

МРНТИ 28.23.15
УДК 004.93

<https://doi.org/10.55452/1998-6688-2025-22-1-74-83>

¹**Скаков М.К.,**

докторант, ORCID ID: 0009-0003-2074-3629,

e-mail: skakov90@gmail.com

^{2*}**Исламгожаев Т.У.,**

PhD, ассист. профессор, ORCID ID: 0000-0001-7891-242X

*e-mail: talgat.islamgozhayev@astanait.edu.kz

¹**Абдилдаева А.А.,**

PhD, ассоц. профессор, ORCID ID: 0000-0002-6381-9350

e-mail: asselabdildayeva5@gmail.com

¹Казахский национальный университет им. аль-Фараби, г. Алматы, Казахстан

²Astana IT University, г. Астана, Казахстан

НАБОР ДАННЫХ ИЗОБРАЖЕНИЙ ДЛЯ ОБНАРУЖЕНИЯ И КЛАССИФИКАЦИИ ВОЕННОЙ ТЕХНИКИ (НДОКВТ)

Аннотация

Технология интеллектуального обнаружения объектов военных транспортных средств уже стала основой для задач разведки и отслеживания оружия и оборудования, что необходимо для осведомленности о ситуации в современных интеллектуальных войнах. За последние два десятилетия было собрано множество наборов данных изображений военных транспортных средств, основное внимание в которых уделялось классификации различных категорий военных транспортных средств, и почти все они недоступны обществу, а доступные имеют проблемы с качеством аннотаций. В ответ на отсутствие набора данных и качество существующих общедоступных наборов данных мы предлагаем специализированный набор данных, основанный на собственной коллекции изображений, а также на общедоступных. Для разработки методов автоматического обнаружения военных транспортных средств различных категорий и отличия их от гражданских мы создали новый набор изображений военных транспортных средств (НДОКВТ). Он состоит из 5899 изображений военных транспортных средств, собранных с использованием трех различных методов: автоматизированного скрапинга, ручного подбора результатов поиска изображений и техник увеличения данных. Насколько нам известно, наш набор данных НДОКВТ является единственным общедоступным набором данных военных транспортных средств, который учитывает скрытые и частично видимые объекты. Этот набор данных будет способствовать созданию моделей компьютерного зрения для обнаружения военных транспортных средств.

Ключевые слова: объекты военных транспортных средств, обнаружение объектов, набор изображений военных транспортных средств, ограничивающая рамка, машинное обучение, компьютерное зрение, классификация, глубокое обучение.

Введение

В системах наблюдения и безопасности автоматическое обнаружение военных транспортных средств особенно важно на пограничных постах. Современные достижения в глубоком обучении улучшили технологии обнаружения объектов, но задачи, связанные с военными транспортными средствами, требуют специализированных наборов данных. Эта статья представляет набор изображений военных транспортных средств (НДОКВТ) [1], включающий 5899 изображений. Эти изображения классифицированы на военные (танки, грузовики, самолеты, вертолеты) и гражданские транспортные средства (автомобили, самолеты). Набор данных содержит 6528 объектов, разделенных для обучения, валидации и тестирования, что

делает его надежным ресурсом для улучшения алгоритмов обнаружения. НДОКВТ устраняет ограничения существующих наборов данных, такие как исключение изображений с открытыми транспортными средствами, и включает разнообразные условия изображения для моделирования реальных сред. Набор данных был собран с использованием общественных источников и скрапинга, прошел строгую предварительную обработку и увеличение данных для повышения эффективности обучения моделей. НДОКВТ предоставляет исследователям инструмент для оценки и улучшения алгоритмов обнаружения, адаптированных к задачам военных транспортных средств.

Обзор литературы

Автоматическое обнаружение военных транспортных средств с использованием глубокого обучения стало значительным направлением в области компьютерного зрения, особенно с учетом его последствий для усиления мер пограничной безопасности и наблюдения. Несколько заметных исследований и технологических достижений заложили основу, на которой строится набор изображений военных транспортных средств (НДОКВТ). В следующем обзоре освещены ключевые методологии и достижения в области обнаружения транспортных средств, особенно военных транспортных средств.

Общие методы обнаружения транспортных средств

YOLO (You Only Look Once) стал краеугольным камнем в реальном времени обнаружения объектов, ценится за способность достигать высоких кадровых частот и точности с помощью одной сверточной сети [2].

