

Проектирование компонента классификации объектов и интерпретации их действий с использованием методов компьютерного зрения и машинного обучения

Д.П. Зубова, С.А. Семенист, С.Н. Широбокова

*Южно-Российский государственный политехнический университет (НПИ)
имени М.И. Платова, Новочеркасск*

Аннотация: В статье приведены аспекты проектирования модуля работы с искусственным интеллектом для анализа видеопотоков с камер наблюдения с целью классификации объектов и интерпретации их действий в рамках задачи сбора статистической информации и фиксации сведений об аномальной активности объектов наблюдения. Представлены диаграмма последовательности процесса работы пользователя с активными мониторингами с помощью Telegram-бота и концептуальная схема взаимодействия информационно-аналитической системы питомника породистых собак на платформе «1С:Предприятие» с внешними сервисами.

Ключевые слова: компьютерное зрение, машинное обучение, нейронные сети, искусственный интеллект, распознавание действий, классификация объектов, YOLO, поведенческие паттерны, поиск ключевых точек, 1С: Предприятие, Telegram-бот.

Устойчивое и активное развитие технологий искусственного интеллекта все больше расширяет спектр задач, выполнение которых может быть делегировано технологиям, позволяя эффективнее распределять людские ресурсы компании. Одним из актуальных и перспективных направлений является классификация объектов наблюдения и распознавание их действий [1] с применением методов компьютерного зрения и машинного обучения. На примере питомника породистых собак данную задачу можно свести к необходимости классификации человека в кадре с целью фиксации факта посещения животного персоналом питомника, а также к распознаванию действий самого животного с помощью метода поиска ключевых точек [2].

В информационно-аналитической системе питомника породистых собак на платформе «1С:Предприятие» модуль работы с искусственным интеллектом позволяет осуществлять анализ видеопотоков с камер наблюдения, размещенных в вольерах животных. Целью является обеспечение автоматизированного контроля за поведением животного, отслеживание подозрительных активностей [3], фиксирование времени и частоты посещения

питомца персоналом организации за указанный временной интервал. Концептуальная схема взаимодействия информационно-аналитической системы питомника породистых собак на платформе «1С: Предприятие» с внешними сервисами представлена на рис. 1.

