

Раздел IV. Техническое зрение

УДК 004.932.2

DOI 10.18522/2311-3103-2026-2-250-258

С.Е. Бабин, П.А. Гессен, В.А. Павлова, В.А. Тупиков

ПОВЫШЕНИЕ УСТОЙЧИВОСТИ ДЕТЕКЦИИ СИАМСКИХ СЕТЕЙ ПРИ ОГРАНИЧЕННОЙ РАЗМЕТКЕ С ОТБОРОМ ТРУДНЫХ ПРИМЕРОВ

Цель работы – разработка и экспериментальная оценка метода повышения качества детекции морских судов на основе архитектуры Attention RPN при малом количестве обучающих данных с интеграцией межклассового выбора трудных отрицательных примеров. Для достижения цели реализованы: адаптация модели в режиме тонкой настройки на основе пяти примеров, формирование специализированного датасета из 9 категорий судов (637 обучающих и 175 валидационных изображений), а также алгоритм отбора трудных отрицательных пар по косинусной мере сходства. Методологическую основу составили подходы глубокого обучения для детекции объектов на основе сиамских нейронных сетей, анализ распределений признаков в пространстве и алгоритмы создания трудных отрицательных примеров с вычислением межклассовых расстояний на основе косинусной метрики. Экспериментальная оценка выполнена на стандартных бенчмарках для обнаружения объектов при малом количестве данных. Внедрение межклассового выбора трудных отрицательных примеров позволило сократить долю ложноположительных детекций на ~18% и повысить показатель mAP в среднем на ~6% по сравнению с базовой моделью без отбора трудных негативных примеров. Наибольший прирост (до 9,1%) зафиксирован для классов с крупными объектами интереса. Полученные результаты демонстрируют, что целенаправленный отбор межклассовых трудных отрицательных примеров существенно повышает устойчивость метрико-ориентированных детекторов в условиях сильного дисбаланса и высокой внутриклассовой схожести. Практическая значимость работы заключается в возможности применения предложенного подхода для развертывания систем видеонаблюдения в задачах при отсутствии большой обучающей выборки и минимальном объеме разметки.

Сиамские нейронные сети; глубокое обучение; трудные отрицательные примеры.

S.E. Babin, P.A. Gessen, V.A. Pavlova, V.A. Tupikov

IMPROVING THE ROBUSTNESS OF SIAMESE NETWORK DETECTION WITH LIMITED ANNOTATION THROUGH HARD EXAMPLE MINING

The aim of this study is to develop and experimentally evaluate a method for improving the quality of marine vessel detection based on the Attention RPN architecture under conditions of limited training data, with the integration of inter-class hard negative mining. To achieve this goal, the following were implemented: model adaptation in a fine-tuning mode based on five examples, creation of a specialized dataset consisting of 9 vessel categories (637 training and 175 validation images), and an algorithm for selecting hard negative pairs using cosine similarity. The methodological foundation includes deep learning approaches for object detection based on Siamese neural networks, analysis of feature distributions in embedding space, and algorithms for generating hard negative samples by computing inter-class distances using the cosine metric. Experimental evaluation was carried out on standard few-shot object detection benchmarks. The integration of cross-class hard negative mining reduced the proportion of false positive detections by approximately 18% and increased the mean Average Precision (mAP) by about 6% on average compared to the baseline model without hard negative mining. The greatest improvement (up to 9.1%) was observed for classes with large target objects. The results demonstrate that targeted selection of inter-class hard negative examples significantly improves the robustness of metric-based detectors under conditions of strong class imbalance and high intra-class similarity. The practical significance of the work lies in the applicability of the proposed approach for deploying video surveillance systems in tasks where large training datasets and extensive annotation are unavailable.

Siamese neural networks; deep learning; hard negative mining.

Введение. Системы видеонаблюдения стали неотъемлемой частью современного мира и активно применяются в различных сферах жизни, таких как: обеспечение безопасности [1], промышленное производство [2, 3], транспорт [4, 5], здравоохранение [6, 7] и других областях. Масштабное использование камер и устройств видеонаблюдения привело к необходимости автоматизированного анализа большого количества поступающей информации в реальном времени. Эту проблему активно решают методы компьютерного зрения, получая и анализируя значимую информацию из сплошного потока данных. Одними из ключевых задач компьютерного зрения являются задачи обнаружения и классификации объектов интереса.

