

SISTEMA DE NAVEGAÇÃO INTERNA UTILIZANDO APRENDIZADO POR REFORÇO¹

Rogério Barbosa Chaves², Alessandra Oliveira Andrade³

RESUMO:

Esta pesquisa encontra-se em andamento e busca apresentar um sistema de navegação autônoma para robôs em ambientes internos desconhecidos. A partir dos recentes avanços no campo do aprendizado por reforço profundo e no enorme desenvolvimento do campo de navegação autônomas de robôs busca-se aplicar o algoritmo Twin Delayed DDPG, também chamado de TD3, e avaliar o seu desempenho da navegação do robô, isto é, a capacidade do robô de navegar em ambientes internos evitando colisões com objetos.

PALAVRAS-CHAVE: aprendizado por reforço, navegação autônoma, robótica.

ABSTRACT:

This research is ongoing and seeks to present an autonomous navigation system for robots in unknown indoor environments. Based on recent advances in the field of deep reinforcement learning and the enormous development in the field of autonomous robot navigation, we seek to apply the Twin Delayed DDPG algorithm, also called TD3, and evaluate its robot navigation performance, i.e., the robot's ability to navigate indoor environments avoiding collisions with objects.

KEYWORDS: reinforcement learning, autonomous navigation, robotics.

INTRODUÇÃO:

A capacidade de robôs se moverem de forma autônoma em um ambiente interno desconhecido, sem colidir ou ficar preso em objetos, é uma importante tarefa para robôs móveis. Conforme (LEE, YUSUF, 2022), navegação é o mecanismo pela qual um robô móvel desempenha uma certa tarefa em um ambiente. Além disso, navegação autônoma é quando um robô se move em um ambiente sem intervenção de um controlador externo, sendo esta, um tópico de pesquisa chave para o campo de robôs móveis.

Embora, nos últimos anos tenha havido um enorme desenvolvimento no campo de navegação autônoma de robôs móveis, com os benefícios dos avanços do campo da IA e de visão computacional, ainda há desafios para permitir que robôs possam navegar autonomamente no mundo real (LEE, YUSUF, 2022). Com isso, cabe destacar, o grande aumento da popularidade do aprendizado por reforço, iniciado com a combinação de aprendizado por reforço e deep learning atingindo ótimos resultados.

No aprendizado por reforço, o agente (em nossa situação, o robô), irá interagir continuamente com o ambiente, primeiramente obtendo as informações do estado atual ambiente, em seguida o agente toma uma ação neste ambiente que o move para um novo estado e retorna

¹ Fundação de Amparo à Pesquisa do Estado da Bahia – FAPESB.

² Discente do curso de Ciência da Computação da Universidade Estadual do Sudoeste da Bahia.

³ Docente do Departamento de Ciências Exatas e Tecnológicas da Universidade Estadual do Sudoeste da Bahia.

uma recompensa correspondente. Assim, a tarefa do agente é maximizar a recompensa recebida ao longo tempo.

