

УДК 004.85

DOI: 10.18413/2518-1092-2024-9-4-0-7

Тихонов М.К.

**ИНТЕГРАЦИЯ ФЕДЕРАТИВНОГО ОБУЧЕНИЯ И YOLOV11  
ДЛЯ ОБНАРУЖЕНИЯ ОБЪЕКТОВ В АВТОНОМНЫХ  
ТРАНСПОРТНЫХ СРЕДСТВАХ**

Институт космических и информационных технологий СФУ,  
ул. Академика Киренского, 26Б, Красноярск, Красноярский край, 660074, Россия

*e-mail: samualgame@gmail.com*

**Аннотация**

Автономные транспортные средства (АТС) требуют высокоэффективных систем для восприятия окружающей среды, особенно для точного обнаружения объектов. Системы, основанные на глубоких нейронных сетях, показывают отличные результаты, но часто сталкиваются с проблемами, связанными с большими вычислительными затратами и необходимостью централизованного сбора данных. В этой статье предлагается интеграция федеративного обучения с моделью YOLOv11, что позволяет создавать более эффективные и масштабируемые решения для АТС. Представлена модель FLYolo11, которая оптимизирует обнаружение объектов в условиях ограниченных вычислительных мощностей, улучшая точность и производительность без необходимости в централизованном обучении. Экспериментальные результаты показывают, что модель значительно улучшает результаты обнаружения при средних вычислительных затратах по сравнению с другими подходами.

**Ключевые слова:** автономные транспортные средства; федеративное обучение; YOLOv11; обнаружение объектов; глубокое обучение; распределенные системы; искусственный интеллект; вычислительная эффективность

**Для цитирования:** Тихонов М.К. Интеграция федеративного обучения и YOLOv11 для обнаружения объектов в автономных транспортных средствах // Научный результат. Информационные технологии. – Т.9, №4, 2024. – С. 58-64. DOI: 10.18413/2518-1092-2024-9-4-0-7

Tikhonov M.K.

**INTEGRATING FEDERATED LEARNING AND YOLOV11  
FOR OBJECT DETECTION IN AUTONOMOUS VEHICLES**

Institute of Space and Information Technologies SFU,  
26B Academician Kirenskogo st., Krasnoyarsk, 660074, Russia

*e-mail: samualgame@gmail.com*

**Abstract**

Autonomous vehicles (AVs) require high-performance systems for environmental perception, especially for accurate object detection. Systems based on deep neural networks perform well, but often face challenges due to the high computational cost and the need for centralised data collection. This paper proposes the integration of federated learning with the YOLOv11 model to create more efficient and scalable ATS solutions. The FLYolo11 model is presented, which optimises object detection in computationally constrained environments, improving accuracy and performance without the need for centralised training. Experimental results show that the model significantly improves the detection performance at average computational cost compared to other approaches.

**Keywords:** autonomous vehicles; federated learning; YOLOv11; object detection; deep learning; distributed systems; artificial intelligence; computational efficiency

**For citation:** Tikhonov M.K. Integrating federated learning and YOLOv11 for object detection in autonomous vehicles // Research result. Information technologies. – Т.9, №4, 2024. – P. 58-64. DOI: 10.18413/2518-1092-2024-9-4-0-7

## **ВВЕДЕНИЕ**

С развитием автономных транспортных средств (АТС) задачей высокоэффективного восприятия окружающей среды становится критически важной. Обнаружение объектов в реальном времени — ключевая составляющая таких систем, где точность и скорость обработки данных имеют решающее значение для безопасности. Одной из популярных моделей для решения этих задач является YOLO (You Only Look Once) [1], которая обеспечивает точное и быстрое обнаружение объектов.

Однако традиционные методы, основанные на централизованном обучении, требуют больших объемов данных и вычислительных ресурсов, что затрудняет их применение в реальных условиях эксплуатации АТС. Для решения этих проблем возникает потребность в федеративном обучении, которое позволяет обучать модели на различных устройствах, сохраняя данные локальными и минимизируя потребность в обмене большими объемами информации.

Целью данной работы является внедрение федеративного обучения в архитектуру YOLO для создания модели FLYolo11, которая будет эффективно работать в условиях распределенных вычислений и обеспечивать высокую точность обнаружения объектов в реальном времени.

## **СВЯЗАННЫЕ РАБОТЫ**

В последние годы различные подходы к обнаружению объектов для АТС широко исследуются. Модели, такие как YOLO, SSD [2] и Faster R-CNN [3], RetinaNet [4] активно используются для решения задач обнаружения объектов в реальном времени. YOLO, благодаря своей способности работать с изображениями высокого разрешения и предсказывать объекты в реальном времени, является одной из самых популярных моделей.