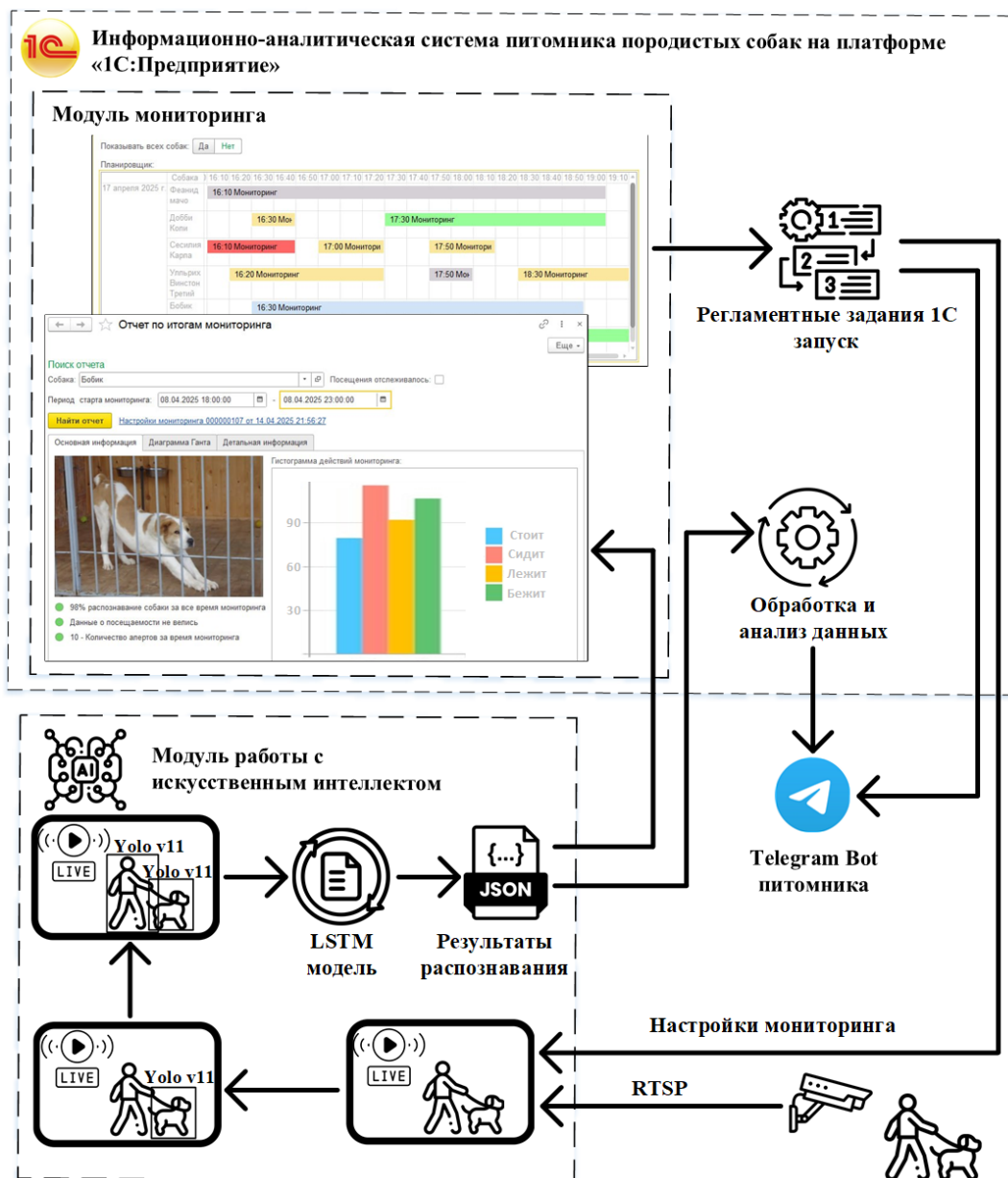


Рис. 1. – Концептуальная схема взаимодействия информационно-аналитической системы с внешними сервисами

В системе предусмотрен механизм логирования результатов мониторинга, анализ полученных данных с рассылкой уведомлений о начале и завершении запланированного мониторинга, отправкой отчетов и

мгновенных сообщений, предупреждающих об отклонении состояний объектов наблюдения от ожидаемых результатов. Рассылка уведомлений предусмотрена с использованием чат-бота [4, 5], который способен эмулировать живое человеческое общение с иллюзией диалога, следуя по структурированному алгоритму. Диаграмма последовательности процесса работы пользователя с активными мониторингами с помощью *Telegram*-бота представлена на рис. 2.

