

## DAFTAR PUSTAKA

- [1] D. Baldazo, J. Parras, and S. Zazo, “Decentralized Multi-Agent Deep Reinforcement Learning in Swarms of Drones for Flood Monitoring,” in *2019 27th European Signal Processing Conference (EUSIPCO)*, 2019, pp. 1–5, doi: 10.23919/EUSIPCO.2019.8903067.
- [2] J. Yang, L. Liu, Q. Zhang, and C. Liu, “Research on Autonomous Navigation Control of Unmanned Ship Based on Unity3D,” in *2019 5th International Conference on Control, Automation and Robotics (ICCAR)*, 2019, pp. 422–426, doi: 10.1109/ICCAR.2019.8813722.
- [3] T. Campbell and J. M. Hereford, “Scalability of robot swarms when applied to maze solving,” in *SoutheastCon 2015*, 2015, pp. 1–7, doi: 10.1109/SECON.2015.7132906.
- [4] N. Saputro, “Game-Theoretic and Genetic-Based Approach for Cooperative Mission-Oriented Swarms of Drones,” in *2019 International Conference on Mechatronics, Robotics and Systems Engineering (MoRSE)*, 2019, pp. 163–168, doi: 10.1109/MoRSE48060.2019.8998726.
- [5] M. O. A. Aqel, A. Issa, M. Khair, M. ElHabbash, M. AbuBaker, and M. Massoud, “Intelligent Maze Solving Robot Based on Image Processing and Graph Theory Algorithms,” in *2017 International Conference on Promising Electronic Technologies (ICPET)*, 2017, pp. 48–53, doi: 10.1109/ICPET.2017.15.
- [6] E. Kivelevitch and K. Cohen, “Multi-Agent Maze Exploration,” *J. Aerosp. Comput. Inf. Commun.*, vol. 7, pp. 391–405, 2010, doi: 10.2514/1.46304.
- [7] M. Rafiei and M. Sina, “Choatic GA Based Q-Learning in Nondeterministic Maze Benchmark,” in *2015 15th International Conference on Computational Science and Its Applications*, 2015, pp. 114–118, doi: 10.1109/ICCSA.2015.27.
- [8] J. Hu, H. Niu, J. Carrasco, B. Lennox, and F. Arvin, “Voronoi-Based Multi-Robot Autonomous Exploration in Unknown Environments via Deep Reinforcement Learning,” *IEEE Trans. Veh. Technol.*, vol. 69, no. 12, pp. 14413–14423, 2020, doi: 10.1109/TVT.2020.3034800.
- [9] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 3rd ed. Prentice Hall, 2010.
- [10] R. Bellman, “A Markovian Decision Process,” *J. Math. Mech.*, vol. 6, no. 5, pp. 679–684, 1957, [Online]. Available: <https://www.jstor.org/stable/24900506>.
- [11] D. Lee, H. Seo, and M. W. Jung, “Neural Basis of Reinforcement Learning and Decision Making,” *Annu. Rev. Neurosci.*, vol. 35, no. 1, pp. 287–308,

- 2012, doi: 10.1146/annurev-neuro-062111-150512.
- [12] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, Second. The MIT Press, 2018.
  - [13] C. J. C. H. Watkins and P. Dayan, “Technical Note: Q-Learning,” *Mach. Learn.*, vol. 8, no. 3, pp. 279–292, 1992, doi: 10.1023/A:1022676722315.
  - [14] D. Pandey and P. Pandey, “Approximate Q-Learning: An Introduction,” in *2010 Second International Conference on Machine Learning and Computing*, Feb. 2010, pp. 317–320, doi: 10.1109/ICMLC.2010.38.
  - [15] C. Gaskett, “Q-Learning for Robot Control,” 2002.
  - [16] V. Konda and J. Tsitsiklis, “Actor-Critic Algorithms,” *Soc. Ind. Appl. Math.*, vol. 42, 2001.
  - [17] J. Wang and I. C. Paschalidis, “An Actor-Critic Algorithm With Second-Order Actor and Critic,” *IEEE Trans. Automat. Contr.*, vol. 62, no. 6, pp. 2689–2703, Jun. 2017, doi: 10.1109/TAC.2016.2616384.
  - [18] Y. P. Awate, “Policy-Gradient Based Actor-Critic Algorithms,” in *2009 WRI Global Congress on Intelligent Systems*, May 2009, vol. 3, pp. 505–509, doi: 10.1109/GCIS.2009.372.
  - [19] S. Li, “Multi-Agent Deep Deterministic Policy Gradient for Traffic Signal Control on Urban Road Network,” in *2020 IEEE International Conference on Advances in Electrical Engineering and Computer Applications (AEECA)*, 2020, pp. 896–900, doi: 10.1109/AEECA49918.2020.9213523.
  - [20] R. Lowe, Y. Wu, A. Tamar, J. Harb, P. Abbeel, and I. Mordatch, “Multi-Agent Actor-Critic for Mixed Cooperative-Competitive Environments,” *Neural Inf. Process. Syst.*, 2017.
  - [21] W. KONG, D. ZHOU, and Z. YANG, “Air Combat Strategies Generation of CGF Based on MADDPG and Reward Shaping,” in *2020 International Conference on Computer Vision, Image and Deep Learning (CVIDL)*, 2020, pp. 651–655, doi: 10.1109/CVIDL51233.2020.000-7.
  - [22] P. Zhu, W. Dai, W. Yao, J. Ma, Z. Zeng, and H. Lu, “Multi-Robot Flocking Control Based on Deep Reinforcement Learning,” *IEEE Access*, vol. 8, pp. 150397–150406, 2020, doi: 10.1109/ACCESS.2020.3016951.
  - [23] L. Wang, K. Wang, C. Pan, W. Xu, N. Aslam, and L. Hanzo, “Multi-Agent Deep Reinforcement Learning Based Trajectory Planning for Multi-UAV Assisted Mobile Edge Computing,” *IEEE Trans. Cogn. Commun. Netw.*, p. 1, 2020, doi: 10.1109/TCCN.2020.3027695.
  - [24] L. Angel, C. Hernández, and C. Díaz-Quintero, “Modeling, simulation and control of a differential steering type mobile robot,” in *Proceedings of the 32nd Chinese Control Conference*, 2013, pp. 8757–8762.
  - [25] G. Dudek and M. Jenkin, *Computational Principles of Mobile Robotics*, 2nd ed. USA: Cambridge University Press, 2010.