

РАЗРАБОТКА НАВИГАЦИОННОЙ СИСТЕМЫ НА ОСНОВЕ РАСПОЗНАВАНИЯ ЭМОЦИЙ ПЕШЕХОДОВ

Цзи Юнань¹

Научный руководитель: Поберезкина Н.²

¹ТПУ, ИШИТР, зр. 158Т11, e-mail: ty05@tpu.ru

² ТПУ, ИШИТР, ОАР, старший преподаватель, e-mail: nikpi@tpu.ru

Аннотация

На основе YOLOv5s и ROS разработана навигационная система, учитывающая эмоциональное состояние пешеходов для корректировки траектории робота.

Ключевые слова: робототехническая навигация, распознавание эмоций, YOLOv5s, ROS.

Введение

В связи с широким спектром применения мобильных роботов в последние десятилетия наблюдается растущий интерес к исследованиям в области робототехники. От промышленных роботизированных рук и беспилотников до современных автономных наземных транспортных средств и разнообразных бытовых роботов – типы и функции роботов становятся все более разнообразными.

Люди не только уделяют все больше внимания использованию роботов, но и выдвигают более высокие требования к автономности, интеллекту и человечности роботов. Начиная с первых исследований автономной навигации мобильных роботов и заканчивая широкомасштабным появлением мобильных сервисных роботов, люди все больше склоняются к изучению взаимодействия между людьми и роботами с целью значительного повышения комфорта взаимодействия между людьми и роботами.

Традиционные алгоритмы навигации ориентированы только на обход препятствий и зачастую не учитывают эмоциональное состояние человека, что может приводить к возникновению конфликтных ситуаций и снижению качества взаимодействия. В результате этого клиенты могут понизить свою оценку, особенно в таких местах обслуживания, как гостиницы и рестораны. Поэтому в последние годы ученые стали рассматривать человеческие эмоции как важный фактор, влияющий на изучение роботов.

Нараянан и др. (2020) предложили метод анализа походки с использованием сверточной сети с несколькими видами графов, достигающий точности прогнозирования эмоций 82–85 % [1]. Бера и др. (2019) объединили байесовское мышление, глубокие нейронные сети и модель PAD для корректировки поведения робота на основе эмоциональных реакций других людей [2], в то время как Биеле и др. (2024) разработали систему, которая динамически регулирует область около тела, чтобы увеличить расстояние во время отрицательных эмоций и уменьшить расстояние во время положительных эмоций [3]. Но ни один из них не был реализован в реальных роботах.

В данной статье предлагается система, объединяющая модуль распознавания эмоций пешеходов на основе модели YOLOv5s с платформой ROS. В основе навигационного модуля ROS лежат проверенные алгоритмы глобального планирования (A*) и локального управления (DWA). Модуль распознавания эмоций обучается на наборе данных с использованием зрелой модели YOLO. Система позволяет роботу корректировать траекторию движения на основе эмоциональных реакций людей, что помогает повысить безопасность и улучшить взаимодействие с людьми. А упрощение и модульность системы способствуют массовому производству в реальных условиях.

Целью данного исследования было создание модели, которую можно было бы легко применить к роботам, а также продемонстрировать возможность учета эмоциональных

факторов при навигации робота, что открывает перспективы для дальнейшей интеграции и оптимизации системы в реальных условиях эксплуатации.

Общая архитектура системы и модульное разделение

Система имеет модульную конструкцию. Модуль сбора данных использует датчики, включая лидар и глубинные камеры, для получения информации в режиме реального времени и генерации данных об окружающей среде и целевых изображений.

Модуль распознавания эмоций на основе модели YOLOv5s идентифицирует эмоции пешеходов на основе полученных изображений и выводит категории эмоций, такие как положительные, нейтральные и отрицательные, а также соответствующие значения уверенности. Используем модель YOLOv5s, так как ее размер невелик, имеет высокую скорость отклика и подходит для развертывания на устройствах с меньшей вычислительной мощностью, таких как встроенные микроконтроллеры (Рис. 1).