SSD (Single Shot MultiBox Detector) предлагает хорошо сбалансированный подход, предоставляя конкурентный компромисс между скоростью и точностью обнаружения [3].

Faster R-CNN был важным в продвижении границ точности в двухэтапных процессах обнаружения, хотя он работает медленнее, чем его одношаговые аналоги [4].

Специфические достижения в обнаружении военных транспортных средств

MCUnet выделяется как специализированная сеть для классификации военных транспортных средств, показывая высокую точность по различным наборам данных, специально разработанным для военных приложений [5].

Методы переноса обучения продемонстрировали значительные преимущества в этой области с моделями, такими как ResNet и InceptionV3, адаптированными для улучшения возможностей извлечения признаков для изображений военных транспортных средств [6].

Пользовательские модели: проводились целенаправленные исследования по разработке моделей, специализирующихся на обнаружении определенных типов военных транспортных средств, таких как танки и самолеты. Эти усилия часто включают пользовательские модификации стандартных архитектур для удовлетворения уникальных характеристик военных транспортных средств [7].

Существующие наборы данных и их ограничения

Предыдущие наборы данных, такие как у Гупта и др. [8, 17, 18, 19, 20, 21, 22], содержат 6772 изображения для идентификации военных и гражданских транспортных средств, включая грузовики, танки, самолеты, вертолеты, автомобили. Из них 6642 изображения положительные, содержащие 11 528 объектов, и 130 отрицательных для распознавания сцен без целевых объектов. Все изображения преобразованы в формат .jpg для облегчения обработки моделями обнаружения объектов. Набор данных подготовлен путем сбора [22, 23], предварительной обработки и форматирования файлов. Однако он имеет ограничения, такие как отсутствие аннотаций для скрытых изображений, что снижает надежность моделей в реальных условиях.

Описание набора данных

Набор изображений военных транспортных средств (НДОКВТ) [1] специально разработан для поддержки разработки и валидации моделей машинного обучения, направленных на обнаружение военных транспортных средств по изображениям или видео с камер. Набор данных включает 5899 изображений, разделенных на 6 классов в двух основных категориях: военные и гражданские транспортные средства. Военные транспортные средства включают танки, грузовики, самолеты и вертолеты, в то время как гражданские категории охватывают автомобили и самолеты. Это разнообразие гарантирует, что модели, обученные на этом наборе данных, могут хорошо обобщать данные по различным типам транспортных средств в различных сценариях.

Таблица 1 – Количество изображений валидационного набора для каждого класса/метки

| Метка класса | Количество |
|---------------------|------------|
| танк | 422 |
| самолет | 794 |
| грузовик | 269 |
| вертолет | 730 |
| гражданское авто | 147 |
| гражданский самолет | 149 |

Материалы и методы

Сбор данных

НДОКВТ был собран через систематический процесс сбора данных, используя общедоступные ресурсы и целевые методы скрапинга. Используя комбинацию библиотек BeautifulSoup и Requests, изображения были получены на основе конкретных ключевых слов, связанных с военными и гражданскими транспортными средствами, что обеспечивает разнообразный и всеобъемлющий набор данных.

Предварительная обработка данных

Предварительная обработка данных является критическим шагом в подготовке набора изображений военных транспортных средств (НДОКВТ) [1] для эффективного использования в моделях машинного обучения. Этот процесс обеспечивает, что данные, подаваемые в модели, являются однородными и способствуют изучению закономерностей, а не шуму. Все собранные изображения были стандартизированы для обеспечения однородности, что является важным для последующего обучения моделей машинного обучения.

Нормализация разрешения. Чтобы добиться однородности по всему набору данных, все изображения были изменены до согласованного разрешения 640x640 пикселей. Этот шаг важен для обеспечения того, чтобы все объекты на изображениях имели одинаковый масштаб, что помогает в обучении сверточных нейронных сетей (CNN), предоставляя им стандартизированные размеры входных данных.

Коррекция ориентации. Были применены автоматические корректировки ориентации, чтобы все изображения были выровнены в одной ориентации. Этот шаг нормализации помогает предотвратить обучение модели неправильным позиционным смещениям, связанным с ориентацией транспортных средств.

Стандартизация формата. Учитывая разнообразие источников для изображений, они первоначально были в различных форматах (JPEG, PNG, BMP и т.д.). Все изображения были преобразованы в формат JPEG, чтобы уменьшить сложность входного конвейера и оптимизировать хранение.