В последнее время был достигнут существенный прогресс в упомянутых задачах. С одной стороны он обусловлен ростом вычислительных мощностей и связанным с этим развитием глубоких нейронных сетей, таких как Faster R-CNN [8], семейство YOLO [9, 10], нейросети типа трансформер [11, 12], которые демонстрируют высокую точность, а порой и скорость работы. Современные архитектуры демонстрируют высокие показатели точности и устойчивости к вариативности сцены, освещенности и ракурсов. Улучшение метрик качества стало возможным благодаря усложнению моделей и использованию продвинутой стратегии обучения. С другой стороны это развитие стало возможным благодаря созданию больших наборов данных (датасетов) с уже размеченными экземплярами объектов, таких как MS COCO [13], PASCAL VOC [14]. И это же стало одной из проблем: достигнутый прогресс опирается на наличие значительных объемов размеченных данных, подготовка которых требует существенных временных затрат. В реальных ситуациях, особенно при появлении новых классов объектов или специфических условий съемки, собрать репрезентативную обучающую выборку зачастую невозможно. Зависимость современных моделей от больших датасетов становится ключевым ограничением их использования, что обуславливает необходимость разработки методов, обеспечивающих высокое качество распознавания при ограниченном количестве аннотационных данных.

Задача детекции объектов при ограниченном объеме размеченных данных (Few-Shot Object Detection, FSOD [15]) направлена на обход проблемы большой выборки данных и рассматривает модели, способные обнаруживать и распознавать новые виды объектов при наличии лишь нескольких размеченных примеров для каждого класса (обычно 1–10 изображений). Такие модели предварительно обучаются на базовых классах с достаточным объемом данных, после чего адаптируются к новым классам, для которых доступно лишь ограниченное число примеров без полного переобучения сети.

Одним из представителей данной группы методов являются сиамские нейронные сети – это архитектуры, предназначенные для обучения функции сходства между объектами. Их ключевая особенность заключается в использовании двух (или более) идентичных подсетей с общими весами, которые обрабатывают разные входы, отображая их в общее пространство признаков. Основная концепция состоит не в прямой классификации, а в минимизации расстояния между схожими объектами и увеличении расстояния между различными, что делает их отличным выбором в условиях ограниченного количества данных.

Предлагаемое решение. В данной работе исследуется архитектура сети внимания [16], представленная в библиотеке mmfewshot [17]. Ключевой особенностью метода является интеграция механизма внимания в сеть обнаружения регионов интереса, что позволяет использовать признаки нового изображения интереса (support) для фильтрации предполагаемых объектов на основном изображении (query). Эксперименты проводятся в рамках стратегии few-shot finetuning, предполагающей предварительное обучение на базовых классах с последующей донастройкой на целевых классах при ограниченном числе размеченных примеров.

В данной работе авторами предлагается рассмотреть работу описанной сети в рамках распознавания объектов интереса на целевом наборе изображений морских судов и улучшить качество обнаружения и при помощи добавления механизма межклассовых трудных отрицательных примеров.

Стратегия обучения

Основные моменты по архитектуре модели описаны в оригинальной статье [16], нам же для понимания работы необходимо разобрать как работает обучение сети для дальнейшего улучшения подаваемого набора данных. Для обучения модели применяется стратегии двустороннего контрастивного обучения. Обучающая выборка формируется в виде триплетов (q_c, s_c, s_n) , где q_c – изображение-запрос с объектом класса c ; s_c – положительный образец того же класса c , а s_n – отрицательный образец класса n ($n \neq c$). Такой подход учит модель не только находить соответствия между объектами одного класса, но и различать разные классы. Функция потерь L , вычисляемая для каждого предложения, определяется как сумма функции потерь сопоставления и функции потерь регрессии описывающего прямоугольника:

$$L = L_{\text{matching}} + L_{\text{box}},$$

где L_{matching} – бинарная кросс-энтропия. Она штрафует модель, если оценка сходства низкая для положительных пар (query и support одного класса) или высокая для отрицательных пар (разные классы); L_{box} – функция потерь регрессии описывающего прямоугольника, аналогичная используемой в Fast R-CNN [18]. Она применяется только к положительным парам для уточнения координат объекта.