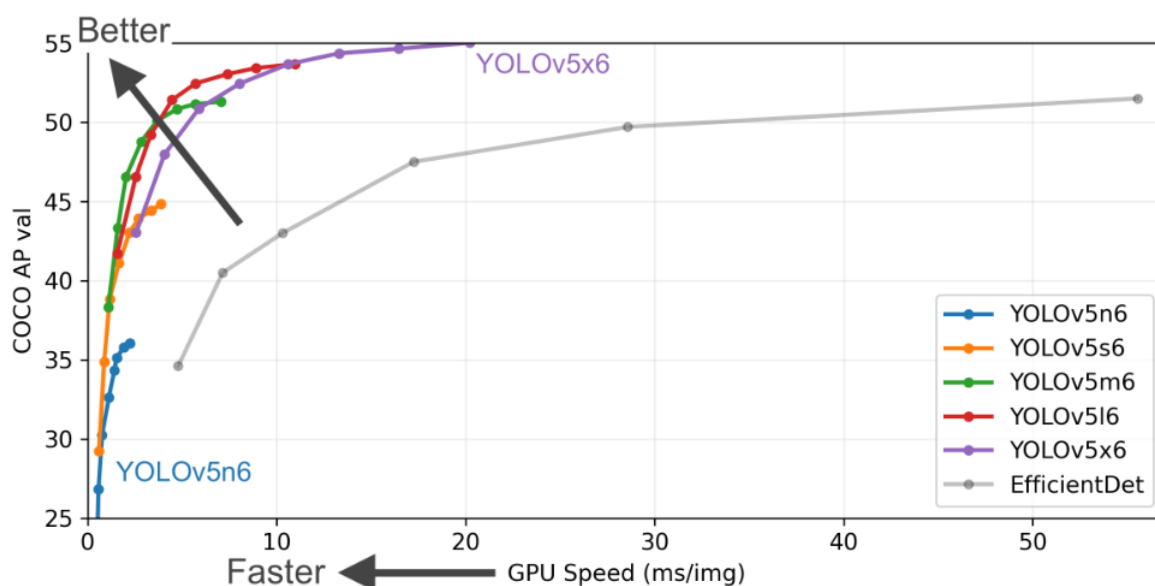


Рис. 1. Сравнение характеристик моделей YOLOv5s [4]

Модуль глобального планирования пути использует усовершенствованный алгоритм A* для прокладки глобального маршрута от текущего положения робота до целевой точки на основе карты окружающей среды, созданной на основе данных датчиков. Сначала система получает информацию об окружающей среде и данные о человеческих эмоциях с помощью глубинных камер и радиолокационных модулей; затем на основе полученной карты окружающей среды формируется карта первоначальных затрат. В процессе расчета стоимости алгоритм A* вычисляет расстояние между роботом и обнаруженным местоположением человека и добавляет дополнительное штрафное значение к стоимости соответствующей области на основе эмоционального состояния. Поэтому, если в окружающей среде присутствует отрицательная эмоция (например, гнев), алгоритм увеличивает стоимость области рядом с человеком, заставляя глобального планировщика выбирать объезд; в положительных эмоциональных состояниях значение штрафа уменьшается, заставляя робота приближаться. Динамически корректируя веса на основе информации о настройках, алгоритм планирования может обновлять глобальный маршрут в режиме реального времени в зависимости от текущей ситуации.

Модуль локального планирования использует алгоритм DWA для оптимизации локальной траектории и обрабатывает данные с лидара в режиме реального времени, чтобы избегать динамических препятствий.

Модуль управления навигацией и объединения данных отвечает за объединение результатов распознавания эмоций с глобальными и локальными данными планирования, формирование карт стоимости различных радиусов вокруг пешеходов с помощью заданных функций, тем самым корректируя стратегию движения робота и осуществляя обмен данными между модулями через ROS.