Увеличение данных

Техники увеличения данных были применены к изображениям в наборе данных для улучшения устойчивости и способности к обобщению обученных моделей. Эти техники имитируют различные реальные условия, с которыми может столкнуться система обнаружения военных транспортных средств. Стратегии увеличения включали:

Геометрические трансформации. Изображения случайным образом переворачивались горизонтально и поворачивались в пределах от -15 до $+15$ градусов. Эти трансформации помогают модели научиться распознавать транспортные средства с разных углов и ориентаций, что важно для реальных приложений, как показано на рисунке 3(с).

Фотометрические трансформации. Были применены корректировки, такие как изменение яркости, контрастности и насыщенности, а также случайное преобразование некоторых изображений в черно-белые. Эти изменения подготавливают модель к работе в различных условиях освещения и когда информация о цвете ненадежна, как показано на рисунке 3(б).

Искусственное добавление шума и размытие. Чтобы имитировать реальные несовершенства в получении изображений, были введены искусственные шумы и легкое размытие (до 2,5 пикселей). Этот шаг тренирует модель быть эффективной даже когда входные изображения невысокого качества из-за факторов окружающей среды или ограничений сенсоров, как показано на рисунке 3(а).

Аннотирование/Разметка данных

Аннотации играют жизненно важную роль в обучении с учителем, так как они предоставляют исходные данные, из которых модели учатся. Была настроена удаленная серверная Label Studio [12] для обеспечения одновременных параллельных возможностей аннотирования. Изображения были аннотированы ограничивающими рамками и метками командой из шести аннотаторов, включая военных экспертов. Этот тщательный процесс аннотирования обеспечивает высокую точность и надежность данных для обучения, что важно для разработки эффективных моделей обнаружения. Этот процесс аннотирования для НДОКВТ был тщательно разработан для обеспечения высококачественных меток:

Ограничивающие рамки: Каждое транспортное средство в наборе данных было аккуратно заключено в ограничивающую рамку обученными аннотаторами. Координаты каждой рамки были записаны в метаданные, связанные с изображениями, как показано на рисунке 1(а) и рисунке 1(б).

Точность меток. Команда из шести аннотаторов, включая двух с военным опытом, использовала Label Studio [12] для аннотирования изображений. Этот подход гарантировал, что транспортные средства были не только правильно идентифицированы, но и что ограничивающие рамки были точно размещены, отражая истинные размеры и положения транспортных средств.

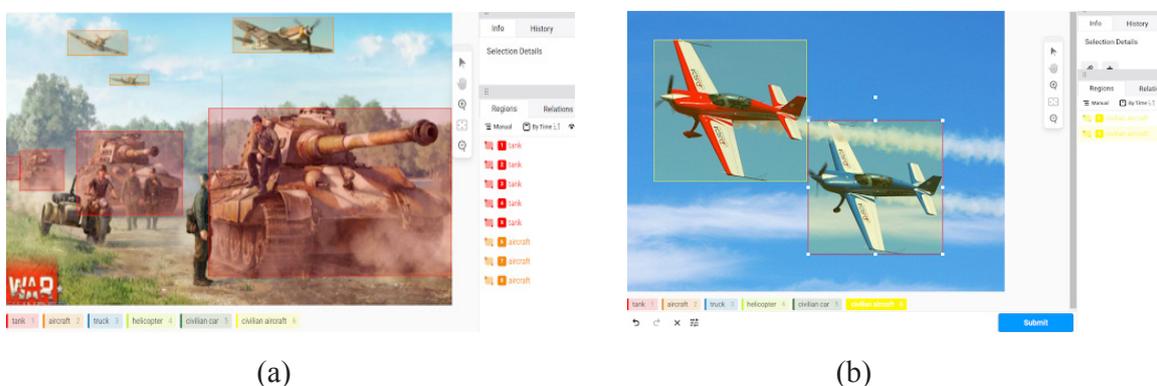


Рисунок 1 – Аннотирование/маркировка данных: (а) несколько различных классов на одном изображении; (б) гражданские самолеты

Механизм проверки. Для обеспечения согласованности и точности аннотаций была внедрена двухуровневая система проверки. После первоначальных аннотаций отдельная команда проверяла случайную выборку изображений на наличие ошибок или несоответствий, обеспечивая дополнительный уровень контроля качества.