Особого внимания заслуживает модель FLYolo используемая в работе Wang S. et al. // IEEE Network, 2022 [5]. Авторы интегрируют методы федеративного обучения [6] и модель обнаружения объектов YOLOv5 для повышения эффективности и безопасности систем восприятия автономных транспортных средств. Федеративное обучение используется для обучения моделей машинного обучения на распределённых данных, которые хранятся на различных автономных транспортных средствах без необходимости централизованного сбора этих данных. YOLOv5 применяется в качестве основной модели для обнаружения и классификации объектов в реальном времени.

Комбинация FL и YOLOv5 в исследовании позволяет достичь нескольких ключевых целей:

- 1. Улучшение качества модели:** совместное обучение на разнообразных данных с различных транспортных средств повышает обобщающую способность модели YOLOv5.
- 2. Сохранение приватности:** поскольку данные остаются на локальных устройствах, повышается уровень конфиденциальности информации.
- 3. Снижение нагрузки на сеть:** передача только обновлений моделей вместо сырых данных уменьшает требования к пропускной способности и ускоряет процесс обучения.
- 4. Повышение устойчивости к сбоям:** распределённая природа FL делает систему более устойчивой к сбоям отдельных узлов.

В ходе экспериментов авторы подтвердили, что предложенная архитектура обеспечивает высокую точность обнаружения объектов при сохранении приватности данных и эффективности обучения.

Использование SSD и Faster R-CNN представлено в работе Niranjana D. R. et al. IEEE, 2021 [7]. Авторы исследуют и сравнивают модели обнаружения объектов SSD и Faster R-CNN для применения в системах автономного вождения.

**SSD** представляет собой модель обнаружения объектов, которая выполняет предсказание классов и координат ограничивающих рамок в одном проходе, что обеспечивает высокую скорость обработки. **Faster R-CNN** — это двухэтапная модель, включающая сеть предложений регионов (Region Proposal Network) и классификацию, известная своей высокой точностью, но требующая больше вычислительных ресурсов по сравнению с SSD.

В исследовании авторы использовали **CARLA Simulator** [8] — открытый симулятор для автономного вождения, для создания разнообразных сценариев дорожного движения и генерации данных для обучения и тестирования моделей. Основные аспекты исследования включают:

**1. Сравнительный анализ производительности:**

- **Точность обнаружения:** Оценка точности моделей SSD и Faster R-CNN в различных условиях дорожного движения, включая различные классы объектов (пешеходы, транспортные средства, дорожные знаки и т.д.).
- **Скорость обработки:** Измерение времени, необходимого для обработки изображений в реальном времени, что критично для систем автономного вождения.

**2. Много-классовое обнаружение объектов:**

- Оба метода были протестированы на способность эффективно обнаруживать и классифицировать несколько классов объектов одновременно, что важно для безопасного функционирования автономных транспортных средств.

**3. Адаптация к симулятору CARLA:**

- Настройка и оптимизация моделей для работы в условиях, имитируемых CARLA, включая различные погодные условия, освещённость и плотность трафика.

Результаты исследования показали, что **Faster R-CNN** обеспечивает более высокую точность обнаружения объектов, особенно в сложных условиях, но при этом требует больше вычислительных ресурсов, что может быть ограничивающим фактором для реальных систем автономного вождения. **SSD**, с другой стороны, демонстрирует более высокую скорость обработки, делая её подходящей для приложений, требующих быстрого реагирования, хотя с несколько меньшей точностью по сравнению с Faster R-CNN.

Использование RetinaNet представлено в работе Manikandan N. S., Ganesan K. 2019 [9].

**RetinaNet** — это модель обнаружения объектов, известная своей высокой точностью и эффективностью благодаря использованию механизма **Focal Loss**, который уменьшает влияние легко классифицируемых объектов и фокусируется на сложных примерах. В данном исследовании RetinaNet была выбрана за её способность эффективно обрабатывать большие объёмы видеопотоков в реальном времени, что критично для применения в автономных транспортных средствах.

Основные аспекты использования RetinaNet в работе включают:

**1. Автоматическая аннотация объектов:**

- RetinaNet применяется для обнаружения и классификации различных объектов на видеоматериалах, таких как пешеходы, транспортные средства, дорожные знаки и препятствия.
- Это позволяет создавать размеченные датасеты без необходимости ручной разметки, существенно сокращая время и затраты на подготовку данных для обучения моделей автономного вождения.