При построении функции мы ссылались на известную теорию социального пространства. Социальное пространство относится к пространственной дистанции, которую люди поддерживают в различных социальных и межличностных отношениях. Первоначально этот термин был придуман Холлом для описания контроля человека над пространством [5]. Пространственные интерактивные отношения между людьми включают интимные отношения, личные отношения, социальные отношения и общественные отношения (Рис. 2). Когда люди находятся слишком близко друг к другу, они часто чувствуют себя некомфортно, и их личное пространство нарушается. Поэтому крайне важно по возможности избегать вторжения в личное пространство других людей.

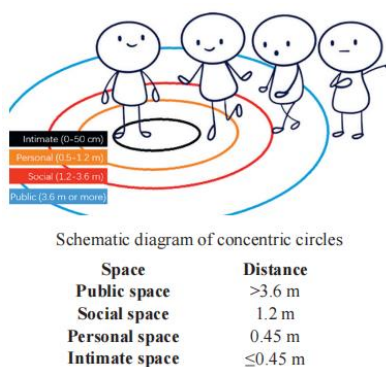


Рис. 2. Социальное пространство [5]

Эти расстояния были выбраны для имитации идеального диапазона взаимодействия между роботами и людьми, способствуя естественному и комфортному взаимодействию для обеих сторон. Ниже приведена схематическая диаграмма архитектуры системы, иллюстрирующая поток данных и интерфейсы между модулями:

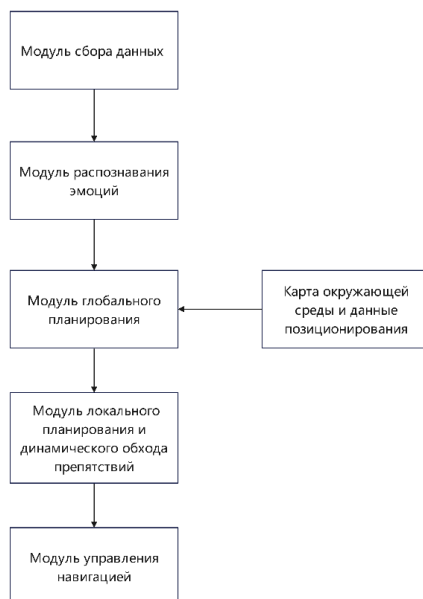


Рис. 3. Архитектура системы

Результат

1. Анализ эффективности модели распознавания эмоций:

Для обучения модели распознавания эмоций на основе YOLO использовали исходную модель YOLOv5s и набор данных COCO и обучили ее в течение 268 эпох с размером изображения 96×96 и размером пакета 64. Включили кэширование изображений для ускорения загрузки данных, отключили обучение со смешанной точностью для обеспечения стабильности параметров и установили задержку для буферизации нагрузки на сервер во время непрерывного обучения. Такая конфигурация позволяет модели эффективно различать различные эмоциональные состояния, обеспечивая при этом баланс между эффективностью обучения и точностью модели.

(1) Анализ матрицы ошибок

Из нормализованной матрицы путаницы мы можем видеть, что модель демонстрирует наилучшие результаты классификации в категории «Happy» с точностью 83%, что намного выше, чем в других категориях, что свидетельствует о том, что модель обладает сильной способностью различать положительные эмоциональные проявления. Эффекты распознавания «Contempt» (75%) и «Disgust» (66%) также вполне идеальны.

Относительно говоря, точность распознавания «Neutral» (52%), «Sad» (57%) и «Fear» (59%) относительно низкая, а доля «Гнева», ошибочно классифицированного как фон, достигает 28%, что свидетельствует о том, что модель испытывает большие трудности с обработкой категорий отрицательных или слабо выраженных эмоций. Более того, существует путаница между некоторыми категориями, такими как «Гнев» и «Отвращение», «Грусть» и «Neutral».

