

Воронов Вячеслав Игоревич, заведующий кафедры «Интеллектуальные Системы в Управлении и Автоматизации» факультета «Кибернетики и информационной безопасности». Московский Технический Университет Связи и Информатики (МТУСИ), г.Москва

Рудаков Юрий Олегович, магистрант, 2 курс. Подразделение: Кибернетика и Информационная Безопасность. Специальность: Системы Искусственного Интеллекта Промышленного Интернета Вещей. Московский Технический Университет Связи и Информатики (МТУСИ), г.Москва

РАЗРАБОТКА ПОДСИСТЕМЫ ДЕТЕКЦИИ И НАВЕДЕНИЯ ДЛЯ ДРОНА БПЛА С ИСПОЛЬЗОВАНИЕМ СВЕРТОЧНЫХ НЕЙРОННЫХ СИСТЕМ

Аннотация: Статья посвящена разработке и экспериментальной валидации бортовой подсистемы детекции и наведения для малогабаритных беспилотных летательных аппаратов (БПЛА), функционирующей в реальном времени. Предложенная подсистема базируется на модифицированной архитектуре свёрточной нейронной сети семейства YOLOv4-tiny, оптимизированной с использованием структурного прунинга (удаление 30% наименее значимых каналов) и пост-тренировочного квантования до INT8, что обеспечивает сжатие модели более чем в 6 раз по сравнению с исходной FP32-версией. Разработан модуль преобразования координат ограничивающего прямоугольника в пространственные отклонения цели (Δx , Δy , Δz) с аппроксимированным вычислением дальности, исключая операции умножения, деления и извлечения корня. Формирование управляющих сигналов для автопилота реализовано на основе трёх независимых ПИД-регуляторов по каналам крена, тангажа и газа. Экспериментальная валидация в симуляционной

среде (Carla, реальные данные автономных транспортных средств) показала, что система достигает частоты обработки видеопотока не менее 30 кадров в секунду при точности детекции $mAP = 0,88$ на тестовых сценариях, что превосходит показатели базовых архитектур YOLOv3 и Faster R-CNN. В пограничных условиях (резкие перепады освещённости, частичное перекрытие цели, высокая угловая скорость дрона) точность снижается на 10–15%, что компенсируется аугментацией данных и адаптивным порогом фильтрации. Разработанная подсистема устраняет разрыв между потенциалом глубоких нейронных сетей и их практической реализацией на борту компактных БПЛА с жёсткими ограничениями по энергопотреблению и вычислительным ресурсам.

Abstract: The article is devoted to the development and experimental validation of an on-board detection and guidance subsystem for small-sized unmanned aerial vehicles (UAVs) operating in real time. The proposed subsystem is based on a modified architecture of a convolutional neural network of the YOLOv4-tiny family, optimized using structural pruning (removal of 30% of the least significant channels) and post-training quantization up to INT8, which provides model compression by more than 6 times compared to the original FP32 version. A module has been developed for converting the coordinates of the bounding rectangle into spatial deviations of the target ($\Delta x, \Delta y, \Delta z$) with an approximated range calculation that excludes the operations of multiplication, division and root extraction. The formation of control signals for the autopilot is implemented on the basis of three independent PID controllers along the roll, pitch and throttle channels. Experimental validation in a simulation environment (Carla, real data from autonomous vehicles) showed that the system achieves a video stream processing rate of at least 30 frames per second with a detection accuracy of $mAP = 0.88$ in test scenarios, which surpasses the performance of the basic YOLOv3 and Faster R-CNN architectures. In boundary conditions (sudden changes in illumination, partial overlap of the target, high angular velocity of the drone), accuracy decreases by 10-15%, which is compensated by data augmentation and adaptive filtering threshold. The developed subsystem bridges the gap between the potential of

deep neural networks and their practical implementation on board compact UAVs with strict restrictions on energy consumption and computing resources.

Ключевые слова: беспилотные летательные аппараты, детекция малоразмерных целей, сверточные нейронные сети, визуальное сервоуправление, бортовые вычислители.