Для компенсации вклада различных типов пар в обучение применяется отбор данных с поддержанием соотношения 1:2:1 между следующими группами:

- положительные предложения + положительный образец;
- фоновые предложения + положительный образец;
- любые предложения + отрицательный образец (negative support pairs).

Сеть внимания и многорегиональный детектор обучаются совместно, а общая функция потерь объединяет компоненты обоих модулей.

Сам процесс обучения делится на два этапа:

- ♦ предварительное обучение – модель обучается на полных данных базовых классов. На этом этапе настраиваются веса основополагающей сети и модулей сети внимания для извлечения общих признаков объектов и формирования метрики сходства;
- ♦ дообучение на малом числе примеров – проводится тонкая настройка модели на сбалансированной выборке, включающей как базовые, так и новые классы, где каждая категория представлена лишь ограниченным числом примеров.

Отбор трудных отрицательных примеров (Hard Negative Mining, HNM)

Это метод обучения моделей, при котором особое внимание уделяется таким примерам, которые не являются искомым объектом, но ошибочно классифицируются как положительный результат. Вместо того чтобы случайно выбирать отрицательные примеры, HNM отбирает регионы или объекты, наиболее похожие на целевой класс, но не являющиеся им, и использует их для усиления градиентов при обучении. Общая последовательность действий заключается в следующих этапах:

- ♦ для каждого конкретного класса формируются потенциально близкие отрицательные пары путем выбора визуально схожих объектов других классов;
- ♦ извлекаются их признаки с помощью обученной основной сети;
- ♦ измеряется сходство между признаками целевого класса и отрицательного примера путем использования косинусного расстояния;
- ♦ наиболее близкие примеры включаются в обучающий набор с сохранением метки о том, что данный пример не принадлежит интересующему классу.

Эксперименты

Описание датасета

Для оценки эффективности использования архитектуры сети внимания на задачах специализированной предметной области был сформирован небольшой датасет, собранный из открытых источников. Датасет содержит 9 категорий судов: Passenger ship (пассажирское судно), Unidentified (неопознанное судно), bulk-ship (балкер), container-ship (контейнеровоз), military-ship (военное судно), motor-ship (моторное судно), sailing-ship (парусное судно), tanker-ship (танкер) и small motor-ship (малое моторное судно). Обучающая выборка включает 637 изображений, а валидационная – 175 изображений.

Особенностью данного датасета является существенный дисбаланс классов (рис. 1). Наиболее многочисленным классом выступает *tanker-ship* (395 объектов в сумме), за ним следуют *sailing-ship* (383) и *small motor-ship* (267). Наименее представлены *motor-ship* (8 объектов), *Passenger ship* (18) и *bulk-ship* (31), что приводит к разнице между максимумом и минимумом почти в 50 раз.

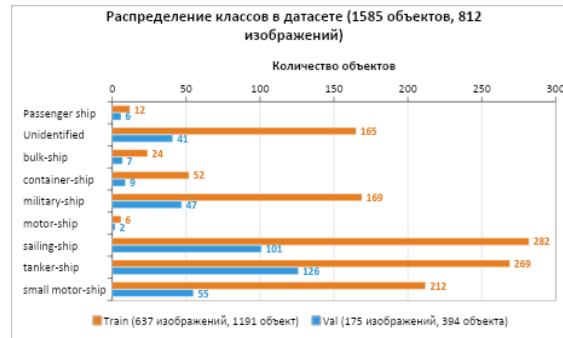


Рис. 1. Распределение объектов по классам в датасете

Датасет ориентирован на задачу *few-shot* обучения: для каждой из 9 категорий подготовлено по 5 *support*-изображений, используемых для инициализации модели. Такой дисбаланс отражает типичные условия реальных приложений, где данные по категориям распределены неравномерно. Особый интерес представляет обработка малочисленных классов (*motor-ship* – 8 объектов, *Passenger ship* – 18), позволяющая проверить обобщающую способность метода при минимальном объеме примеров.