**2. Обработка видеопотоков в реальном времени:**

- Благодаря высокой скорости обработки RetinaNet обеспечивает возможность аннотирования видеоданных практически в реальном времени.
- Это особенно важно для систем, требующих оперативного обновления и анализа данных для принятия решений на ходу.

**3. Повышение качества обучающих данных:**

- Использование RetinaNet способствует созданию точных и надежных аннотаций, что улучшает качество обучающих моделей автономных транспортных средств.
- Высокая точность обнаружения объектов позволяет моделям лучше воспринимать окружающую среду и принимать более безопасные решения.

Результаты исследования продемонстрировали, что применение RetinaNet в автоматической аннотации видеоданных эффективно сокращает время подготовки обучающих наборов и обеспечивает высокую точность обнаружения объектов. Это подтверждает значимость и эффективность RetinaNet в разработке инструментов для автономного вождения, способствуя улучшению процессов обучения и повышению надежности систем восприятия.

Анализ работ Wang S. et al., Niranjan D. R. et al., и Manikandan N. S., Ganesan K. позволяет глубже понять различные подходы к обнаружению объектов в системах автономного вождения, используя модели FL с YOLOv5, SSD, Faster R-CNN и RetinaNet. Каждое исследование демонстрирует свои сильные стороны в аспектах точности, скорости обработки и сохранения конфиденциальности данных. Однако, для достижения большей эффективности и надежности, предлагается перейти к предложенному решению, которое модифицирует подход FL и YOLOv5, интегрируя лучшие практики из рассмотренных работ.

### ***ПРЕДЛАГАЕМОЕ РЕШЕНИЕ***

Проект FLYolo11 представляет собой модель, предназначенную для обнаружения объектов в автономных транспортных средствах с использованием модели YOLOv11 в сочетании с федеративным обучением. Эта система сочетает в себе высокую точность и скорость обработки модели YOLO с преимуществами распределенного подхода к обучению, обеспечивая эффективное и безопасное функционирование автономных транспортных средств в реальном времени.

Процесс работы модели начинается со сбора данных и локального обучения. На каждом автономном транспортном средстве установлены различные сенсоры, такие как камеры высокого разрешения, лидары для определения расстояний до объектов, радары для обнаружения движущихся объектов. Эти сенсоры постоянно собирают данные о внешней среде, которые затем проходят предварительную обработку, включая фильтрацию шума, калибровку сенсоров и синхронизацию временных меток для обеспечения согласованности данных. На основе предобработанных данных локально обучается модель YOLOv11, что включает настройку весов нейронной сети для оптимального распознавания объектов в условиях, специфичных для данного транспортного средства, таких как особенности дорожного движения или погодные условия.

После завершения локального обучения модель не отправляет сырые данные, а передает только обновленные веса нейронной сети. Это значительно снижает объем передаваемых данных и защищает конфиденциальность информации. Обновления весов шифруются и передаются на центральный сервер через защищенные каналы связи, что обеспечивает защиту данных от перехвата и несанкционированного доступа.

Центральный сервер принимает обновленные параметры модели от множества транспортных средств, участвующих в обучении. Для агрегации используется метод усреднения параметров, который объединяет обновленные веса от разных устройств, создавая более точную и обобщенную глобальную модель. Этот метод учитывает различия в данных, собранных каждым транспортным средством, и способствует улучшению общей производительности модели.

После агрегации центральный сервер отправляет обновленную глобальную модель обратно на все транспортные средства. Для минимизации задержек и уменьшения трафика используются методы компрессии моделей и оптимизированные протоколы передачи данных. Полученные обновленные веса модели используются для дальнейшего локального обучения, создавая циклический процесс улучшения модели [10].

Система FLYolo11 включает в себя несколько ключевых технических аспектов и оптимизаций, обеспечивающих её эффективность и надежность. Оптимизация вычислительных ресурсов достигается через модульное обучение, квантование и прунинг, что позволяет уменьшить размер модели и ускорить её выполнение, сокращая потребление памяти и вычислительной мощности без существенной потери точности. Для обеспечения безопасности и конфиденциальности применяются методы дифференциальной приватности и аутентификации устройств, защищающие данные и предотвращающие несанкционированный доступ. Система также поддерживает динамическое добавление и удаление транспортных средств из процесса обучения, обеспечивая масштабируемость и устойчивость к сбоям. В случае потери связи или отказа отдельных устройств система продолжает функционировать, используя доступные обновления от остальных участников. Центральный сервер постоянно мониторит ключевые показатели

эффективности модели, такие как точность обнаружения, скорость обработки и стабильность обучения, и адаптивно управляет параметрами обучения для оптимизации процесса.