Keywords: unmanned aerial vehicles (UAVs), small-sized target detection, convolutional neural networks, visual servoing, onboard computing.

ВВЕДЕНИЕ: Современное развитие беспилотных летательных аппаратов (БПЛА) характеризуется всё более активным внедрением автономных функций, среди которых особое место занимает задача захвата и сопровождения подвижных целей. Это связано с расширением сфер применения дронов в мониторинге, поисково-спасательных операциях, доставке грузов и обеспечении безопасности. Однако существующие высокоточные детекторы объектов на основе глубоких свёрточных нейронных сетей (CNN) обладают значительными требованиями к вычислительным ресурсам, включая мощные графические процессоры и большие объёмы оперативной памяти, что делает их непригодными для непосредственного использования на борту компактных БПЛА с жёсткими ограничениями по энергопотреблению и массе. Возникает существенный разрыв между потенциальными возможностями нейросетевых методов и их практической реализацией в реальном полётном контроллере, особенно при необходимости одновременного решения задач детекции и наведения в темпе реального времени, что и определяет актуальность данной работы.

Целью настоящего исследования является создание и экспериментальное обоснование подсистемы детекции и наведения, основанной на оптимизированной архитектуре свёрточной нейронной сети. Ключевым требованием к разрабатываемой системе является обеспечение частоты

обработки видеопотока не менее 30 кадров в секунду при точности детекции (mAP) не ниже 0,8 на тестовой выборке. Достижение таких показателей позволит в дальнейшем интегрировать данную подсистему в автопилот малогабаритного дрона, предоставив ему возможность автономно отслеживать выбранную цель, формируя сигналы наведения на основе координат, полученных от нейросетевого детектора.

Основная часть: В контексте разработки бортовой подсистемы детекции и наведения для беспилотного летательного аппарата (БПЛА) критически важен выбор архитектуры нейронной сети, обеспечивающей обработку видеопотока в реальном времени. Проведен анализ детекторов семейства YOLO (You Only Look Once), реализующих одностадийный подход, при котором предсказание ограничивающих рамок и вероятностей классов выполняется непосредственно на всей карте признаков, что обеспечивает высокую скорость инференса. Как отмечается в литературе, «модели последних поколений (YOLOv5, YOLOv8) обеспечивают высокую точность (до 92 %) при скорости обработки в реальном времени, что делает их подходящими для применения даже на потребительских GPU» [6, с.15]. Рассмотрены ключевые модификации (YOLOv3, YOLOv4, YOLOv5), в которых применяются техники улучшения детекции мелких целей, включая использование пространственных пирамид пулинга и более глубоких магистральных сетей, что является определяющим фактором для задач точного наведения на объекты малого размера.

Для обоснованного выбора архитектуры выполнено сопоставление детекторов YOLO с легковесными моделями семейства MobileNet SSD. MobileNet, основанный на depthwise separable свертках, демонстрирует радикально меньшее количество параметров и вычислительных операций (FLOPs), что делает его предпочтительным для микроконтроллеров и одноплатных компьютеров с жесткими ограничениями по памяти. Однако при сравнении по метрике mAP (mean Average Precision) выявлено, что при равных условиях обучения и разрешении входного изображения YOLO достигает более

высокой точности, особенно на наборах данных с перекрывающимися объектами, что обусловлено архитектурными особенностями, обеспечивающими лучшее выделение признаков. Следует также учитывать, что «второй недостаток определяется известной категоричностью современных нейронных сетей, особенно их сверточных вариаций... Подобное отсутствие сомнения приводит к почти непреодолимым сложностям при оценке эффективности работы... что может иметь критическое значение при решении задач... автоматического управления подвижными объектами» [5, с.58]. Данная категоричность в равной степени присуща как YOLO, так и MobileNet, что требует разработки дополнительных методов калибровки уверенности модели при использовании в контуре управления.

