1. *Формирование исходного состава компонент и требуемых характеристик распознающей системы.*

а) *Требуемые характеристики.* На данном этапе задаются необходимые характеристики распознающей системы.

б) *Формирование множества вычислительных устройств.* Множество вычислительных устройств может быть сформировано на основе различных требований.

в) *Формирование множества архитектур СНС.*

г) *Формирование данных.*

2. *Формирование набора.* Из множества вычислительных устройств выбирается непроанализированное вычислительное устройство с программной поддержкой. Если рассмотрены все вычислительные устройства, то происходит переход к шагу 3.

а) *Выбор архитектуры СНС.* Из множества архитектур СНС выбирается непроанализированная архитектура СНС. Если рассмотрены все архитектуры СНС, то происходит переход к шагу 2.

б) *Обучение модели СНС.*

в) *Выбор разрядности вычислений.* На этом этапе алгоритма происходит выбор разрядности вычислений, который поддерживается устройством. Если рассмотрены все поддерживаемые разрядности вычислений для данного устройства, то происходит переход к шагу 2 а.

г) *Построение модели СНС.* На выбранном вычислительном устройстве происходит построение обученной модели СНС с учётом выбранной разрядности вычислений.

д) *Оценка времени выполнения СНС.* На этом шаге происходит оценка времени выполнения модели СНС на вычислительном устройстве. В случае

неудовлетворительного времени выполнения модели происходит переход к шагу 2 в.

е) *Оценка качества распознавания изображений объектов.* В случае неудовлетворительного качества распознавания модели СНС, происходит переход к шагу 2 а.

ж) *Буферизация набора.* Если все характеристики пространной системы удовлетворяют заданным критериям, то составные части набора буферизируются.

3. *Выбор набора.* После того как будут рассмотрены все удовлетворяющие комбинации, наступает заключительный этап выбора лучшего набора на основе полученных оценок характеристик распознающей системы методом взвешенной суммы нормированных неоднородных критериев:

$$F(\hat{f}|w) = w_0\hat{f}_0 + \dots + w_m\hat{f}_m, \quad (9)$$

где w – весовые коэффициенты (коэффициенты важности) критерия, \hat{f} – нормированные значения неоднородных критериев.

На основе рассчитанных показателей значимости F выбирается набор с максимальным значением.

С помощью сформированного набора можно построить интеллектуальную систему распознавания малоразмерных изображений, удовлетворяющую заданным характеристикам согласно следующему порядку действий:

1. Набор для создания распознающей системы.
2. Обучение модели СНС.
3. Переход к вычислениям с выбранной разрядностью.
4. Построение модели СНС на вычислительном устройстве.

В четвёртой главе представлена апробация разработанного метода и алгоритмов с помощью проведения вычислительных экспериментов. В рамках экспериментов сформировано множество тестовых архитектур СНС: YOLOv3-Tiny, YOLOv5s, YOLOv5l, YOLOv8s. На основе разработанного алгоритма формирования аннотированных данных условно-реальных изображений получены три датасета аннотированных условно-реальных изображений. Первый датасет состоит из изображений в градациях серого, второй датасет из комплексированных изображений с одним разностным каналом, а третий датасет с двумя разностными каналами.

Каждая СНС обучена на трёх датасетах. Проведена количественная оценка качества распознавания обученных СНС с помощью критериев mAP и вероятности правильного распознавания $P_{\text{рас}}$ при фиксированной частоте ложной тревоги $F_{\text{лт}} = 5 \cdot 10^{-4}$.

Информативность системы признаков отражает количество ошибок распознавания алгоритма, поэтому используется частота ошибок $ErrorRate$, которая отражает совокупность ошибок классификации первого и второго рода. СНС, обученные на датасетах с двумя разностными каналами, уменьшили частоту ошибок распознавания на 11.5 – 14.5 % по сравнению с СНС, обученными на

исходном датасете. Исходя из того, что информативность системы признакового пространства характеризует величина *Error Rate*, то признаковое пространство, представленное изображениями с двумя разностными каналами, имеет большую информативность. Таким образом, предложенный метод расширения исходного признакового пространства, с помощью добавления признаков движения для нескольких моментов времени, позволяет сформировать систему признаков, обладающую большей информативностью.

