REFERENCES

[1] Yossi Adi, Carsten Baum, Moustapha Cisse, Benny Pinkas, and Joseph Keshet. 2018. Turning your weakness into a strength: Watermarking deep neural networks by backdooring. In *USENIX Security Symposium*. 1615–1631.

[2] William Aiken, Hyoungshick Kim, and Simon Woo. 2020. Neural Network Laundering: Removing Black-Box Backdoor Watermarks from Deep Neural Networks. *arXiv preprint arXiv:2004.11368* (2020).

[3] Vahid Behzadan and William Hsu. 2019. Sequential triggers for watermarking of deep reinforcement learning policies. *arXiv preprint arXiv:1906.01126* (2019).

[4] Greg Brockman, Vicki Cheung, Ludwig Pettersson, Jonas Schneider, John Schulman, Jie Tang, and Wojciech Zaremba. 2016. OpenAI Gym. arXiv:arXiv:1606.01540

[5] Anthony R Cassandra. 1998. A survey of POMDP applications. In *Working notes of AAAI 1998 Fall Symposium on Planning with Partially Observable Markov Decision Processes*, Vol. 1724.

[6] Xinyun Chen, Wenxiao Wang, Chris Bender, Yiming Ding, Ruoxi Jia, Bo Li, and Dawn Song. 2019. REFIT: A Unified Watermark Removal Framework for Deep Learning Systems with Limited Data. *arXiv preprint arXiv:1911.07205* (2019).

[7] Ingemar J Cox, Joe Kilian, F Thomson Leighton, and Talal Shamoon. 1997. Secure spread spectrum watermarking for multimedia. *IEEE Transactions on Image Processing* 6, 12 (1997), 1673–1687.

[8] Prafulla Dhariwal, Christopher Hesse, Oleg Klimov, Alex Nichol, Matthias Plappert, Alec Radford, John Schulman, Szymon Sidor, Yuhuai Wu, and Peter Zhokhov. 2017. OpenAI Baselines. https://github.com/openai/baselines.

[9] Shixiang Gu, Ethan Holly, Timothy Lillicrap, and Sergey Levine. 2017. Deep reinforcement learning for robotic manipulation with asynchronous off-policy updates. In *IEEE International Conference on Robotics and Automation*. 3389–3396.

[10] Shangwei Guo, Tianwei Zhang, Han Qiu, Yi Zeng, Tao Xiang, and Yang Liu. 2020. The Hidden Vulnerability of Watermarking for Deep Neural Networks. *arXiv preprint arXiv:2009.08697* (2020).

[11] Suyog Gupta, Ankur Agrawal, Kailash Gopalakrishnan, and Pritish Narayanan. 2015. Deep learning with limited numerical precision. In *International Conference on Machine Learning*. 1737–1746.

[12] Song Han, Huizi Mao, and William J Dally. 2015. Deep compression: Compressing deep neural networks with pruning, trained quantization and huffman coding. *arXiv preprint arXiv:1510.00149* (2015).

[13] Panagiota Kiourti, Kacper Wardega, Susmit Jha, and Wenchao Li. 2019. TrojDRL: Trojan Attacks on Deep Reinforcement Learning Agents. arXiv:1903.06638

[14] Solomon Kullback and Richard A Leibler. 1951. On information and sufficiency. *The annals of mathematical statistics* 22, 1 (1951), 79–86.

[15] Erwan Le Merrer, Patrick Perez, and Gilles Trédan. 2019. Adversarial frontier stitching for remote neural network watermarking. *Neural Computing and Applications* (2019), 1–12.

[16] Zheng Li, Chengyu Hu, Yang Zhang, and Shanqing Guo. 2019. How to prove your model belongs to you: A blind-watermark based framework to protect intellectual property of DNN. In *Annual Computer Security Applications Conference*. 126–137.

[17] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al. 2015. Human-level control through deep reinforcement learning. *Nature* 518, 7540 (2015), 529–533.

[18] Ryota Namba and Jun Sakuma. 2019. Robust watermarking of neural network with exponential weighting. In *ACM Asia Conference on Computer and Communications Security*. 228–240.

[19] Bita Darvish Rouhani, Huili Chen, and Farinaz Koushanfar. 2019. Deepsigns: An end-to-end watermarking framework for protecting the ownership of deep neural networks. In *ACM International Conference on Architectural Support for Programming Languages and Operating Systems*.

[20] Ahmad EL Sallab, Mohammed Abdou, Etienne Perot, and Senthil Yogamani. 2017. Deep reinforcement learning framework for autonomous driving. *Electronic Imaging* 2017, 19 (2017), 70–76.

[21] Tom Schaul, John Quan, Ioannis Antonoglou, and David Silver. 2015. Prioritized experience replay. *arXiv preprint arXiv:1511.05952* (2015).

[22] Ruimin Shen, Yan Zheng, Jianye Hao, Zhaopeng Meng, Yingfeng Chen, Changjie Fan, and Yang Liu. [n.d.]. Generating Behavior-Diverse Game AIs with Evolutionary Multi-Objective Deep Reinforcement Learning. ([n. d.]).

[23] Karen Simonyan and Andrew Zisserman. 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556* (2014).

[24] Jianwen Sun, Tianwei Zhang, Xiaofei Xie, Lei Ma, Yan Zheng, Kangjie Chen, and Yang Liu. 2020. Stealthy and efficient adversarial attacks against deep reinforcement learning. *arXiv preprint arXiv:2005.07099* (2020).

[25] Richard S Sutton. 1988. Learning to predict by the methods of temporal differences. *Machine learning* 3, 1 (1988), 9–44.

[26] Sebastian Szyller, Buse Gul Atli, Samuel Marchal, and N Asokan. 2019. Dawn: Dynamic adversarial watermarking of neural networks. *arXiv preprint arXiv:1906.00830* (2019).

[27] Yusuke Uchida, Yuki Nagai, Shigeyuki Sakazawa, and Shin'ichi Satoh. 2017. Embedding watermarks into deep neural networks. In *ACM on International Conference on Multimedia Retrieval*. 269–277.

[28] Yue Wang, Esha Sarkar, Michail Maniatakos, and Saif Eddin Jabari. 2020. Stop-and-Go: Exploring Backdoor Attacks on Deep Reinforcement Learning-based Traffic Congestion Control Systems. arXiv:2003.07859

[29] Ronald J Williams. 1992. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine learning* 8, 3-4 (1992), 229–256.

[30] Jialong Zhang, Zhongshu Gu, Jiyong Jang, Hui Wu, Marc Ph Stoecklin, Heqing Huang, and Ian Molloy. 2018. Protecting intellectual property of deep neural networks with watermarking. In *ACM Asia Conference on Computer and Communications Security*. 159–172.

[31] Yan Zheng, Xiaofei Xie, Ting Su, Lei Ma, Jianye Hao, Zhaopeng Meng, Yang Liu, Ruimin Shen, Yingfeng Chen, and Changjie Fan. 2019. Wuji: Automatic online combat game testing using evolutionary deep reinforcement learning. In *2019 34th IEEE/ACM International Conference on Automated Software Engineering (ASE)*. IEEE, 772–784.