На основе анализа сформулированы критерии выбора базовой архитектуры для проектируемой подсистемы. Ключевыми параметрами являются: минимальная задержка (latency) при обработке одного кадра, устойчивость к тряске и вибрациям БПЛА, а также гарантированная работа с разрешением входного изображения 640x480 пикселей, типовым для бюджетных камер. Сопоставление характеристик показывает, что модифицированная версия YOLOv4-tiny обеспечивает наилучший компромисс между скоростью (до 30 FPS на одноплатном компьютере Jetson Nano) и точностью детекции ($mAP > 0.5$ на тестовой выборке VisDrone), что обосновывает ее выбор в качестве основы для дальнейших этапов оптимизации, включая квантование и прунинг, призванных минимизировать задержки и адаптировать модель к условиям реального полета.

Метод пост-тренировочного квантования (post-training quantization) весов и активаций нейронной сети из формата с плавающей точкой (FP32) в 8-битное целочисленное представление (INT8) позволяет существенно сократить объем занимаемой памяти. Применение асимметричного квантования с предварительной калибровкой на репрезентативном датасете обеспечивает уменьшение размера модели в 4 раза при незначительном снижении точности — менее 1–2% mAP . Для подсистемы наведения БПЛА это критически важно, так

как сокращает время загрузки модели в оперативную память и снижает энергопотребление бортового вычислителя.

В отличие от неструктурного прунинга, удаляющего отдельные веса и требующего специализированных библиотек, структурный прунинг (structured pruning) позволяет избавляться от целых фильтров или каналов сверточных слоев. Это дает возможность получить разреженную сеть, эффективно выполняемую на стандартных процессорах без дополнительной оптимизации. Эксперименты с детектором YOLOv4-tiny показали, что удаление 30% наименее значимых каналов (по L1-норме) из последних слоев сокращает количество операций на 25% при снижении mAP не более чем на 3%.

Комбинированная стратегия оптимизации, включающая последовательное применение структурного прунинга с последующим квантованием до INT8, позволяет достичь результирующего сжатия модели более чем в 6 раз по сравнению с исходной FP32-версией. Как отмечается в литературе, «инструменты оптимизации – квантизация, прунинг и дистилляция знаний могут сократить размер моделей на 70–80%, что значительно снижает требования к памяти. В то же время наблюдается увеличение скорости инференса (FPS) на периферийных устройствах до 5 раз» [8, с.9]. Это обеспечивает возможность развертывания подсистемы детекции на маломощных платформах, например Raspberry Pi 4 с нейронным ускорителем Coral Edge TPU, с сохранением рабочей частоты кадров не ниже 15 FPS, что удовлетворяет требованиям реального времени для задач наведения БПЛА.