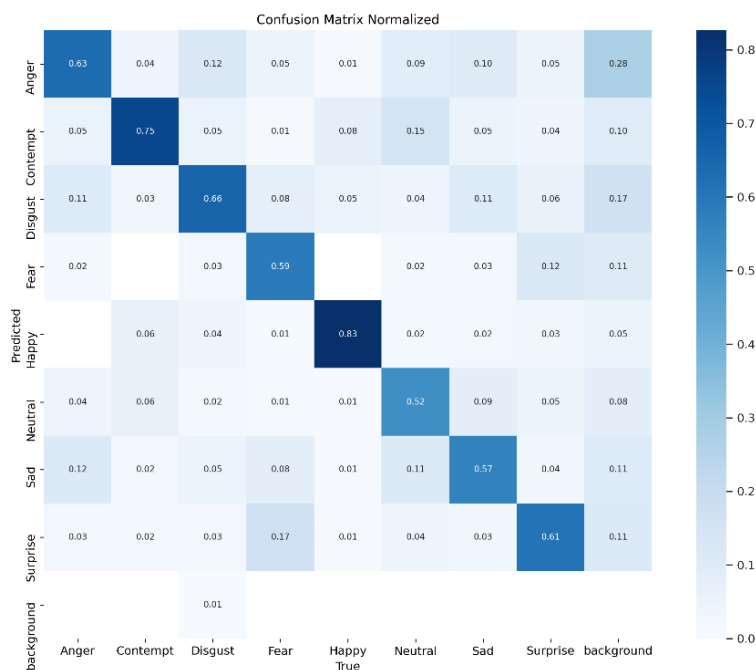


Рис. 4. Матрица ошибок

(2) Кривая F1-Confidence

Кривая F1-уверенности показывает, что при установке порога уверенности на уровне 0,388 средний балл F1 модели достигает оптимального значения 0,66. Стоит отметить, что категория «Happy» показывает высокое значение F1 в каждом доверительном интервале, что подтверждает, что она имеет самую сильную различимость в модели. Напротив, значения F1 таких категорий, как «Neutral», «Sad» и «Fear», значительно снизились в средних и высоких доверительных интервалах, что указывает на то, что модель не может уверенно различать эти эмоции (рис. 5, а).

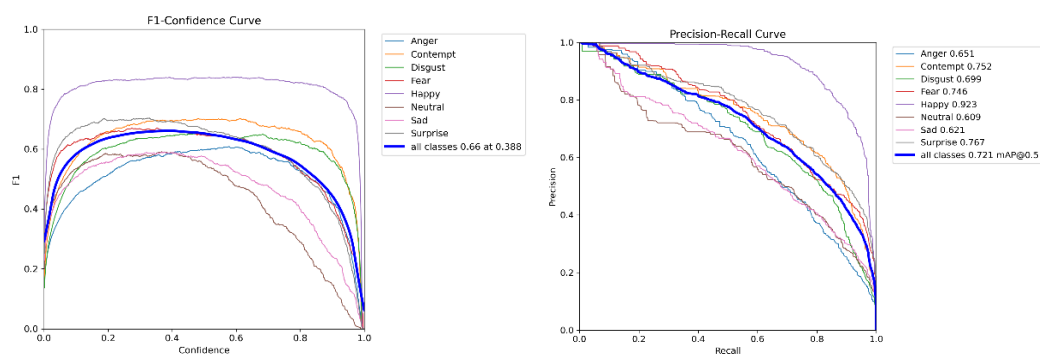


Рис. 5. Кривая F1-уверенности

(3) Точность и отзыв

mAP (@0,5) каждого типа эмоций, рассчитанный на основе кривой PR, выглядит следующим образом (рис. 5, б):

Таблица 1 – Результат mAP@0.5

Категория	mAP@0.5
Happy	0.923
Surprise	0.767
Contempt	0,752
Fear	0,746
Disgust	0,699
Anger	0,651
Sad	0.621
Neutral	0,609

Общее значение mAP@0.5 модели составляет 0,721, что указывает на высокую общую точность обнаружения, особенно при определении очень экспрессивных эмоций. Однако для нейтральных и отрицательных эмоций есть возможности для дальнейшего улучшения точности и отзывчивости модели.