Эти подробные процессы подготовки данных и аннотирования гарантируют, что НДОКВТ является надежным ресурсом, готовым поддерживать передовые исследования в области обнаружения военных транспортных средств с помощью машинного обучения. Этот комплексный подход увеличивает полезность набора данных для разработки систем обнаружения, эффективных в различных и сложных операционных условиях.

Обучение моделей

Выбор архитектуры

Для НДОКВТ мы использовали передовые архитектуры нейронных сетей, известные своей эффективностью в задачах обнаружения объектов. Основной архитектурой, используемой в этом проекте, была YOLOv8 [2], выбранная за баланс между скоростью и точностью, что делает ее подходящей для сценариев реального времени. Кроме того, мы экспериментировали с вариантами SSD и Faster R-CNN для сравнения производительности различных архитектурных дизайнов. Каждая из двух предобученных моделей (включая YOLOv8n и YOLOv8x) [2] была настроена для оптимизации точности обнаружения для классов как военных, так и гражданских транспортных средств, включая корректировки конфигураций слоев и функций активации для лучшей обработки специфических задач набора данных.

Оптимизация гиперпараметров

Тонкая настройка гиперпараметров критически важна для оптимизации процесса обучения и достижения наилучших результатов модели. Мы провели обширные эксперименты для нахождения оптимальных настроек:

Скорость обучения. Изначально была установлена скорость обучения 0.01 с корректировками на основе адаптивных техник скорости обучения для снижения риска превышения минимальных значений.

Размер пакета. Были протестированы различные размеры пакетов для баланса между вычислительной эффективностью и ограничениями памяти. Настройки были 32 для NVIDIA RTX 3090 и 16 для двух NVIDIA RTX 3060 GPUs [16] для двух предобученных моделей.

Оптимизатор. Stochastic Gradient Descent (SGD) использовался с импульсом 0.9 и распадом веса $5e-4$ для обеспечения плавной сходимости.

Процедура обучения

Обучение проводилось с использованием систематического подхода для обеспечения всестороннего обучения.

Разделение данных. Набор данных был разделен на 85% для обучения, 10% для валидации и 5% для тестирования. Это разделение было выбрано для предоставления значительного обучающего набора, при этом позволяя проводить тщательную валидацию и тестирование.

Кросс-валидация. Мы реализовали стратегию 5-кратной кросс-валидации для оценки производительности модели на различных поднаборах данных, что повышает надежность оценочных метрик.

Раннее прекращение. Чтобы предотвратить переобучение и обеспечить обобщение, было внедрено раннее прекращение на основе потерь при валидации. Обучение прекращалось, если не наблюдалось улучшения в течение 20 последовательных эпох.

Описания обученных моделей

Модель-1: YOLOv8n.

Детали архитектуры: модель-1, основанная на архитектуре YOLOv8n, была настроена для высокоскоростного обнаружения с уменьшенным числом параметров для повышения скорости вывода при сохранении разумной точности. Эта модель состоит из 168 слоев, оптимизированных для быстрого извлечения признаков и обработки в реальном времени.

Параметры. Она состоит примерно из 3,006,818 обучаемых параметров, что обеспечивает баланс между глубиной и вычислительной эффективностью, делая ее подходящей для развертывания в условиях реального времени.

Продолжительность обучения. Общая продолжительность обучения для этой модели составила примерно 38,1 минуты на GPU NVIDIA RTX 3090, что демонстрирует ее эффективность в быстром достижении решения благодаря оптимизированной архитектуре.

Модель-2: YOLOv8x.

Детали архитектуры. Модель-2, использующая архитектуру YOLOv8x, была разработана для более сложных задач обнаружения с акцентом на точность. Она включает 268 слоев, значительно больше, чем у модели-1, что позволяет ей захватывать более тонкие детали и нюансы различных типов транспортных средств.

Параметры. Эта модель значительно больше, содержит около 68,129,346 параметров, что позволяет ей достигать более высокой точности и лучше справляться с разнообразными и сложными аспектами обнаружения военных транспортных средств.

Продолжительность обучения. Обширная архитектура потребовала более длительного времени обучения, всего около 7,845 ч на той же установке NVIDIA. Это длительное время обучения оправдано значительными улучшениями в точности обнаружения и надежности модели.