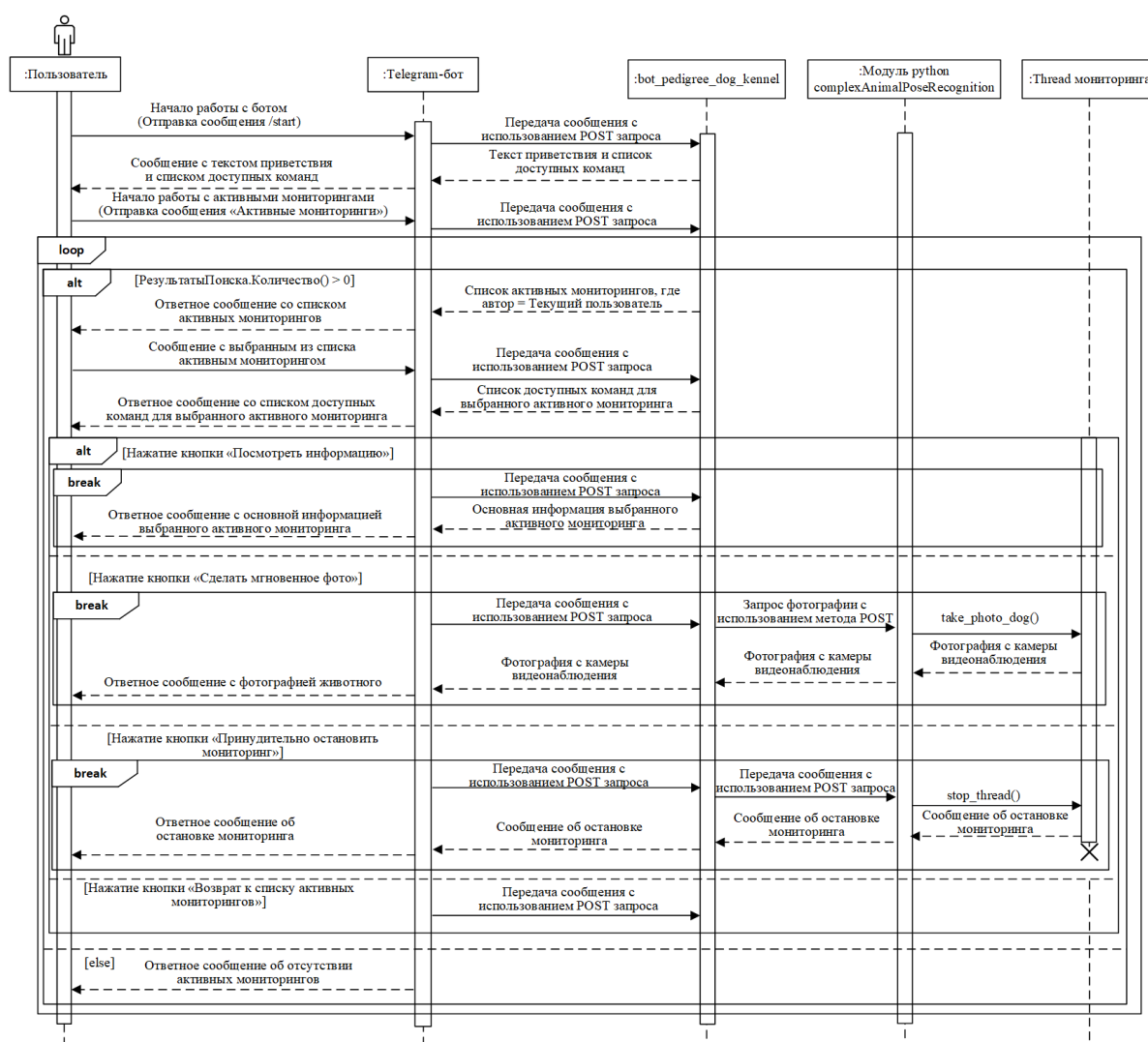


Рис. 2. – Диаграмма последовательности процесса работы пользователя с активными мониторингами с помощью *Telegram*-бота

Обнаружение животного в кадре для анализа действий объекта наблюдения осуществляется с помощью нейронной сети *YOLO* версии 11, ядром архитектуры которой является единая свёрточная нейронная сеть

(*Convolutional Neural Networks, CNN*) [6]. На сегодняшний день *CNN* это один из самых популярных алгоритмов глубокого обучения, позволяющий классифицировать и отличать изображения, при этом сам *YOLO* оперирует с полным кадром как в процессе обучения, так и в процессе тестирования [7]. Набор данных «*dog-pose.yaml*», который используется *YOLOv11 pose* для распознавания собак включает в себя 24 ключевые точки, 21 из них относится к анатомическим ориентирам животного, а остальные являются техническими и используются для нормализации позы и стабилизации обучения. Анатомические точки распределены следующим образом (см. табл. 1).

Таблица № 1

Распределение ключевых анатомических точек собаки

Количество ключевых точек	по 1 ключевой точке	по 2 ключевые точки	по 3 ключевые точки
Части тела собаки	Нос	Левое ухо	Левая передняя лапа
	Подбородок	Правое ухо	Правая передняя лапа
	Шея	Хвост	Левая задняя лапа
	-	-	Правая задняя лапа
Итого	3	6	12

Каждая из ключевых точек характеризуется 3 измерениями – x , y и видимость. Данный подход обеспечивает устойчивую к частичным окклюзиям [8] высоко детализированную позу, которая позволяет получить точные аннотации для обучения моделей.

Поскольку для корректного и эффективного обучения модели компьютерного зрения качество подготовки тренировочного и тестового датасетов имеет колоссальное значение, необходимо обеспечить разнообразие, четкость данных, нормализацию координат и предусмотреть балансировку классов во избежание искажения обучающей функции и игнорирования предсказаний редких классов. Для корректной оценки распознавания поз животных большую часть материала, который использовался для генерации данных обучения и тренировочных данных, составляли те видео, на которых животное было представлено одной особью.

Массив точек, полученных с каждого кадра видео обучающей выборки с помощью *YOLOv11 pose*, служит входными данными для рекуррентной нейронной сети *Long Short-Term Memory (LSTM)* [9]. Данная модель способна анализировать данные и изучать долгосрочные зависимости с целью выявления закономерности между движениями и их предсказания.