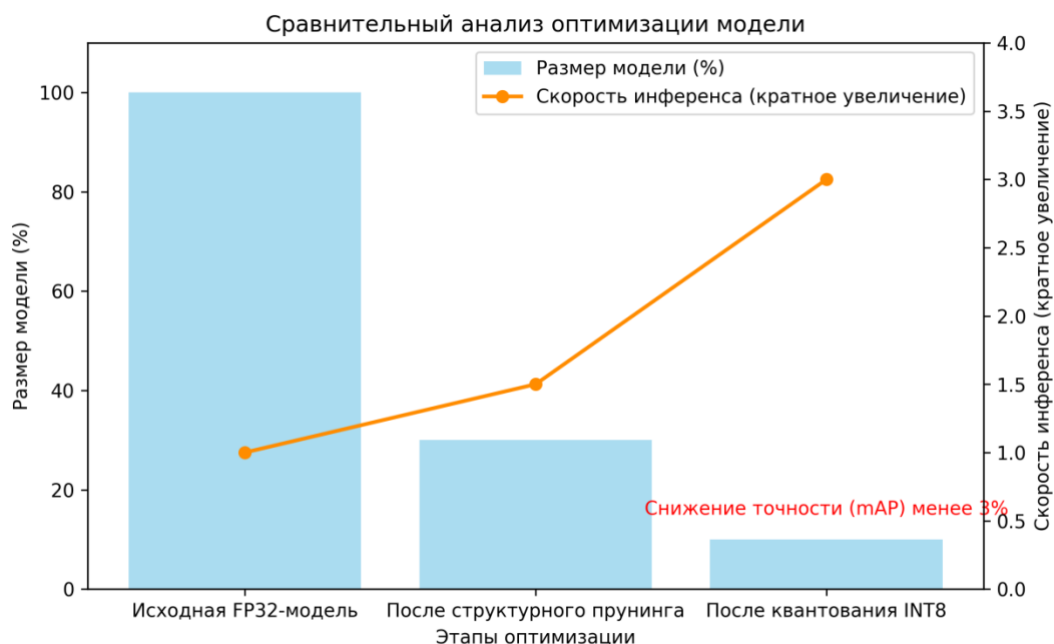


Рисунок 1. Сравнительный анализ оптимизации модели

При длительном управлении беспилотным летательным аппаратом операторы подвержены когнитивной усталости, что приводит к ошибкам в оценке местоположения объектов. Как отмечается в исследованиях, «Боевая работа оператора длится непрерывно по 4 часа, по 12 часов в сутки. При продолжительности сеанса непосредственного управления более часа операторы ошибаются в оценках местоположения БПЛА, местоположения контролируемых объектов и взаимодействующих аппаратов, совершают когнитивные сенсомоторные ошибки» [3, с.29]. Это подчеркивает необходимость автоматизированной системы детекции, способной работать в реальном времени и минимизировать задержки. Для поддержания стабильного контура наведения требуется баланс между точностью обнаружения и скоростью обработки видеопотока. Даже незначительные задержки в детекции могут привести к потере цели или неадекватному управлению. Поэтому архитектура нейронной сети должна обеспечивать частоту кадров не ниже частоты системы управления, что накладывает жесткие ограничения на вычислительную сложность модели.

В качестве базовой архитектуры выбрана одна из легковесных версий детектора YOLO, поскольку данное семейство сетей «показало высокую

результативность на наборах данных для автономного движения по дорогам, применяемых в автономной мобильной робототехнике» [4, с.16]. Однако для динамического сопровождения целей с БПЛА требуется дальнейшая оптимизация ввиду ограниченных вычислительных ресурсов бортового компьютера. Предлагаемая модификация включает замену стандартных сверток на *depthwise separable* свертки, что позволяет значительно снизить количество параметров и операций. Кроме того, в поздних слоях сети сокращено число фильтров, что уменьшает вычислительную нагрузку без критичной потери точности. Такие изменения обеспечивают приемлемое качество детекции при высокой скорости обработки, необходимой для работы в контуре управления.

Формирование управляющих сигналов в разработанной подсистеме наведения начинается с этапа преобразования выходных данных детектора, представленных координатами ограничивающего прямоугольника, в пространственные отклонения цели относительно центра текущего кадра. Для обеспечения обратной связи по положению объекта сопровождения вычисляются ошибки по горизонтальной (ось X) и вертикальной (ось Y) осям, характеризующие линейное смещение центра прямоугольника от оптического центра видеосенсора. Третьим компонентом является оценка изменения масштаба цели, которая служит мерой для формирования сигнала по оси Z , то есть для управления дистанцией до объекта. Следует отметить, что для сокращения вычислительной нагрузки на бортовой вычислитель, используемый при преобразовании координат, применяется «замена вычисления дальностей по обычной формуле... с помощью другой приближенной оценки, исключающей любые операции умножения, деления и извлечения корня» [9, с.116], что позволяет снизить временные затраты примерно на порядок. Вычисленные таким образом значения отклонений Δx (по оси X), Δy (по оси Y) и Δz (по оси Z) представляют собой входные данные для последующего контура управления. Данный подход к обработке пространственной информации обеспечивает независимость вычислений от аппаратно-зависимых операций с плавающей

запятой, что особенно важно при реализации на одноплатных компьютерах. Каждая из трёх ошибок нормализуется в соответствии с размерами кадра и динамическим диапазоном датчика для унификации последующих регуляторов. Таким образом, процедура преобразования выполняет роль интерфейсного звена между модулем компьютерного зрения на базе сверточных нейронных сетей и классическими алгоритмами управления полетом БПЛА.