Результаты

Оценка обученных моделей на наборе данных изображений военных транспортных средств (НДОКВТ) дала обнадеживающие результаты. Обе модели (YOLOv8n и YOLOv8x) были оценены на основе метрик потерь, точности, полноты и mAP в ходе этапов обучения и валидации, как показано в таблице 2 и на рисунках 2 и 3.

Таблица 2 – Результаты модели-1 (YOLOv8n) и модели-1 (YOLOv8x)

| Модель | Изображений | Экземпляров | Box (P) | R | mAP | mAP50-95 |
|---------|-------------|-------------|---------|-------|-------|----------|
| YOLOv8n | 501 | 1034 | 0.826 | 0.707 | 0.771 | 0.554 |
| YOLOv8x | 501 | 1034 | 0.87 | 0.765 | 0.816 | 0.625 |

5.1. Метрики производительности моделей

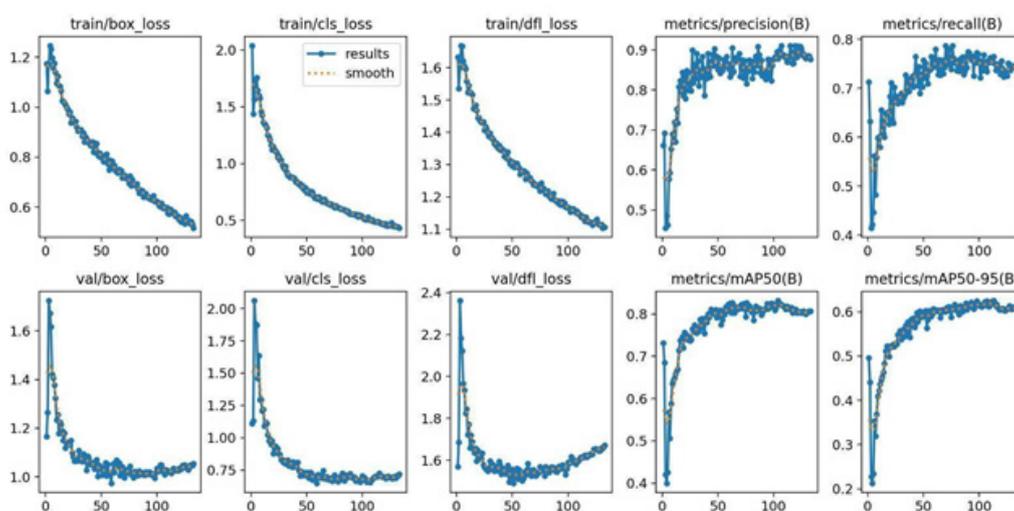


Рисунок 2 – Потери при обучении и валидации по эпохам из левых двух колонок и строк; точность и полнота по эпохам в верхних правых двух колонках в первой строке; классы ограничивающих рамок и распределение фокальных потерь и mAP по эпохам