В рассматриваемой архитектуре предусмотрено 5 классов поведения животного, такие как «Бежит», «Лежит», «Сидит», «Стоит» и «Отсутствует в кадре». При планировании мониторинга система позволяет пользователю выбрать несколько настроек, отслеживание показателей которых будет осуществляться весь период логирования, обеспечивая тем самым гибкость и вариативность наблюдения за объектами в кадре. По умолчанию модуль «*complexAnimalPoseRecognition*», написанный на языке *Python*, подгружает только модель *YOLOv11 pose* для обнаружения ключевых точек собаки, но, при включенной настройке необходимости отслеживания количества посещений животного персоналом питомника, подгружается и модель *YOLOv11 Object Detection* для поиска человека в кадре. Принцип работы модели *Object Detection* сводится к выделению прямоугольных областей, очерчивающих искомые объекты на цифровом изображении и их классификации [10]. В случае успешного обнаружения человека в кадре, система фиксирует дату и время последнего навешания животного для корректной оценки общего числа посещений и возможности получения детализированной информации о каждом событии. Экранные формы процесса распознавания и фиксации ключевых сведений о событиях в кадре представлены на рис. 3.

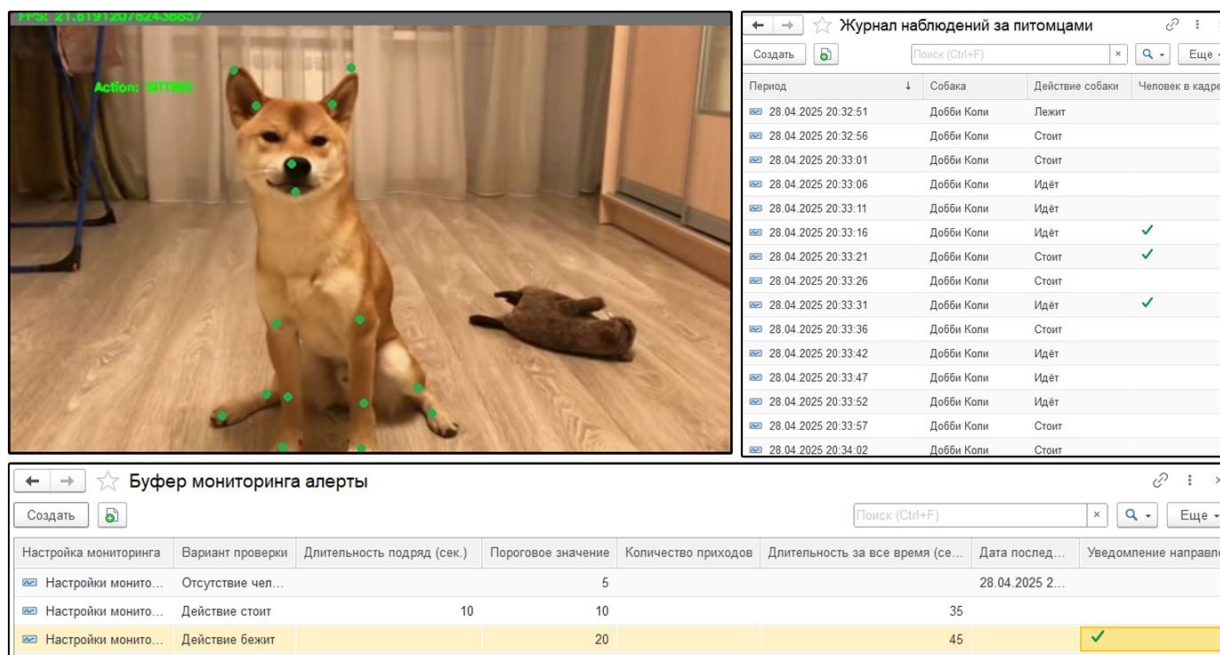


Рис. 3. – Экранные формы системы

Данная разработка позволяет автоматизировать процессы мониторинга поведенческих сценариев, исключая необходимость привлечения дополнительных людских ресурсов компании для решения данного вопроса. Компонент поддерживает работу в многопоточном режиме, обеспечивая тем самым параллельную обработку данных с нескольких источников. Таким образом, система способна непрерывно отслеживать поведенческие паттерны всех наблюдаемых объектов, обеспечивая актуальность сведений о состоянии каждого из них.