Кроме количественного дисбаланса, датасет характеризуется высоким визуальным сходством категорий. В отличие от бенчмарков вроде PASCAL VOC или MS COCO, где новые классы существенно различаются (например, *bird*, *bus*, *cow*, *sofa*), здесь все категории относятся к общему классу морских судов и различаются в основном формой, размерами и пропорциями корпуса.

Детали реализации

Эксперимент проводился в режиме тонкой настройки на основе пяти примеров на основе модели сети внимания с базовой сверточной сетью извлечения признаков ResNet-50 [19] конфигурации C4, предварительно обученной на базовых классах MS COCO. Замороженными оставались все слои базовой сверточной сети; дообучению подвергались модули сети внимания и многореляционного детектора.

На каждой итерации обучения модель получала два типа входных данных: *query*-изображения, на которых нужно найти объекты, и наборы *support*-изображений – эталонных примеров, показывающих модели, как выглядят объекты разных классов. Размер батча составлял 2: каждый батч содержал 2 *query*-изображения, для каждого из которых формировался свой набор *support*-примеров. Из 9 классов датасета для каждого *query* случайным образом выбирались 6, для каждого из которых подавалось по 5 эталонных изображений. Таким образом, каждое *query*-изображение сопровождалось 30 *support*-примерами: 5 – с примерами целевого класса и 25 – с примерами пяти других классов, что позволяло модели учиться отличать нужный класс от остальных. Общее число итераций составило 3000.

Оптимизация проводилась методом стохастического градиентного спуска [20] с начальной скоростью обучения $\text{lr} = 0.001$, импульсом 0.9 и затуханием весов 10^{-4} . Для головы детектора применялся удвоенный множитель скорости обучения, что ускорило адаптацию классификационных и регрессионных весов многореляционного детектора.

Финальная оценка произведена на полном валидационном наборе (175 изображений). Метрики рассчитаны по протоколу COCO: $\text{mAP@}[\cdot 50:95]$, AP50, AP75, а также AP по размерам объектов (малый, средний, большой).

Результаты. Кривые функции потерь представлены на рис. 2 и 3. Суммарная потеря уменьшилась с 11.37 на начальных итерациях до 0.18 на 3000-й, что соответствует снижению в 62 раза. Наиболее сильное падение произошло в фазе разогрева (первые 200 итераций), после чего потери убывали монотонно. Функция потерь сети внимания стабилизировались к 300-й итерации, указывая на быструю адаптацию механизма внимания к новым классам.



Рис. 2. Суммарная функция потерь с EMA-сглаживанием



Рис. 3. Компоненты функции потерь

В ходе обучения валидация проводилась на подмножестве из 40 изображений каждые 500 итераций (рис. 4). Динамика метрик была немонотонной: AP50 снизился с 35.8% (итерация 500) до 31.4% (итерация 1000), что может быть обусловлено стохастичностью few-shot обучения и малым размером валидационной выборки (40 изображений). В интервале 1000-2000 итераций наблюдалось постепенное восстановление метрик, после чего снижение $\lg$ на итерации 2000 обеспечило рост: AP50 увеличился с 34.0% (итерация 2000) до 43.9% (итерация 3000), т.е. на +9.9 п.п. Та же ситуация была с $mAP@[.50:.95]$: значение выросло с 21.6% (итерация 2000) до 26.1% (итерация 3000), +4.5 п.п. Это подтверждает эффективность ступенчатого уменьшения скорости обучения для более точной адаптации модели в режиме few-shot.