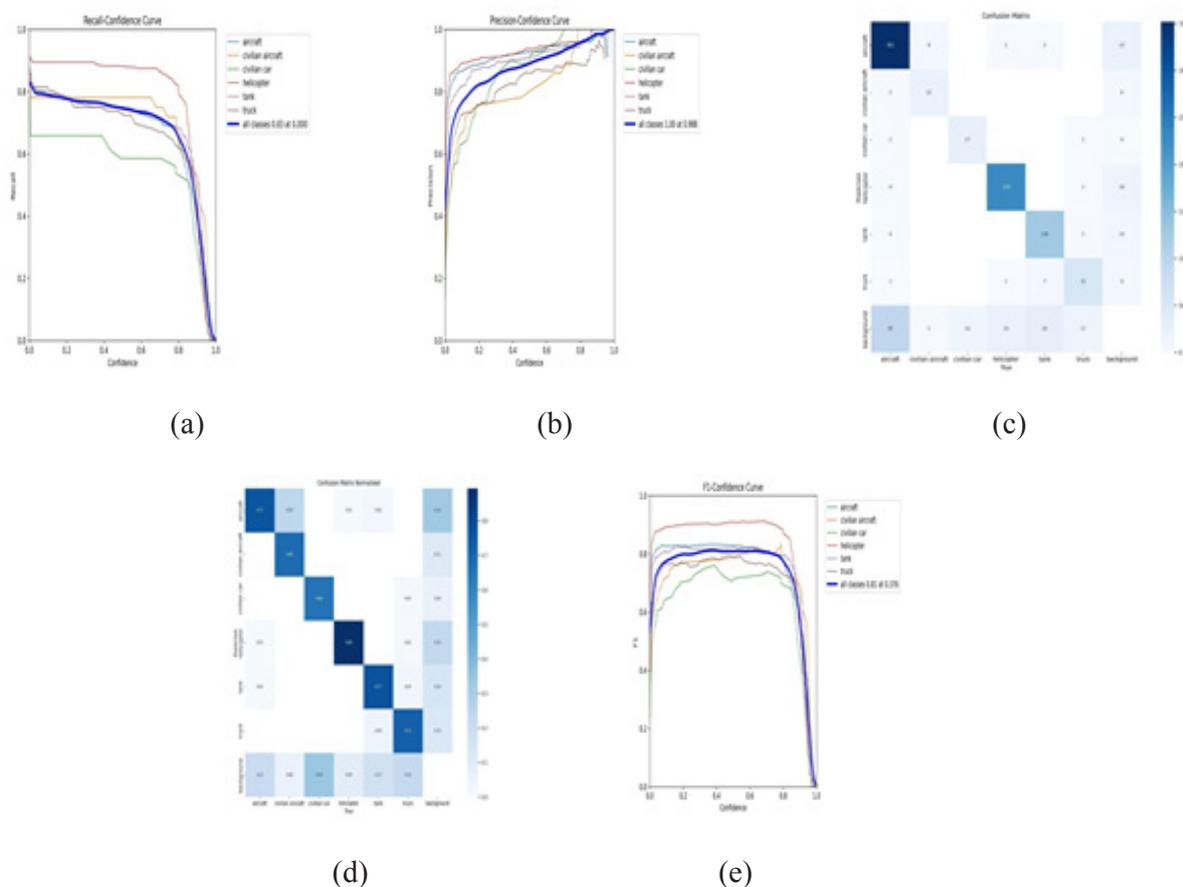


Рисунок 3 – Аннотирование данных/маркировка: (а) кривая полноты–уверенности; (б) кривая точности–уверенности; (с) матрица путаницы; (д) нормализованная матрица путаницы; (е) кривая F1–уверенности

Кривые полноты–уверенности и точности–уверенности демонстрируют, что модели достигли значительной полноты и точности на различных порогах уверенности (рисунки 3(а) и 3(б)). Кривая F1–уверенности указывает на баланс между точностью и полнотой, причем модель-2 превосходит модель-1, особенно в обнаружении мелких или частично скрытых транспортных средств (рисунок 3).

Матрица путаницы и нормализованная матрица путаницы предоставляют представление о производительности по классам, подчеркивая возможности и ограничения моделей в различении различных типов транспортных средств (рисунки 3(с) и 3(д)).

Обсуждение

Процесс обучения показал последовательные улучшения в снижении потерь и стабильности метрик точности и полноты. Модель-1 продемонстрировала способность к быстрому обучению, что идеально для приложений в реальном времени, где критична скорость. В то время как модель-2 показала превосходную производительность в точности, что полезно для приложений, требующих высокой надежности.

Критический анализ матриц путаницы показал, что, хотя модели могли эффективно различать военные и гражданские транспортные средства, наблюдалась некоторая путаница между похожими классами, такими как грузовики и танки. Эти выводы указывают на потенциальные области для дальнейшего совершенствования модели.

Заклучение

В этом исследовании представлены две модели нейронных сетей, обученные на новом наборе данных НДОКВТ. Модель-1 предлагает быстрое и эффективное решение для обнаружения военных транспортных средств в реальном времени, в то время как модель-2 обеспечивает более высокую точность, что делает ее подходящей для приложений, где точность имеет первостепенное значение.

Было показано, что модели способны хорошо обобщать данные по различным типам транспортных средств и условиям, делая значительные шаги вперед по сравнению с предыдущими моделями, обученными на менее разнообразных наборах данных. Будущие шаги включают дальнейшее улучшение возможностей моделей с помощью передовых методов обучения и расширение набора данных для включения более широкого разнообразия экземпляров транспортных средств.

Наши результаты подчеркивают важность комплексных наборов данных и специализированных моделей в области автоматического обнаружения военных транспортных средств, прокладывая путь к улучшенной пограничной безопасности и приложениям наблюдения.