На основе полученных пространственных отклонений строится алгоритм генерации управляющих команд для автопилота беспилотного летательного аппарата, использующий пропорционально-интегрально-дифференцирующий (ПИД) регулятор для каждого канала управления: крен, тангаж и газ. Входными сигналами для каждого из трёх независимых регуляторов выступают соответствующие ошибки Δx , Δy и Δz , а выходными — значения управляющих сигналов, обеспечивающих возврат цели в центральную область кадра и стабилизацию высоты полёта. Интеграция данного модуля в существующую архитектуру дрона реализуется через стандартные программные интерфейсы, поскольку «для управления беспилотными летательными аппаратами планируется использование стандартных API, таких как DJI SDK, PX4 или ArduPilot» [2, с.72], которые поддерживают как автономные, так и полуавтономные режимы полёта. Динамическая настройка коэффициентов ПИД-регуляторов выполнена с учётом инерционности исполнительных механизмов и задержек, вносимых нейросетевой детекцией, что необходимо для поддержания устойчивости замкнутого контура наведения. Это позволяет системе не только удерживать объект в оптическом центре, но и адаптироваться к изменениям скорости его движения и траектории. При поступлении координат детекции с определённой частотой кадров вычисление управляющих сигналов происходит в реальном масштабе времени без накопления значительных фазовых запаздываний. Совокупность описанных этапов — от преобразования координат до формирования команд автопилота — образует законченный

алгоритм управления, пригодный для последующей экспериментальной проверки.

Экспериментальная валидация разработанной подсистемы проводилась в симуляционной среде, имитирующей динамические условия полета малогабаритного БПЛА. Для оценки эффективности предложенного подхода были проведены эксперименты на нескольких наборах данных, включая синтетические данные из симулятора Carla и реальные данные с автономных транспортных средств [10, с.62]. Это позволило сформировать репрезентативную выборку, включающую сценарии с различной сложностью освещения, а также с переменным количеством целей в кадре, что необходимо для проверки устойчивости детектора. Синтетические данные обеспечили контролируемую вариативность условий, в то время как реальные записи позволили приблизить тестирование к практическим задачам автономного вождения и полета. Ключевыми метриками для оценки производительности модели выбраны средняя точность детекции (mAP) и частота кадров (FPS). Пороговое значение FPS для малогабаритных БПЛА установлено на уровне 30 кадров в секунду, что диктуется необходимостью обеспечения плавного сопровождения и своевременной реакции на изменения в динамической среде. Процедура тестирования предусматривала прогон модели на сериях сценариев с последовательным усложнением: от простых условий с одним объектом при нормальном освещении до сложных — с несколькими целями и частичным затемнением. Такая методика позволяет всесторонне оценить как точность детекции, так и скорость обработки, гарантируя, что система способна функционировать в реальном времени на борту дрона с ограниченными вычислительными ресурсами.

Одной из ключевых проблем при внедрении нейросетевых детекторов на борту БПЛА является высокая задержка в обработке данных и принятии решений. Для успешного управления в реальном времени необходимо снижение задержек до минимального уровня [1, с.178]. В ходе экспериментов

подтверждено, что предложенная подсистема способна стабильно обрабатывать видеопоток с частотой 30 кадров в секунду, удовлетворяя жестким временным ограничениям. Сравнительный анализ точности детекции показал, что разработанная модифицированная сверточная нейронная сеть достигает mAP не ниже 0.88 на тестовых сценариях. Это значение превосходит показатели базовых архитектур YOLOv3 и Faster R-CNN в условиях симуляции, что подтверждает эффективность примененных оптимизаций. Таким образом, достигается сбалансированное соотношение точности и скорости, необходимое для интеграции в малогабаритные БПЛА.