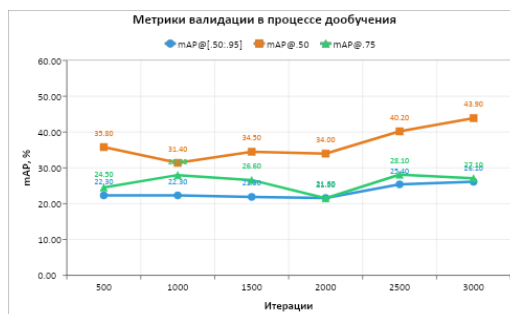


Рис. 4. Метрики валидации ($mAP@[.50:.95]$, AP50, AP75) в процессе дообучения

После завершения обучения была проведена оценка модели на полном валидационном наборе из 175 изображений (394 аннотации) с использованием стандартных метрик COCO. Результаты приведены в табл. 1.

Таблица 1

Результаты детекции на полном валидационном наборе

Метрика	Значение
$mAP@[.50:.95]$	28.7%
AP@.50	46.6%
AP@.75	32.4%
AP (small)	0.0%
AP (medium)	31.9%
AP (large)	32.7%
AR@100	35.5%

Расширение валидационной выборки с 40 до 175 изображений привело к повышению метрик (mAP: +2.6 п.п., AP50: +2.7 п.п.), что свидетельствует об устойчивости модели и отсутствии переобучения на ограниченном валидационном подмножестве.

На рис. 5 представлено распределение mAP по размерам объектов на каждом этапе промежуточной валидации. Малые объекты (площадь меньше 1024 пикселей) не обнаруживаются моделью на протяжении всего обучения ($AP_s = 0\%$). Это обусловлено архитектурой ResNet-50 конфигурации C4, которая извлекает признаки с единственного уровня карты признаков, что недостаточно для детекции малых объектов. На полном валидационном наборе (175 изображений) средние объекты достигают $AP_m = 31.9\%$, крупные – $AP_l = 32.7\%$, что подтверждает эффективность применения модели для объектов среднего и крупного масштаба.

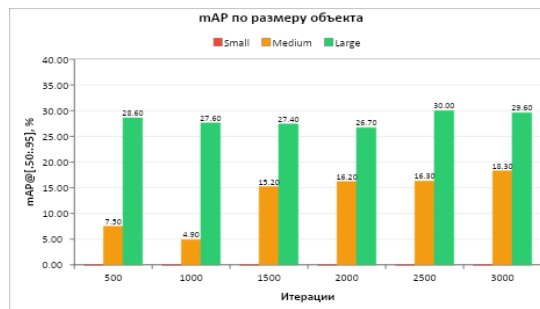


Рис. 5. $mAP@[.50:.95]$ по размеру объекта на каждом этапе валидации

На рис. 6 представлены результаты запуска обученной сети на изображениях валидационной выборки с порогом уверенности 0.7. Модель успешно локализует крупные суда различных типов, однако наблюдаются ошибки при разграничении визуально сходных классов (tanker-ship и small motor-ship).

В табл. 2 приведено сравнение полученных результатов с официальными тестами Attention RPN на стандартных датасетах. Источником данных являются результаты, опубликованные в mmfewshot [17].