REFERENCES

- 1 Dataset of images for detection and classification of military equipment (NDOMV), 2024. Available online: https://drive.google.com/file/d/1m8MIo06XmgtYsiks0LDrjKh_STD9L4Vb/view?usp=drive_link.
- 2 Redmon J., Divvala S., Girshick R., and Farhadi A. You Only Look Once: Unified, Real-Time Object Detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 779–788.
- 3 Liu W., Anguelov D., Erhan D., Szegedy C., Reed S., Fu C.Y., and Berg, A.C. (2016). SSD: Single Shot MultiBox Detector. Proceedings of the European Conference on Computer Vision (ECCV), pp. 21–37.
- 4 Ren S., He K., Girshick R., and Sun J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. Advances in Neural Information Processing Systems (NeurIPS), 2015, no. 28.
- 5 MCUnet Technology, 2020. Available online: <https://hanlab.mit.edu/projects/mcunet>.
- 6 He K., Zhang X., Ren, S., and Sun, J. Deep Residual Learning for Image Recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778.
- 7 Custom models designed for military vehicle detection.
- 8 Gupta P., Mandal R., and Sinha, D. Military and Civilian Vehicles Classification. Mendeley Data, 2021, vol.1.
- 9 Dosovitskiy A., Beyer L., and Kolesnikov A., et al. An image is worth 16x16 words: Transformers for image recognition at scale. arXiv preprint arXiv:2010.11929, 2019.
- 10 Tan M., and Le, Q. EfficientNet: Rethinking model scaling for convolutional neural networks. Proceedings of the International Conference on Machine Learning (ICML), 2019, pp. 6105–6114.
- 11 He K., Zhang X., Ren S., and Sun J. Deep Residual Learning for Image Recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778.
- 12 Label Studio Technology. Available online: <https://labelstud.io/>.
- 13 JPEG Image Format. Available online: <https://jpeg.org/>.
- 14 PNG Image Format. Available online: <https://en.wikipedia.org/wiki/PNG>.
- 15 BMP Image Format. Available online: https://en.wikipedia.org/wiki/BMP_file_format.
- 16 Nvidia Graphics Cards. Available online: <https://www.nvidia.com/ru-ru/geforce/graphics-cards/>.
- 17 CapstoneProject. Russian-Military-Vehicles Dataset. Available online: <https://universe.roboflow.com/capstoneproject/russian-military-vehicles> (accessed on 30 July 2024).
- 18 Bose A.R. Normal vs Military Vehicles Dataset. Available online: <https://www.kaggle.com/datasets/amanrajbose/normal-vs-military-vehicles> (accessed on 30 July 2024).
- 19 CapstoneProject. Russian-Military-Annotated Dataset. Available online: <https://universe.roboflow.com/capstoneproject/russian-military-annotated> (accessed on 30 July 2024).

20 Verone M. Temp Dataset. Available online: <https://universe.roboflow.com/mike-verone-wxrqi/temp-uanyu> (accessed on 30 July 2024).

21 MVDdetection. MV_Detection-FECL Dataset. Available online: https://universe.roboflow.com/mvddetection/mv_detection-fecl (accessed on 30 July 2024).

22 Fernandes A.A.A., Koehler, M., Konstantinou N., Pankin P., Paton N.W., and Sakellariou, R. Data Preparation: A Technological Perspective and Review. SN Comput. Sci., 2023, vol. 4, no. 425. <https://doi.org/10.1007/s42979-023-01828-8>.

23 Aguinis H., Hill N.S., and Bailey, J.R. Best Practices in Data Collection and Preparation: Recommendations for Reviewers, Editors, and Authors. Organ. Res. Methods, 2021, vol. 24, pp. 678–693. <https://doi.org/10.1177/1094428119836485>.

¹Скаков М.Қ.,

докторант, ORCID ID: 0009-0003-2074-3629,

e-mail: skakov90@gmail.com

^{2*}Исламғожаев Т.У.,

PhD, ассист. профессор, ORCID ID: 0000-0001-7891-242X

*e-mail: talgat.islamgozhayev@astanait.edu.kz

¹Абдилдаева А.А.,

PhD, қауымдастырылған профессор, ORCID ID: 0000-0002-6381-9350

e-mail: asselabdildayeva5@gmail.com

¹әл-Фараби атындағы Қазақ ұлттық университеті, Алматы қ., Қазақстан

²Astana IT University, Астана қ., Қазақстан

ӘСКЕРИ КӨЛІКТЕРДІ АНЫҚТАУ ЖӘНЕ ЖІКТЕЛУГЕ АРНАЛҒАН БЕЙНЕЛЕР ДЕРЕКТЕР ЖИНАҒЫ (ӘКАЖДЖ)