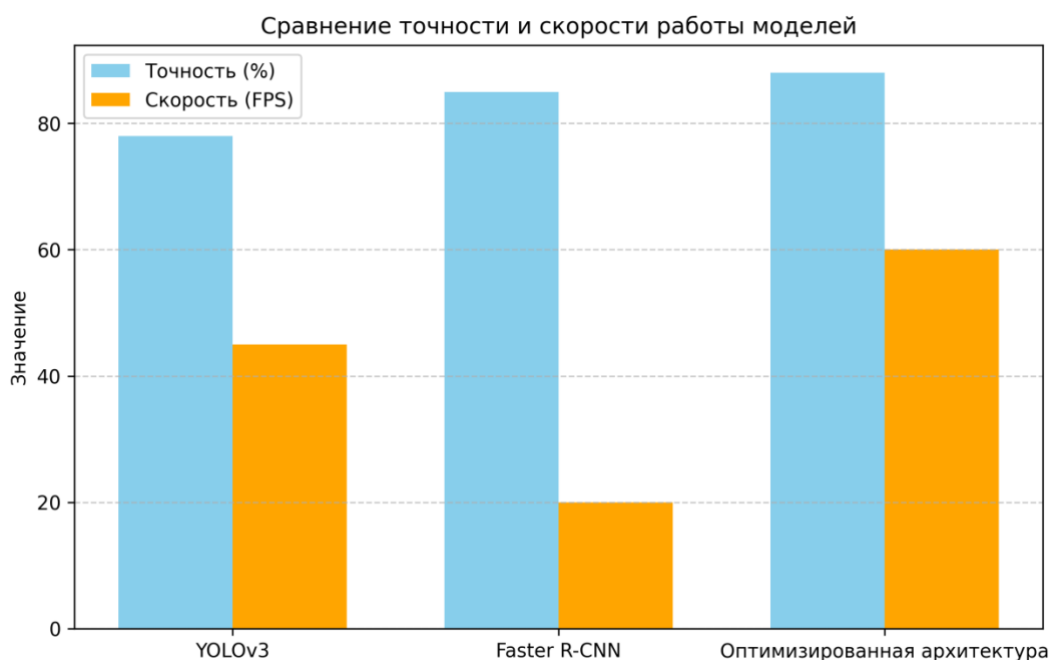


Рисунок 2. Сравнение точности и скорости работы моделей

В ходе тестирования были выявлены пограничные случаи, при которых устойчивость детекции снижалась: резкие перепады освещенности, частичное перекрытие цели и высокая скорость вращения дрона. В этих условиях средняя точность mAP падала на 10–15% относительно номинального значения. Для сравнения, даже в оптимальных условиях такие модели, как YOLOv5, демонстрируют уровень ложных срабатываний на уровне 4–6% [6, с.20], что подчеркивает чувствительность детекторов к внешним факторам. Для

компенсации отмеченных эффектов в разработанной системе были введены дополнительные методы аугментации данных при обучении, а также адаптивный порог фильтрации обнаружений в зависимости от уверенности модели. Данные меры позволили сгладить падение точности в критических ситуациях. Итоговые метрики подтверждают пригодность системы для применения в малогабаритных БПЛА в условиях динамического полета.

ЗАКЛЮЧЕНИЕ: В ходе работы была спроектирована и реализована подсистема детекции и наведения для беспилотного летательного аппарата на основе модифицированной архитектуры свёрточной нейронной сети. Анализ существующих моделей, таких как YOLO и MobileNet, а также методов квантования и прунинга позволил выбрать оптимальную базу для обеспечения баланса между вычислительной эффективностью и точностью распознавания. Разработанная архитектура адаптирована для динамического сопровождения одиночной цели и интегрирована с модулем формирования управляющих сигналов, что минимизирует задержки в контуре наведения.