Проведено исследование влияния разработанного алгоритма предварительной обработки изображений с целью комплексирования признаков движения на качество распознавания малоразмерных изображений объектов с помощью вычислительного эксперимента, по результатам которого можно сделать выводы, что СНС, обученные на изображениях с двумя разностными каналами, повысили вероятность распознавания на 14.9 – 20.1 % и на 13.4 – 18.8 % *mAP*, чем СНС, обученные на исходных изображениях.

Проведён вычислительный эксперимент с целью исследования временных характеристик разработанного алгоритма предварительной обработки изображений для нейронного ускорителя. В рамках эксперимента сформировано множество вычислительных устройств с аппаратной поддержкой графического и нейронного процессоров: NVIDIA Jetson Nano, NVIDIA Jetson TX2, NVIDIA Jetson AGX Xavier и Rockchip RK3588.

На центральном и нейронных ускорителях всех вычислительных устройств построен алгоритм предварительной обработки изображений с целью комплексирования признаков движения. По результатам оценки времени выполнения алгоритма можно сделать вывод, что применение алгоритма позволило увеличить скорость обработки данных в 1.3 – 3.2 раза. Проведение целочисленных вычислений привело к снижению качества распознавания СНС на 6.5 – 19.1 % *mAP* или 5.4 – 15.9 % $P_{\text{рас}}$.

На основе разработанного алгоритма построения интеллектуальной системы, которая удовлетворяет требованиям: потребляемая мощность не более 30 Вт, вероятность распознавания не менее 85 % при частоте ложной тревоги $5 \cdot 10^{-4}$, время выполнения не более 60 мс. Распознающую интеллектуальную систему можно построить на основе сформированного набора, состоящего из вычислительного устройства NVIDIA Jetson AGX Xavier, фреймворка TensorRT, модифицированной архитектуры СНС YOLOv5s-M, половинной разрядности вычислений (Floating Point 16) и датасетов комплексированных изображений с двумя разностными каналами. В конечном итоге распознающая система обеспечивает вероятность распознавания 86.4 %, время выполнения 22 мс, потребляемая мощность устройства не превышает 30 Вт.

ЗАКЛЮЧЕНИЕ

В ходе выполнения диссертационного исследования достигнута поставленная цель, заключающаяся в повышении качества распознавания малоразмерных изображений объектов свёрточными нейронными сетями на устройствах с ограниченным вычислительным ресурсом. Научные и практические результаты диссертационного исследования имеют большое значение для развития технологий искусственного интеллекта при обработке и анализе изображений, успешно апробированы на конференциях и симпозиуме, а также внедрены в научно-исследовательской работе и учебный процесс.

Основные научные и практические результаты диссертационного исследования заключаются в следующем:

1. Разработан метод расширения признакового пространства с целью повышения информативности системы признаков. Расширение признакового пространства осуществляется с помощью предварительной обработки изображений, которая выделяет признаки движения и комплексирования с исходными признаками изображения. В расширенном признаковом пространстве минимизированы зоны неоднозначности за счёт увеличения расстояния между границами образов.

2. Разработан алгоритм предварительной обработки изображений с целью комплексирования признаков движения. Алгоритм формирует комплексированное изображение, в котором учтены исходные признаки изображения с признаками движения, полученными в процессе продолжительного наблюдения. Результирующее изображение является псевдоцветным изображением с цветными кластерами, характеризующими признаки движения объектов.

3. Для минимизации стоимости разработки распознающей системы на этапе обучения свёрточных нейронных сетей разработан алгоритм формирования условно-реальных разномодальных изображений, который позволяет в автоматическом режиме генерировать аннотированные изображения, что позволяет уменьшить стоимость процедуры формирования датасетов.

4. Разработан алгоритм предварительной обработки изображений с целью комплексирования признаков движения на нейронном ускорителе. Алгоритм предварительной обработки изображений построен в рамках операций, поддерживаемых нейронными ускорителями, и является составной частью экстрактора признаков свёрточной нейронной сети, что позволяет обрабатывать данные на вычислительных мощностях нейронного ускорителя.