(4) Динамика функции потерь

Графики изменений box_loss, cls_loss и dfl_loss во время обучения и проверки показывают стабильную сходимость и отсутствие признаков переобучения. Показатели точности, полноты и mAP также неуклонно улучшаются, что подтверждает правильность выбора параметров обучения и общую устойчивость модели (рис. 6).

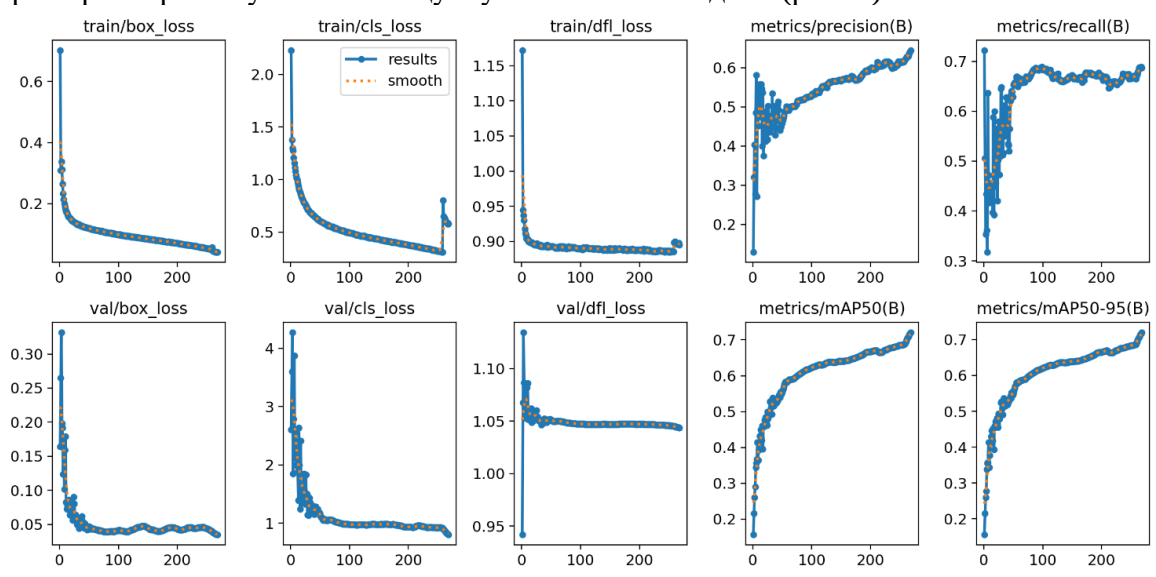


Рис. 6. Результаты тестирования модели YOLOv5s

Ниже приведен результат обнаружения одного изображения.



Рис. 7. Результат модели распознавания эмоций

2. Анализ навигационной модели:

Алгоритм A^* был выбран в качестве глобального планирования, так как он быстро реагирует на окружающую среду, его метод поиска прост и широко используется.

Для локального планирования была опробована модель прогностического управления MPC. Во время работы робот перестает двигаться в относительно узком пространстве, хотя на самом деле он может через него пройти (рис. 8(a)). Это связано с тем, что модель MPC основана на точной динамической модели (такой как уравнения состояния и соотношения «вход-выход»), а ошибки модели могут привести к сбою управления. Более того, оптимизация, основанная на градиентном спуске, склонна приводить к локальным оптимальным решениям (например, к путям, упирающимся в препятствия) (рис. 8(б)).