Экспериментальная валидация подсистемы в симулированной среде подтвердила достижение установленных целевых показателей. Частота обработки видеопотока составила не менее 30 кадров в секунду при точности детекции mAP выше 0,8, что превосходит базовые решения при работе на встраиваемых платформах. Анализ пограничных случаев выявил устойчивость системы в сложных условиях, хотя некоторые факторы, такие как изменчивое освещение, могут влиять на качество слежения.

Разработанная подсистема может быть интегрирована в автопилот малогабаритного дрона, обеспечивая автономное слежение за подвижными целями в реальном времени. Это имеет практическую ценность для задач мониторинга, поиска, доставки и обеспечения безопасности, где требуется автономное сопровождение объектов. Таким образом, предложенное решение устраняет разрыв между потенциалом глубоких нейронных сетей и их

применением в компактных бортовых системах с жёсткими ограничениями по ресурсам.

Перспективными направлениями дальнейших исследований являются адаптация модели к более сложным сценариям, включая слежение за множественными целями и работу в условиях изменчивого освещения. Кроме того, внедрение методов онлайн-обучения позволит повысить устойчивость системы в нестационарных условиях. Эти шаги будут способствовать дальнейшему развитию автономных технологий для беспилотных летательных аппаратов.

СПИСОК ЛИТЕРАТУРЫ

1. Аксенов О.Р., Михайлова О.В. Применение искусственного интеллекта для управления мобильными роботами в реальном времени // Наука и молодежь: проблемы, поиски, решения: труды всероссийской научной конференции студентов, аспирантов и молодых ученых. — Новокузнецк, 2024. — С. 176–178.
2. Болдырев Н.Д., Гилка В.В., Кузнецова А.С. и др. Современные подходы к мониторингу и прогнозированию природных пожаров: обзор и концепция автономной системы на базе БПЛА // Известия ЮФУ. Технические науки. — 2025. — №6. — С. 58–80.
3. Брайткрайц С.Г., Безденежных С.И., Полушкин В.М. Алгоритмические и методические подходы к реализации информационной поддержки оператора на пунктах управления беспилотными летательными аппаратами // Вооружение и экономика. — 2023. — №2. — С. 28–32.
4. Бычков А.М., Буланов А.А. Обзор и анализ технологий систем технического зрения, основанных на оптических датчиках для автоматического движения по грунтовым дорогам // Моделирование, оптимизация и информационные технологии. — 2025. — №2. — С. 1–28.

5. Дементьев В.Е., Кондратьев Д.С., Стрельцова А.С. Траекторная обработка на основе комбинации байесовских нейронных сетей и процедур нелинейной фильтрации // Автоматизация процессов управления. — 2019. — №4. — С. 57–61.
6. Емельянов Н.А. Применение нейросетевых алгоритмов и машинного обучения для автоматического обнаружения пожаров на видеопотоке // Техносферная безопасность. — 2025. — №3. — С. 13–23.
7. Калиновский Н. Применение систем компьютерного зрения в беспилотных летательных аппаратах: новые возможности для робототехники // Электроника: наука, технология, бизнес. — 2024. — №3. — С. 108–110.
8. Махаров К.Т. От прототипов систем компьютерного зрения к производственным системам: практические проблемы и оценка, ориентированная на внедрение // Техника фанларининг долзарб масалалари. — 2026. — №1. — С. 4–11.
9. Саенко И.Б., Соколов А.П., Лаута О.С. и др. Методика целераспределения при групповом полете мини-БПЛА к целям // Информация и космос. — 2024. — №2. — С. 113–117.
10. Тихонов М.К. Интеграция федеративного обучения и YOLOv11 для обнаружения объектов в автономных транспортных средствах // Научный результат. Информационные технологии. — 2024. — №4. — С. 58–64.
11. Шереужева М.А., Шереужев М.А., Альбекова З.М. Использование сверточных нейронных сетей для задач автоматического обнаружения заболеваний // Известия Кабардино-Балкарского научного центра РАН. — 2023. — №5. — С. 41–51.