Литература

1. Shinde S., Kothari A., Gupta V. YOLO based Human Action Recognition and Localization // Procedia Computer Science. 2018, vol. 133, pp. 831-838. DOI: 10.1016/j.procs.2018.07.112.
2. Nguyen T., Scherer R., Van-Hung L. YOLO Series for Human Hand Action Detection and Classification from Egocentric Videos // Sensors. 2023. vol. 23, no. 6. DOI: 10.3390/s23063255.
3. Михайленко И. И., Нерубенко О. Е., Мещеряков О. Д. Процесс сегментации видеоданных для обучения глубоких нейронных сетей

автоматическому распознаванию функциональных состояний КРС // Аграрная Россия, 2024, № 4, С. 33-36. DOI: 10.30906/1999-5636-2024-4-33-36, EDN RBKIKW.

4. Устинова Л. Е., Гончаренко А. Н. Анализ использования Telegram-ботов при выполнении бизнес-задач // Информационные технологии. Проблемы и решения, 2022, № 1(18), С. 80-86, EDN EHKEIN.

5. Билалова И. М., Кувалов Р. С. Эффективность автоматизации процесса предоставления услуг посредством Telegram-бота // Финансовая экономика, 2018, № 7, С. 150-154, EDN YUTAPJ.

6. Ковалев А. В., Исаева А. С. Оценка качества семян пшеницы с использованием сверточной нейронной сети // Инженерный вестник Дона, 2021, № 12. URL: ivdon.ru/ru/magazine/archive/n12y2021/7354/.

7. Никитин Д. В., Тараненко И. С., Катаев А. В. Детектирование дорожных знаков на основе нейросетевой модели YOLO // Инженерный вестник Дона, 2023, № 7. URL: ivdon.ru/ru/magazine/archive/n7y2023/8531/.

8. Бобков А. В., Хтет А. Алгоритм распознавания лиц при их частичном закрытии // Автоматизация. Современные технологии, 2024, Т. 78, № 2, С. 77-87, DOI: 10.36652/0869-4931-2024-78-2-77-87, EDN AWQSLN.

9. Hochreiter S., Schmidhuber J. Long Short-Term Memory // Neural Comput, 1997, N 9 (8), pp. 1735–1780, DOI: 10.1162/neco.1997.9.8.1735.

10. Борзов С. М., Нежевенко Е. С., Потатуркин О. И. Исследование эффективности популярных нейросетевых моделей детектирования объектов в поле наблюдения // Вычислительные технологии, 2023, Т. 28, № 6, С. 68-80, DOI: 10.25743/ICT.2023.28.6.007, EDN BWNIXA.

References

1. Shinde S., Kothari A., Gupta V. Procedia Computer Science. 2018, vol. 133, pp. 831-838. DOI:10.1016/j.procs.2018.07.112.

2. Nguyen T., Scherer R., Van-Hung L. Sensors. 2023. vol. 23, no. 6. DOI: 10.3390/s23063255.

3. Mikhaylenko I. I., Nerubenko O. E., Meshcheryakov O. D. Agrarnaya Rossiya, 2024, no. 4, pp. 33-36, DOI: 10.30906/1999-5636-2024-4-33-36, EDN PBKIKW.

4. Ustinova L. E., Goncharenko A. N. Informacionny`e texnologii. Problemy` i resheniya, 2022, № 1(18), pp. 80-86, EDN EHKEIN.

5. Bilalova I. M., Kovalev R. S. Finansovaya e`konomika, 2018, no. 7, pp. 150-154, EDN YUTAPJ.

6. Kovalev A.V., Isaeva A. S. Inzhenernyj vestnik Dona, 2021, № 12. URL: ivdon.ru/ru/magazine/archive/n12y2021/7354/.

7. Nikitin D. V., Taranenko I. S., Kataev A.V. Inzhenernyj vestnik Dona, 2023, № 7. URL: ivdon.ru/ru/magazine/archive/n7y2023/8531/.

8. Bobkov A.V., Tet A. Avtomatizaciya. Sovremenny`e texnologii, 2024, VOL. 78, no. 2, pp. 77-87, DOI: 10.36652/0869-4931-2024-78-2-77-87, EDN AWQSLN.

9. Hochreiter S., Schmidhuber J. Neural Comput, 1997, N 9 (8), pp. 1735–1780, DOI: 10.1162/neco.1997.9.8.1735.

10. Borzov S. M., Nezhevenko E. S., Potaturkin O. I. Vy`chislitel`ny`e texnologii, 2023, Vol. 28, no. 6, pp. 68-80, DOI: 10.25743/ICT.2023.28.6.007, EDN BWNIXA.

Дата поступления: 9.04.2025

Дата публикации: 25.05.2025