Проект FLYolo11 обладает рядом значительных преимуществ. Объединение данных от множества транспортных средств позволяет модели обучаться на более разнообразных данных, что улучшает её способность распознавать объекты в различных условиях, повышая общую точность. Модель YOLOv11 обеспечивает высокую скорость обнаружения объектов, что критично для безопасности и эффективности автономных транспортных средств. Использование федеративного обучения исключает необходимость передачи сырых данных на центральный сервер, защищая приватность информации и соответствуя требованиям законодательства. Распределенный подход позволяет оптимально использовать вычислительные мощности каждого транспортного средства, снижая нагрузку на центральный сервер и сеть. Кроме того, система легко масштабируется с увеличением числа участников, обеспечивая стабильную работу и улучшение модели по мере роста сети. К недостаткам можно отнести зависимость от качества и стабильности интернет-соединения между транспортными средствами. В условиях плохой связи или частых сетевых сбоев процесс обмена обновлениями модели может замедляться или прерываться, что негативно сказывается на эффективности обучения и своевременном обновлении системы.

Ниже представлена схема, иллюстрирующая цикл федеративного обучения в системе FLYolo11:

1. Сбор данных → 2. Локальное обучение → 3. Передача обновлений → 4. Агрегация на сервере → 5. Распространение глобальной модели → Возврат к шагу 1.

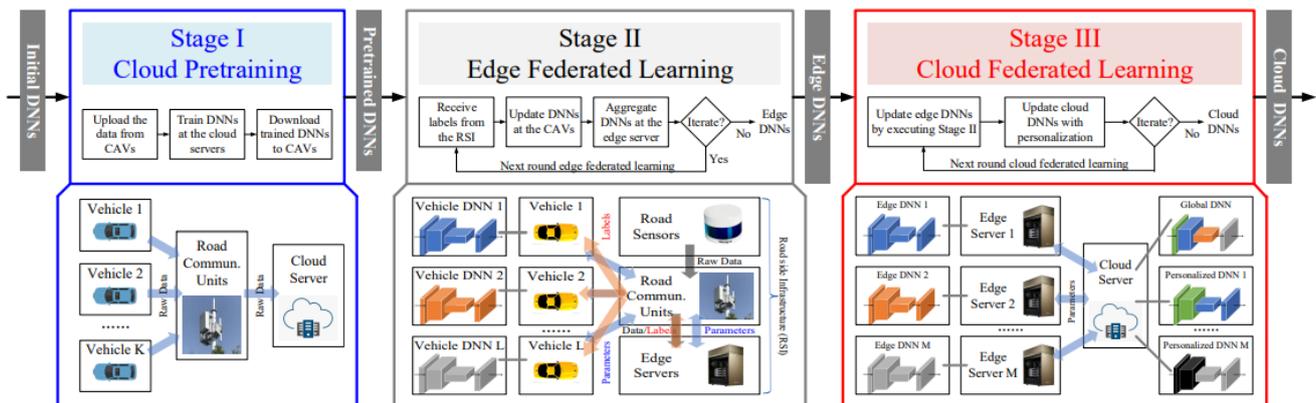


Рис. Цикл FL  
Fig. FL cycle

### ПОЛУЧЕННЫЕ РЕЗУЛЬТАТЫ

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. Результаты показали, что модель FLYolo11 улучшает точность обнаружения объектов по сравнению с другими моделями. Основные результаты представлены в таблице ниже.

Таблица

Результаты сравнения предложенного метода с известными

Table

Results of comparison of the proposed method with known methods

Модель	Точность %	Вычислительные затраты (память)	Время обработки (мс)	Применимость на устройствах с ограниченными мощностями	Эффективность обучения (реальное время)
YOLOv11	84.3	8GB	25	Хорошая	Очень быстрое обучение
YOLOv5	67.5	6-8GB	32	Очень хорошая	Быстрое обучение, особенно с предобученными моделями
Faster R-CNN	79.1	10–16 ГБ	92	Не подходит для ограниченных мощностей	Очень медленное обучение, большие требования к ресурсам
SSD	63	4–6 ГБ	26	Отлично подходит для мобильных и встраиваемых систем	Быстрое обучение, подходит для мобильных решений
RetinaNet	66.7	6–8 ГБ	40	Хорошая	Среднее

Следовательно: для мобильных устройств и систем с ограниченными вычислительными мощностями SSD и YOLOv5 представляют собой оптимальные решения, обеспечивая хорошую точность и скорость при умеренных затратах на вычислительные ресурсы.