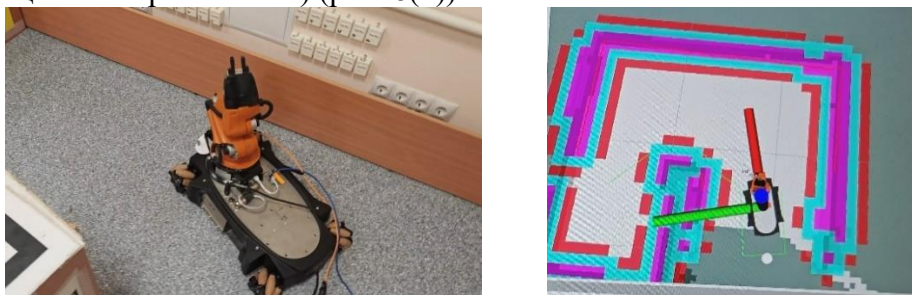


Рис. 8. Проблемы MPC

Затем мы попытались использовать алгоритм DWA вместо MPC. Он обновляет данные о препятствиях на основе последних данных датчиков, поэтому работает в режиме реального времени и может быстро менять стратегию планирования маршрута для динамически меняющихся сценариев. При этом точная динамическая модель не требуется, нужны лишь ограничения по скорости. Более того, DWA может работать в режиме реального времени на устройствах с низкой вычислительной мощностью (таких как STM32), что подходит для потребительских роботов и образовательных роботов. Результаты показаны на рисунке 9.

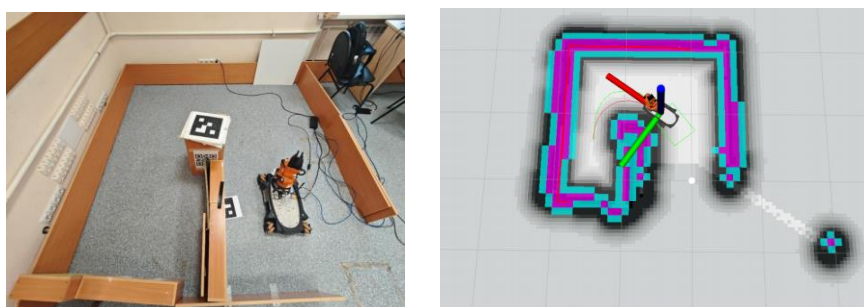


Рис. 9. Результаты навигации робота и обхода препятствий

Тестирование по отдельным модулям является необходимым условием для дальнейшей интеграции моделей в единую систему.

Заключение

Основная область применения данного алгоритма – сервисные роботы в общественных местах. Модульная архитектура обеспечивает гибкость системы и ее адаптируемость к различным сценариям. Первоначальные испытания показали высокую эффективность каждого модуля, подтвердив перспективность данного подхода к созданию систем социальной навигации, учитывающих эмоциональное состояние пешеходов для повышения безопасности и комфорта. Последующие исследования позволят оптимизировать сопоставление данных каждой модели для достижения плавной передачи данных и производительности в реальном времени. И в конечном итоге будет сформирована полная система.

Список использованных источников

1. Narayanan V., Manoghar B.M., Dorbala V.S., et al. Proximo: Gait-based emotion learning and multi-view proxemic fusion for socially-aware robot navigation[C] //2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2020: 8200-8207.
2. Bera A., Randhavane T., Prinja R., et al. The emotionally intelligent robot: Improving social navigation in crowded environments[J]. arXiv preprint arXiv:1903.03217, 2019.
3. Bilen B., Kivrak H., Uluer P., et al. Social robot navigation with adaptive proxemics based on emotions[J]. arXiv preprint arXiv:2401.17663, 2024.
4. YOLOv5 [Электронный ресурс]. – URL: github.com/ultralytics/yolov5 (дата обращения 09.02.2025)
5. Duzhesheng L. Comfort navigation improvement of path planning task in human-robot interaction / L. Duzhesheng, S. A. Chepinskiy, W. Jian // Journal of Instrument Engineering. – 2024. – Vol. 67, No. 6. – P. 481-491. – DOI 10.17586/0021-3454-2024-67-6-481-491. – EDN BDWKXK.