Аңдатпа

Әскери көлік құралдарын интеллектуалды анықтау технологиясы – қазіргі заманғы интеллектуалды соғыстарда қару-жарақ пен жабдықтарды барлау және бақылау үшін қажетті ахуалдық хабардарлықты қамтамасыз етудің негізі. Соңғы жиырма жыл ішінде әскери көлік құралдарының суреттерінен тұратын көптеген деректер жиынтығы жиналды, олардың басым бөлігі әскери көлік құралдарының әртүрлі санаттарын классификациялауға бағытталған. Алайда, бұл деректер жиынтығының көпшілігі көпшілікке қолжетімді емес, ал қолжетімді деректер жиынтығы аннотация сапасындағы мәселелерден зардап шегеді. Осы деректердің жетіспеушілігі мен қолжетімді жиынтықтардың сапасындағы ақауларға байланысты біз өзіміз жинаған суреттер жиынтығы мен қолжетімді деректер негізінде арнайы жиынтық ұсынамыз. Әскери көлік құралдарының әртүрлі санаттарын автоматты түрде анықтау және оларды азаматтық көліктерден ажырату әдістерін жасау мақсатында біз әскери көлік құралдарының жаңа суреттер жиынтығын (ӘКАЖДЖ) [1] әзірледік. Ол автоматтандырылған скрейпинг, сурет іздеу нәтижелерін қолмен таңдау және деректерді ұлғайту әдістерін қолдану арқылы жиналған 5899 әскери көлік құралының суреттерінен тұрады. Біздің білуімізше, ӘКАЖДЖ жиынтығы жасырын және ішінара көрінетін объектілерді ескеретін жалғыз ашық қолжетімді әскери көлік құралдарының деректер жиынтығы. Бұл деректер жиынтығы әскери көлік құралдарын анықтауға арналған компьютерлік көру модельдерін әзірлеуге ықпал етеді.

Тірек сөздер: әскери көлік құралдарының объектілері, объектілерді анықтау, әскери көлік құралдарының бейнелері, шектеу жолағы, машиналық оқыту, компьютерлік көру, классификация, терең оқыту

¹**Skakov M.,**

PhD student, ORCID ID: 0009-0003-2074-3629,
e-mail: skakov90@gmail.com

^{2*}**Islamgozhayev T.,**

PhD, assistant professor, ORCID ID: 0000-0001-7891-242X
*e-mail: talgat.islamgozhayev@astanait.edu.kz

¹**Abdildayeva A.A.,**

PhD, associate professor, ORCID ID: 0000-0002-6381-9350
e-mail: asselabdildayeva5@gmail.com

¹Al-Farabi Kazakh National University, 050040, Almaty, Kazakhstan

²Astana IT University, 010000, Astana, Kazakhstan

AN IMAGE DATASET FOR MILITARY VEHICLE DETECTION AND CLASSIFICATION (DMVDC)

Abstract

The technology of intelligent detection of military vehicle objects has already become the foundation for tasks related to intelligence and tracking of weapons and equipment, which is essential for situational awareness in modern intelligent warfare. Over the past two decades, numerous datasets of military vehicle images have been collected, primarily focusing on the classification of various categories of military vehicles. However, almost all these datasets are not publicly available, and the publicly available ones suffer from annotation quality issues. To address the lack of datasets and the quality of existing public datasets, we propose a specialized dataset based on our own collection of images, as well as publicly available ones. To develop methods for automatic detection of various categories of military vehicles and distinguishing them from civilian vehicles, we created a new dataset for military vehicle detection and classification (DMVDC). It consists of 5,899 images of military vehicles collected using three different methods: automated scraping, manual selection of image search results, and data augmentation techniques. To the best of our knowledge, our DMVDC [1] dataset is the only publicly available dataset of military vehicles that consider hidden and partially visible objects. This dataset will contribute to the development of computer vision models for detecting military vehicles.

Key words: military vehicles, object detection, military vehicle image set, bounding box, machine learning, computer vision, classification, deep learning.

Дата поступления статьи в редакцию: 23.12.2024