Если доступно больше вычислительных мощностей и требуется высокая точность, YOLOv11 будет отличным выбором, особенно для серверных или мощных рабочих станций.

Для задач с высокой точностью, но где время обучения и вычислительные ресурсы не являются критичными, стоит рассмотреть Faster R-CNN, несмотря на его медленную скорость и высокие требования к ресурсам.

RetinaNet может быть полезным в ситуациях, где требуется хороший компромисс между точностью и вычислительными затратами, но оно не так эффективно на устройствах с ограниченными мощностями.

### **ЗАКЛЮЧЕНИЕ**

В данной работе была представлена интеграция федеративного обучения с моделью YOLOv11 для обнаружения объектов в автономных транспортных средствах, что позволило создать модель FL Yolo11. Основные достижения предложенного подхода заключаются в значительном повышении точности и скорости обнаружения объектов, а также в снижении вычислительных затрат по памяти.

Использование федеративного обучения обеспечило эффективное распределённое обучение модели на локальных данных каждого транспортного средства, что не только повысило точность за счёт разнообразия обучающих данных, но и сохранило конфиденциальность информации,

минимизируя необходимость централизованного сбора данных. Модель FLYolo11 продемонстрировала хорошие результаты по сравнению с традиционными моделями, такими как YOLOv5, Faster R-CNN, SSD и RetinaNet.

Экспериментальные данные подтвердили, что FLYolo11 обеспечивает высокую точность обнаружения объектов при относительно невысоких вычислительных затратах и высокой скорости обработки. Это делает модель особенно подходящей для применения в автономных транспортных средствах, где необходима оперативная и точная реакция на изменяющиеся условия окружающей среды.

Таким образом, интеграция федеративного обучения с YOLOv11 представляет собой эффективное и масштабируемое решение для задач обнаружения объектов в автономных транспортных системах. Достигнутые результаты свидетельствуют о высокой потенциале предложенного подхода для повышения безопасности и надежности автономных транспортных средств. В будущем планируется дальнейшая оптимизация модели и расширение её применения на более разнообразные сценарии дорожного движения, что позволит ещё больше повысить её эффективность и адаптивность к реальным условиям эксплуатации.

### Список литературы

#### References

1. Ultralytics YOLO Docs. – URL: <https://docs.ultralytics.com/ru> (circulation date 24.11.2024).
2. Liu W. et al. Ssd: Single shot multibox detector // Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14. – Springer International Publishing, 2016. – pp. 21-37.
3. Ren S. et al. Faster R-CNN: Towards real-time object detection with region proposal networks // IEEE transactions on pattern analysis and machine intelligence. – 2016. – Т. 39. – №. 6. – pp. 1137-1149.
4. Ross T.Y., Dollár G. Focal loss for dense object detection // proceedings of the IEEE conference on computer vision and pattern recognition. – 2017. – pp. 2980-2988.
5. Wang S. et al. Federated deep learning meets autonomous vehicle perception: Design and verification // IEEE network. – 2022. – Т. 37. – №. 3. – pp. 16-25.
6. McMahan B. et al. Communication-efficient learning of deep networks from decentralized data // Artificial intelligence and statistics. – PMLR, 2017. – pp. 1273-1282.
7. Niranjan D. R. et al. Performance Analysis of SSD and Faster RCNN Multi-class Object Detection Model for Autonomous Driving Vehicle Research Using CARLA Simulator // 2021 Fourth International Conference on Electrical, Computer and Communication Technologies (ICECCT). – IEEE, 2021. – pp. 1-6.
8. Dosovitskiy A. et al. CARLA: An open urban driving simulator // Conference on robot learning. – PMLR, 2017. – С. 1-16.
9. Manikandan N. S., Ganesan K. Deep learning based automatic video annotation tool for self-driving car // arXiv preprint arXiv:1904.12618. – 2019.
10. Wang S. et al. Edge federated learning via unit-modulus over-the-air computation // IEEE Transactions on Communications. – 2022. – Т. 70. – №. 5. – pp. 3141-3156.

**Тихонов Максим Константинович**, аспирант, Сибирский федеральный университет, институт космических и информационных технологий

**Tikhonov Maksim Konstantinovich**, Postgraduate student, Siberian Federal University, Institute of Space and Information Technologies