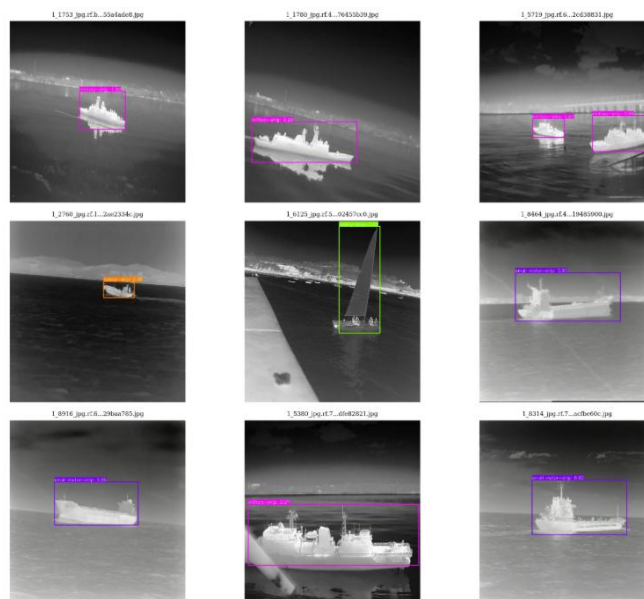


Рис. 6. Примеры обнаружения и классификации

Таблица 2

Сравнение результатов few-shot детекции сети внимания

Метод	Датасет	Novel	Shot	AP50, %	mAP, %
Attention RPN	MS COCO	20	10	–	9.2
Attention RPN	MS COCO	20	30	–	14.8
Attention RPN	VOC Split 1	5	5	51.7	
Attention RPN	VOC Split 2	5	5	37.0	
Attention RPN	VOC Split 3	5	5	48.9	
Attention RPN	Морские суда	9	5	46.6	28.7

Полученное значение $mAP@[.50:.95] = 28.7\%$ при конфигурации с пятью примерами значительно превосходит тесты разработчиков на наборе MS COCO (9.2% при десяти примерах и 14.8% при тридцати), что соответствует превышению в 3.1 раза при вдвое меньшем числе обучающих примеров (5 против 10). Данное различие объясняется меньшим числом целевых категорий (9 против 20) и специализированным характером предметной области, в которой модель может опираться на общие признаки класса «судно», частично сформированные при обучении на базовых классах COCO (в частности, категория boat входит в число базовых).

При сравнении по метрике AP50 с результатами на PASCAL VOC, полученное значение 46.6% превосходит наиболее сложное разбиение (Split 2: 37.0%, novel-классы: aeroplane, bottle, cow, horse, sofa) на 9.6 п.п., превышает среднее по всем трем разбиениям (45.9%) и приближается к результату Split 3 (48.9%), в котором novel-классом является boat – ближайшая категория к нашей предметной области. Split 2 считается наиболее сложным разбиением ввиду визуального сходства между отдельными novel- и base-классами (cow/horse), что аналогично ситуации в нашем датасете, где различные типы судов обладают высокой степенью визуального сходства.

После применения описанного ранее метода HNM на представленном наборе данных были получены результаты, представленные в табл. 3.

Таблица 3

Результаты после улучшения

Метрика	Значение
$mAP@[.50:.95]$	32.6%
$AP@.50$	53.9%
$AP@.75$	39.7%
AP (small)	0.0%
AP (medium)	38.9%
AP (large)	41.8%
AR@100	42.5%

Таким образом внедрение HNM позволило сократить долю ложноположительных детекций на ~18% и в среднем повысить показатель mAP ~6,4% по сравнению с базовой моделью без отбора трудных негативных примеров.

Заключение. Проведенное исследование подтвердило, что применение архитектуры Attention RPN эффективно адаптируется к задаче few-shot детекции объектов, в том числе морских судов, при выраженном межклассовом сходстве и дисбалансе данных.

Обучение в режиме тонкой настройки на пяти примерах на специализированном датасете из 9 категорий позволило получить базовое значение $mAP@[.50:.95]=28,7\%$ и $AP50=46,6\%$ на полном валидационном наборе, что свидетельствует о способности модели формировать устойчивое пространство признаков даже при ограниченном числе опорных примеров.

Интеграция механизма межклассового hard-negative mining продемонстрировала количественно измеримый вклад в повышение качества детекции: $mAP@[.50:.95]$ увеличился до 32,6%, $AP50$ – до 53,9%, $AR@100$ – до 42,5%, а доля ложноположительных срабатываний сократилась приблизительно на 18%. Наиболее выраженный прирост наблюдался для объектов среднего и крупного масштаба ($AP_m=38,9\%$, $AP_l=41,8\%$), что подтверждает эффективность целенаправленного отбора трудных отрицательных примеров при высокой визуальной схожести классов. Отсутствие детекции малых объектов ($AP_s=0\%$) выявило архитектурное ограничение конфигурации C4 и обозначило направление дальнейшей модернизации модели.