5. Разработан алгоритм построения системы распознавания малоразмерных изображений объектов на устройствах с ограниченным вычислительным ресурсом. Алгоритм формирует набор из вычислительного устройства с программной поддержкой, архитектуры свёрточной нейронной сети, датасетов и разрядности вычислений нейронного ускорителя. С помощью определенного порядка действий над составными частями набора можно построить распознающую систему, удовлетворяющую требованиям в качестве распознавания

изображений объектов, времени обработки данных и допустимой потребляемой мощности.

6. Для исследования и апробации разработанного метода и алгоритмов сформированы датасеты аннотированных синтетических изображений для задачи распознавания малоразмерных изображений объектов с помощью разработанного алгоритма формирования аннотированных условно-реальных разномодальных изображений. Первый датасет состоит из исходных изображений в градациях серого, второй датасет состоит из комплексированных изображений с одним разностным каналом, третий состоит из комплексированных изображений с двумя разностными каналами.

7. Проведено исследование разработанного метода расширения признакового пространства с целью повышения информативности системы признаков с помощью вычислительного эксперимента. Сформировано множество тестовых архитектур свёрточных нейронных сетей для задачи распознавания малоразмерных изображений объектов. Применение метода позволило уменьшить частоту ошибок распознавания на 11.5 – 14.5 %. Полученные значения частоты ошибок распознавания говорят о том, что применение метода расширения признакового пространства позволяет сформировать изображения, которые содержат больше различительных признаков.

8. Проведено исследование влияния разработанного алгоритма предварительной обработки изображений на качество распознавания свёрточных нейронных сетей с помощью вычислительного эксперимента. Применение алгоритма позволило повысить вероятность распознавания свёрточных нейронных сетей на 14.9 – 20.1 %. Повышение качества распознавания обусловлено большей информативностью системы признаков, содержащихся в комплексированных изображениях с двумя разностными каналами.

9. Проведено исследование времени выполнения разработанного алгоритма предварительной обработки изображений для нейронного ускорителя с помощью вычислительного эксперимента. На основе полученных оценок времени выполнения можно сделать выводы, что применение предложенного подхода позволило сократить время выполнения алгоритма в 1.3 – 3.2 раза по сравнению с центральным процессором без изменения качества распознавания.

10. Создана интеллектуальная распознающая система, которая удовлетворяет заданным характеристикам в качестве распознавания, времени обработки данных и допустимой потребляемой мощности вычислительного устройства. Система построена на основе набора, состоящего: из вычислительного устройства NVIDIA Jetson AGX Xavier, фреймворка TensorRT, модифицированной архитектуры свёрточной нейронной сети YOLOv5s-M, половинной разрядности вычислений графического процессора и датасетов комплексированных изображений с двумя разностными каналами. Система обеспечивает вероятность распознавания 86.4 % при фиксированной частоте ложной тревоги, время выполнения 22 мс и потребляемую мощность, не превышающую 30 Вт.

Все основные результаты диссертационного исследования направлены на решение научной задачи, достижения цели и являются значимыми для развития технологий искусственного интеллекта и машинного обучения.

Согласованность практических результатов, полученных в результате вычислительного эксперимента, с теоретическими предположениями и известными фактами теории распознавания образов подтверждает достоверность научных результатов.

Дальнейшим направлением исследований является разработка и анализ новых алгоритмов предварительной обработки последовательности кадров, позволяющих сформировать комплексированные изображения с большим количеством различительных признаков движения, для повышения качества распознавания малоразмерных изображений объектов алгоритмами искусственного интеллекта и машинного обучения.

ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИОННОЙ РАБОТЫ

Публикации в изданиях, рекомендованных ВАК РФ:

1. Ковалев В. В. Алгоритм предварительной обработки изображений для снижения вероятности переобучения свёрточных нейронных сетей на нейронном ускорителе [Текст] // Известия ЮФУ. Технические науки. – 2024. – 5 (241). – С. 29–37. (ВАК, К2)

2. Ковалев В. В., Сергеев Н. Е. Расширение признакового пространства в задаче поиска и распознавания малоразмерных объектов на изображениях [Текст] // Известия ЮФУ. Технические науки. – 2024. – № 1. (ВАК, К2)

3. А. Е. Щелкунов, В. В. Ковалев, И. В. Сидько, Н. Е. Сергеев Ускорение прямого прохода при реализации СНС на ограниченном вычислительном ресурсе [Текст] // Известия ЮФУ. Технические науки. – 2022. – № 1. – С. 289–297. (ВАК, К2)

4. Ковалев В. В., Сергеев Н. Е. Реализация сверточных нейронных сетей на встраиваемых устройствах с ограниченным вычислительным ресурсом [Текст] // Известия ЮФУ. Технические науки. – 2021. – 6 (223). – С. 64–72. (ВАК, К2)

5. Ковалев В. В., Сергеев Н. Е. Алгоритм предварительной обработки видеоизображений для повышения точности обнаружения малоразмерных образов [Текст] // Известия ЮФУ. Технические науки. – 2021. – 5 (222). – С. 146–154. (ВАК, К2)

6. А. Е. Щелкунов, В. В. Ковалев, К. И. Морев, И. В. Сидько Метрики оценки алгоритмов автоматического сопровождения [Текст] // Известия ЮФУ. Технические науки. – 2020. – 1 (211). – С. 233–245. (ВАК, К2)

Публикации в других изданиях:

7. Ковалев В. В., Сергеев Н. Е. Построение системы поиска и распознавания объектов на изображениях для устройств с ограниченным вычислительным

ресурсом [Текст] // Сборник научных трудов XII Всероссийской научной конференции «Системный синтез и прикладная синергетика». – 2024. – С. 144–149.

8. *Ковалев В. В., Сергеев Н. Е.* Разработка, исследование алгоритма обнаружения объектов малых размеров на основе сверточных нейронных сетей [Текст] // Информатизация и связь. – 2023. – С. 18–22. (ВАК, К2)

9. *Ковалев В. В.* Разработка и исследование алгоритма повышения точности обнаружения малоразмерных объектов сверточными нейронными сетями [Текст] // Сборник статей IX Всероссийской научно-технической конференции «Фундаментальные и прикладные аспекты компьютерных технологий и информационной безопасности». – 2023. – С. 185–188.

10. *Ковалев В. В.* Методы решения проблемы переобучения нейронных сетей в задаче обнаружения малоразмерных объектов на изображениях [Текст] // Труды международного научно-технического конгресса «Интеллектуальные системы и информационные технологии-2023». – 2023. – С. 38–45.

11. *Ковалев В. В.* Оптимизация гиперпараметров модели сверточной нейронной сети на основе предварительной обработки последовательности изображений [Текст] // Сборник статей VII Всероссийской научно-технической конференции «Фундаментальные и прикладные аспекты компьютерных технологий и информационной безопасности». – 2021. – С. 380–384.

12. *Ковалев В. В.* Алгоритм нелинейного контрастирования видеоизображений для повышения точности обнаружения объектов [Текст] // Материалы Всероссийской научно-технической конференции с международным участием имени профессора О.Н. Пьявченко «КомТех-2021». – 2021. – № 2. – С. 254–260.

13. *Ковалев В. В.* Диагностика работы сверточных нейронных сетей методом GRAD-CAM [Текст] // Сборник трудов XVIII Всероссийской научной конференции молодых ученых, аспирантов и студентов «ИТСАУ-2020». – 2020. – С. 14–18.

Личный вклад автора в работы, выполненные в соавторстве. В работах: [2, 5, 8, 10] – разработан алгоритм предварительной обработки изображений для предотвращения переобучения сверточных нейронных сетей и повышения качества распознавания малоразмерных изображений объектов; [3, 4] – представлены основные принципы построения сверточных нейронных сетей на вычислительных устройствах с ограниченным вычислительным ресурсом; [6] – предложены интегральные критерии для оценки качества сопровожения объектов; [7] – разработан алгоритм построения интеллектуальной распознающей системы.