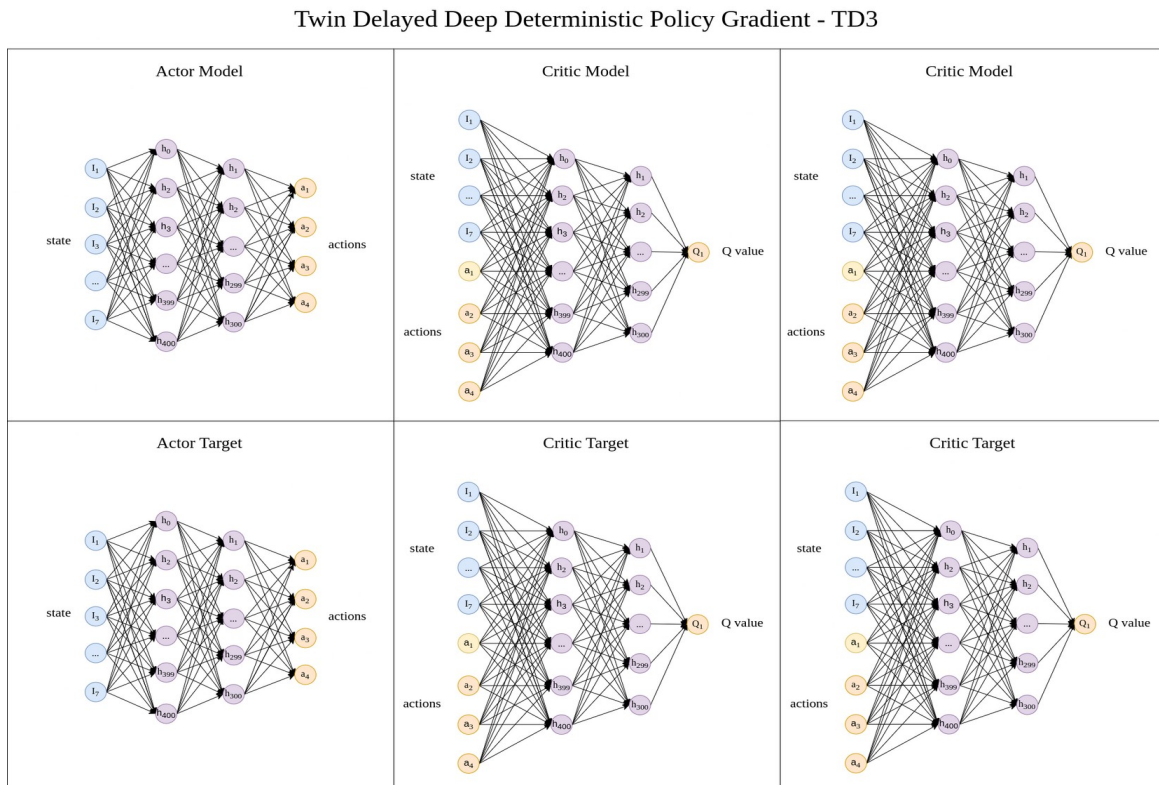
De forma mais específica, conforme (SUTTON, BARTO, 2018), o agente e o ambiente interagem em cada uma de uma sequência discreta de passos de tempo, $t = 0,1,2,3, \dots$. A cada passo de tempo t , o agente recebe alguma representação estado do ambiente, $S_t \in S$, e com base nisso seleciona uma ação A_t , em parte como consequência da sua ação, o agente recebe uma recompensa numérica, R_{t+1} , e se encontra em um novo estado, S_{t+1} .

MATERIAIS E MÉTODOS:

I. Sistema de navegação autônoma:

Foi realizado a implementação do algoritmo de aprendizado por reforço, Twin Delayed Deep Deterministic Policy Gradient, também chamado de TD3 ou Twin Delayed DDPG (FUJIMOTO, HOOF, MEGER, 2018), com a utilização da linguagem Python e com o PyTorch framework, na figura 1 podemos visualizar a arquitetura deste algoritmo; e o algoritmo de detecção de objetos YOLO v4 (BOCHKOVSKIY, WANG, LIAO, 2020), pré-treinado com o conjunto de dados COCO, com a linguagem Python e com a biblioteca OpenCV. Além disso, haverá uma integração entre os algoritmos TD3 e YOLO v4 e com o robô móvel que gerenciará as interações destes no sistema.

FIGURA 1 – Arquitetura da Twin Delayed Deep Determinintic Policy Gradient – TD3

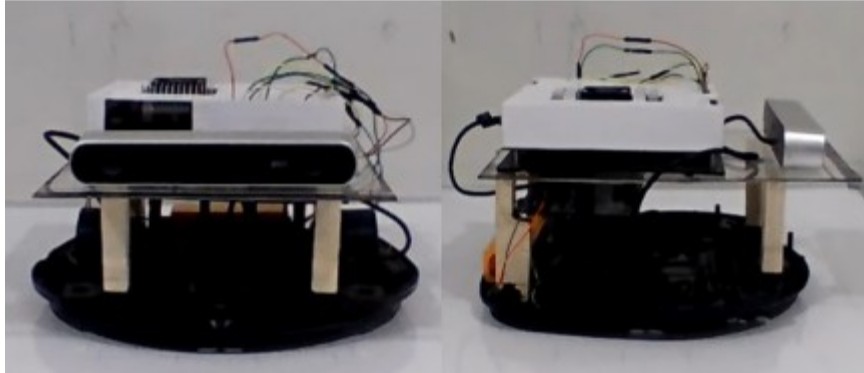


Fonte: De autoria própria.

II. Robô móvel:

Para o desenvolvimento do robô móvel foram utilizados a Placa Nvidia Jetson TX2 juntamente com: o driver controlador de motores CC Ponte H L298N, o sensor de acelerômetro, giroscópio MPU 6050 e a câmera estéreo ZED da StereoLabs. Como pode ser visto na figura 2.