Предложенный подход обеспечивает повышение точности без увеличения объема разметки и может быть применен при разработке прикладных решений мониторинга морской инфраструктуры и судоходства в условиях ограниченных обучающих данных. В рамках дальнейшего развития работы стоит применить иные стратегии отбора данных или изменить аппарат получения признаков объектов интереса на более совершенный.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Zhao J., Masood R., Seneviratne S. A Review of Computer Vision Methods in Network Security, *IEEE Communications Surveys & Tutorials*, 2021, Vol. PP, pp. 1-1. doi: 10.1109/COMST.2021.3086475.
2. Casas E., Ramos L.T., Romero C., Rivas-Echeverria F. A Review of Computer Vision Applications for Asset Inspection in the Oil and Gas Industry, *Journal of Pipeline Science and Engineering*, 2025, Vol. 5(3), pp. 100246. doi: 10.1016/j.jpse.2024.100246.
3. Zhou L., Zhang L., Konz N. Computer Vision Techniques in Manufacturing, *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2022, Vol. 53 (1), pp. 105-117. doi: 10.1109/TSMC.2022.3166397.
4. Suguna A. Computer Vision for Autonomous Driving, *International Journal of Innovative Research in Information Security*, 2025, Vol. 11, pp. 846-850. doi: 10.26562/ijiris.2025.v11i12.03.
5. Dilek E., Dener M. Computer Vision Applications in Intelligent Transportation Systems: A Survey, *Sensors*, 2023, Vol. 23 (6), pp. 2938. doi: 10.3390/s23062938.
6. Kabir M.M., Rahman A., Hasan M.N., Mridha M.F. Computer Vision Algorithms in Healthcare: Recent Advancements and Future Challenges, *Computers in Biology and Medicine*, 2025, Vol. 185, pp. 109531. doi: 10.1016/j.combiomed.2024.109531.
7. Gao J., Yang Y., Lin P., Park D. Computer Vision in Healthcare Applications, *Journal of Healthcare Engineering*, 2018, pp. 1-4. doi: 10.1155/2018/5157020.
8. Ren S., He K., Girshick R., Sun J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks, *Advances in Neural Information Processing Systems (NeurIPS)*, 2015, pp. 91-99.
9. Cong X., Li S., Chen F., Liu C., Meng Y. A Review of YOLO Object Detection Algorithms Based on Deep Learning, *Frontiers in Computing and Intelligent Systems*, 2023, Vol. 4(2), pp. 17-20. doi: 10.54097/fcis.v4i2.9730.
10. Lei M., Li S., Wu Y., Hu H., Zhou Y., Zheng X., Ding G., Du S., Wu Z., Gao Y. YOLOv13: Real-Time Object Detection with Hypergraph-Enhanced Adaptive Visual Perception, *arXiv*, 2025. doi: 10.48550/arXiv.2506.17733.
11. Carion N., Massa F., Synnaeve G., Usunier N., Kirillov A., Zagoruyko S. End-to-End Object Detection with Transformers, *European Conference on Computer Vision (ECCV)*. Springer, 2020, pp. 213-229. doi: 10.1007/978-3-030-58452-8_13.
12. Shehzadi T., Hashmi K.A., Liwicki M., Stricker D., Afzal M.Z. Object Detection with Transformers: A Review, *Sensors*, 2025, Vol. 25, pp. 6025. doi: 10.3390/s25196025.
13. Lin T.Y., Maire M., Belongie S., Hays J., Perona P., Ramanan D., Dollár P., Zitnick C.L. Microsoft COCO: Common Objects in Context, *European Conference on Computer Vision (ECCV)*. Springer, 2014, pp. 740-755.
14. Everingham M., Van Gool L., Williams C.K.I., Winn J., Zisserman A. The Pascal Visual Object Classes (VOC) Challenge, *International Journal of Computer Vision*, 2010, Vol. 88 (2), pp. 303-338.
15. Kohler M., Eisenbach M., Gross H.M. Few-Shot Object Detection: A Comprehensive Survey, *IEEE Transactions on Neural Networks and Learning Systems*, 2023, Vol. PP. doi: 10.1109/TNNLS.2023.3265051.

