

**Universidade de Pernambuco**  
**Programa de Pós-Graduação em Engenharia da**  
**Computação (PPGEC)**

**Proposta de Dissertação de Mestrado**

**Área: Computação Inteligente**

**Título: Abordagem Não-Invasiva de Pessoas por um Robô Usando Aprendizagem por Reforço Profundo**

**Orientador – Bruno José Torres Fernandes ([bjtf@ecomp.poli.br](mailto:bjtf@ecomp.poli.br))**

**Descrição** – Robôs inteligentes deram seus primeiros passos em cenários industriais, sendo este o local da maioria das aplicações de robôs pré-programados até o momento [1]. Atualmente, os robôs estão cada vez mais presentes no cotidiano com o surgimento de novas tecnologias, principalmente assistivas, que almejam fazer do robô um assistente para o dia-a-dia. Espera-se, portanto, que os robôs aprendam a executar tarefas que são consideradas bastante simples para os seres humanos. No entanto, para que um robô atinja um desempenho semelhante ao humano, é necessário realizar diversas subtarefas para concluir satisfatoriamente um determinado objetivo. Essas subtarefas incluem percepção, compreensão do ambiente, estratégias de aprendizagem, representação do conhecimento, conscientização de seu próprio estado, manipulação do ambiente, entre outras [2].

Aprendizagem por Reforço [3] é uma abordagem de aprendizagem que permite a agentes autônomos aprender usando feedback recebido de um ambiente [4]. A ideia é apoiada por psicologia comportamental, baseada em como as pessoas ou os animais aprendem [5]. Em cada etapa de tempo, o agente seleciona uma ação e observa o que acontece no ambiente, às vezes, obtendo uma recompensa ou uma punição. Se as ações levam a melhores situações, há uma tendência de aplicar tal comportamento novamente, caso contrário, a tendência é evitar esse comportamento no futuro. Assim o problema é reduzido para aprender como selecionar ações em cada situação para atingir um determinado objetivo, tentando maximizar a recompensa obtida do ambiente [4]. A aprendizagem do reforço tem sido bastante usada em robôs onde eles tentam executar uma tarefa em um ambiente humano, seja através do desenvolvimento de um trabalho ou interagir com pessoas [6, 7, 8]. As características desses ambientes estão em constante mudança, dificultando que um agente se adapte a essas mudanças e atinja seu objetivo.

Descobertas recentes em visão computacional e reconhecimento de fala têm se baseado em treinar eficientemente redes neurais em conjuntos de treinamento muito grandes. O aprendizado profundo [9] é composto de muitas camadas de processamento e foi testado com sucesso, entre outros, na classificação de imagens, representando diferentes níveis de abstração [10]. Além disso, a aprendizagem de reforço profundo [11] combinou as duas abordagens acima mencionadas, operando diretamente em imagens RGB e processando dados de treinamento com eficiência. Usando uma rede neural profunda, um agente aprende um mapeamento de políticas de um conjunto de situações para um conjunto de ações.

A proposta deste projeto é desenvolver uma abordagem de aprendizagem de reforço profundo aplicada em um problema de interação social. Neste contexto, considera-se um cenário humano-robô, no qual se espera que o robô aprenda abordar pessoas de maneira não invasiva. Assim, é importante que o robô consiga identificar pessoas e suas emoções continuamente à medida que executa sua abordagem, observando os estados do ambiente usando redes neurais de arquitetura profunda [12, 13, 14] e decida as ações a serem executadas por meio do método de aprendizado por reforço.

### Referências Bibliográficas

- [1] T. S. Tadele, T. de Vries, and S. Stramigioli, "The safety of domestic robotics: A survey of various safety-related publications," *IEEE Robot. Autom. Mag.*, vol. 21, no. 3, pp. 134–142, Sep. 2014.
- [2] A. Cangelosi and M. Schlesinger, *Developmental Robotics: From Babies to Robots*. Cambridge, MA, USA: MIT Press, 2015.
- [3] R. S. Sutton and A. G. Barto. *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: Bradford Book, 1998.
- [4] Rieser, V. and Lemon, O. *Reinforcement Learning for Adaptive Dialogue Systems*. Heidelberg, Germany: Springer, 2011.
- [5] Niv, Y. Reinforcement learning in the brain. *Journal of Mathematical Psychology*, 53:139–154, 2009.
- [6] F. Cruz, J. Twiefel, S. Magg, C. Weber, and S. Wermter, "Interactive reinforcement learning through speech guidance in a domestic scenario," in *2015 International Joint Conference on Neural Networks (IJCNN)*, 2015, vol. 2015-Septe, pp. 1–8.
- [7] S. Gu, E. Holly, T. Lillicrap, and S. Levine, "Deep Reinforcement Learning for Robotic Manipulation with Asynchronous Off-Policy Updates," Oct. 2016.
- [8] L. A. León, A. C. Tenorio, and E. F. Morales, "Human Interaction for Effective Reinforcement Learning," in *European Conf. Mach. Learning and Principles and Practice of Knowledge Discovery in Databases (ECMLPKDD 2013)*, 2013.
- [9] I. Goodfellow, Y. Bengio, and A. Courville. *Deep learning*. Cambridge: MIT press, 2015.
- [10] Y. LeCun, Y. Bengio, and G. Hinton. *Deep learning*. *Nature*, Vol. 521, Nr. 7553, pp. 436-444, 2015.
- [11] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller. *Playing atari with deep reinforcement learning*. arXiv preprint arXiv:1312.5602, 2013.
- [12] B. J. T. Fernandes, G. D. C. Cavalcanti e T. I. Ren. Lateral Inhibition Pyramidal Neural Network for Image Classification. *IEEE Transactions on Cybernetics*, v. 43 (6), p. 2082-2092, 2013.
- [13] B. J. T. Fernandes, G. D. C. Cavalcanti e T. I. Ren. Ren. AutoAssociative Pyramidal Neural Network for one class pattern classification with implicit feature extraction. *Expert Systems with Applications*, v. 40 (18) p. 7258-7266, 2013.
- [14] A. M. Soares, B. J. T. Fernandes e C. J. A. Bastos-Filho. Structured Pyramidal Neural Networks. *International Journal of Neural Systems*, v. 1, p. 1750021, 2017.