FIGURA 2 – Robô Móvel.



Fonte: De autoria própria.

Todos os periféricos do robô móvel foram implementados com a linguagem Python com as seguintes bibliotecas: a biblioteca Jeton.GPIO para o driver controlador dos motores, esta biblioteca permite o uso das portas programáveis de entrada e saída de dados da placa Jetson TX2; a biblioteca disponibilizada pela Adafruit para o uso do sensor MPU 6050. Por fim, utilizamos a câmera ZED juntamente com a sua API para Python para obter a distância do objeto imediatamente a frente do robô e para capturar a imagem que será processada pelo algoritmo de detecção de objetos YOLO v4.

RESULTADOS E DISCUSSÕES:

Para desenvolver este projeto foram necessárias muitas pesquisas e avaliações, o que incluiu a definição do uso de aprendizado por reforço no desafio de navegação autônoma. Para isso, foi preciso ser dividido entre três partes: visão computacional, aprendizado por reforço e robótica. Para cada uma destas partes foram realizadas revisões teóricas dos algoritmos e das técnicas utilizadas.

No campo de visão computacional foi realizado estudo teórico e prático sobre o algoritmo de detecção de objetos YOLO v4 e dos algoritmos de segmentação de imagens (técnicas clássicas, segmentação de instância e segmentação semântica).

Para o aprendizado por reforço, foram realizados estudos teóricos sobre o que é o aprendizado por reforço em si e alguns de seus métodos de solução. Podemos citar métodos de solução tabular: bandidos multiarmados e processos de decisão de Markov finitos. Além disso, da mesma forma, foram realizadas revisões teóricas dos métodos: Q-Learning, Deep Q-Learning e TD3 com a escolha deste último para a implementação e teste com o ambiente Ant⁴.

⁴ Ant é um ambiente que é parte dos ambientes do Mujoco que está disponível na biblioteca Gym.

Na parte de robótica foram realizados estudos e pesquisas sobre controle de motores CC, pois a partir do controle destes, o agente poderá realizar as suas ações no ambiente. Além disso, também foram feitos estudos e pesquisas sobre a utilização de sensores, de quais sensores seriam relevantes e de como implementá-los.

Por fim, dado que o projeto ainda está em andamento, tendo sido concluídas as etapas do algoritmo de aprendizado por reforço TD3, do algoritmo de detecção de objetos YOLO v4, dos testes dos sensores e a construção do protótipo do robô móvel, está em andamento a etapa da implementação do sistema que realizará a integração de todas as outras partes para que se prossiga para o treinamento do modelo e do teste.

CONCLUSÕES E CONSIDERAÇÕES:

O projeto encontra-se em andamento, mais especificamente na etapa que antecede as etapas de treinamento e teste do sistema. Claramente, não se trata-se de um processo terminado, pois dentro da literatura existente há um grande espaço para o desenvolvimento de estudos e pesquisas que permitam melhorar a solução aqui apresentada. Cabe destacar que quando tratamos de problemas de navegação autônoma, a percepção das informações do ambiente são cruciais para o sucesso deste. Neste sentido, considero que este trabalho pode obter melhorias, principalmente no que se refere ao uso de outras soluções de visão computacional.

REFERÊNCIAS

1. BOCHKOVSKIY, Alexey; WANG, C-Y; LIAO, H-Y M. YOLOv4: optimal speed and accuracy of object detection. Arxiv, 2004.10934, abril, 2020. Disponível em: <<https://arxiv.org/abs/2004.10934>>. Acesso em: 22 de setembro de 2023.
2. FUJIMOTO, Scott; HOOFF, Herke van; MEGER, David. Addressing Function Approximation Error in Actor-Critic Methods. ArXiv:1802.09477, October, 2018. Disponível em: <<https://arxiv.org/abs/1802.09477>>. Acesso em 28 de setembro de 2023.
3. LEE, M.-F.R.; YUSUF, S.H. Mobile Robot Navigation Using Deep Reinforcement Learning. Processes. 2022. volume 10 (12). Disponível em: <<https://doi.org/10.3390/pr10122748>>. Acesso em: 22 de setembro de 2023.
4. SUTTON, Richard S.; BARTO, Andrew G. Reinforcement Learning: an introduction. 2ª ed. Cambridge, MA: The MIT Press. 2018.