16. Fan Q., Zhuo W., Tang C.K., Tai C.L. Few-Shot Object Detection with Attention-RPN and Multi-Relation Detector, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 4013–4022.
17. OpenMMLab. MMEFWSHOT: A Toolbox for Few Shot Learning Detection and Classification. Available at: <https://mefewshot.readthedocs.io/en/latest/> (accessed 14 January 2026).
18. Girshick R. Fast R-CNN, *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2015, pp. 1440–1448.
19. He K., Zhang X., Ren S., Sun J. Deep Residual Learning for Image Recognition, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778. doi: 10.1109/CVPR.2016.90.
20. Robbins H.E. A Stochastic Approximation Method, *Annals of Mathematical Statistics*, 1951, Vol. 22, pp. 400–407.

Бабин Степан Евгеньевич – АО «Научно-производственное предприятие «Авиационная и морская электроника»; e-mail: babinstepan12@yandex.ru; г. Санкт-Петербург, Россия; тел.: 88123274667; тестировщик программного обеспечения.

Гессен Павел Алексеевич – АО «Научно-производственное предприятие «Авиационная и морская электроника»; e-mail: gessen_pa@nppame.ru; г. Санкт-Петербург, Россия; тел.: 88123274667; зам. начальника центра средств интеллектуальной обработки изображений НПК РТС СН.

Павлова Валерия Анатольевна – АО «Научно-производственное предприятие «Авиационная и морская электроника»; e-mail: pavlova@nppame.ru; г. Санкт-Петербург, Россия; тел.: 88123274667; зам. директора НПК РТС СН по НИОКР.

Тупиков Владимир Алексеевич – АО «Научно-производственное предприятие «Авиационная и морская электроника»; e-mail: tupikov@nppame.ru; г. Санкт-Петербург, Россия; тел.: 88123274667; зам. ген. директора; директор НПК РТС СН.

Babin Stepan Evgenievich – Scientific and Production Enterprise Aviation and Marine Electronics, JSC; e-mail: babinstepan12@yandex.ru; St. Petersburg, Russia; phone: +78123274667; software tester.

Gessen Pavel Alekseevich – Scientific and Production Enterprise Aviation and Marine Electronics, JSC; e-mail: gessen_pa@nppame.ru; St. Petersburg, Russia; phone: +78123274667; deputy head of the Intelligent Image Processing Center of the Scientific and Production Corporation RTS SN.

Pavlova Valeria Anatolyevna – Scientific and Production Enterprise Aviation and Marine Electronics, JSC; e-mail: pavlova@nppame.ru; St. Petersburg, Russia; phone: +78123274667; deputy director for R & D, Scientific and Production Complex of Special-Purpose Robotic Systems.

Tupikov Vladimir Alekseevich – Scientific and Production Enterprise Aviation and Marine Electronics, JSC; e-mail: tupikov@nppame.ru; St. Petersburg, Russia; phone: +78123274667; deputy general director, director of the Scientific and Production Complex of Special-Purpose Robotic Systems.

УДК 004.428

DOI 10.18522/2311-3103-2026-2-258-271

Н.А. Бочаров, К.А. Суминов, М.А. Кирилюк, Н.Б. Парамонов

ИССЛЕДОВАНИЕ ПОДХОДОВ К УНИФИКАЦИИ СВЯЗУЮЩЕГО ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ РОБОТОТЕХНИЧЕСКИХ КОМПЛЕКСОВ ДЛЯ АППАРАТНО-ПРОГРАММНОЙ ПЛАТФОРМЫ «ЭЛЬБРУС»

Импортозамещение аппаратно-программных систем управления является одним из ключевых факторов развития современной отечественной робототехники. Ограниченность программной экосистемы российских вычислительных платформ затрудняет их применение в составе бортовых вычислительных систем (БВС) робототехнических комплексов (РТК), несмотря на появление современных отечественных многоядерных процессоров, например, семейства «Эльбрус», обеспечивающих конкурентоспособные характеристики по производительности и энергоэффективности. ROS (Robot Operating System) и его развитие — ROS 2 — широко используются для разработки программного обеспечения РТК, однако их официальная поддержка ограничена архитек-