# Multimedia Meets Deep Reinforcement Learning

Shu-Ching Chen, *University of Missouri-Kansas City, Kansas City, MO, 64110, USA*

*Multimedia data analysis methods based on artificial intelligence and machine learning have achieved great success in the past decades. However, it remains challenging to generate automated decision-making based on the multimedia data directly. Recent advances in deep reinforcement learning (DRL) have made it a practical framework for solving various sequential decision-making problems. However, it lacks the capability to be effectively used in real-world applications. Multimedia data are useful sources of information to support reliable decision-making. By incorporating DRL with multimedia content analysis, it is promising to develop reliable and effective decision-making systems and frameworks for critical applications, such as disaster management, autonomous driving, healthcare, and so on. Hence, it creates the opportunities for the multimedia community to develop novel techniques to address the existing challenges to further improve the usability and performance of automated decision-making based on multimedia data.*

Multimedia integrates various types of data modalities, including images, videos, audios, texts, etc., to provide a powerful tool to represent and digitalize contents and information. In the past decades, many multimedia researchers have used artificial intelligence and machine learning techniques to analyze and synthesize multimedia data, trying to understand multimedia content as human beings for various applications.[1] However, closing the gaps between multimedia content analyses and decision making, especially for real-world problems, remains challenging. On the other hand, since AlphaGo achieved to beat the professional human player of Go in 2015,[2] deep reinforcement learning (DRL) has become a new paradigm that succeeds in solving sequential decision-making problems. While DRL was initially limited to perfect information board games, such as Go, it has quickly evolved and generalized to imperfect information board games, such as poker,[3] and even video games, such as Atari.[4] DRL has begun to take multimedia data, such as videos, as inputs to guide the decision-making process like humans. However, when multimedia meets DRL, especially for real-world applications, there remain many challenges to be addressed to leverage the rich information contained in the multimedia data and make successful and reliable decisions.

One of the most important challenges is to effectively train the DRL models based on the multimedia data collected from complex environments. To tackle this problem, self-play mechanisms, agent branching methods, and regret minimization have been proposed to enable efficient DRL training without human interventions for games, such as Minecraft and StarCraft.[5–7] By incorporating these DRL techniques, many multimedia applications have recently achieved the state-of-the-art performance, including keyframe selection,[8] low-light image enhancement,[9] video analysis,[10] and so on. However, they rely on large-scale curated datasets or the virtual game environment that can simulate massive numbers of samples.

For critical applications, such as disaster management and healthcare, generating samples could potentially put people's lives at risk. Thus, developing a realistic simulator becomes important to the DRL training for these real-world applications. For example, a healthcare information processing system has been recently proposed to recommend the intervention strategy, where a DNN-based simulator is built to facilitate DRL training based on multimedia data.[11] Similarly, simulators based on the game engine have also been widely used in the autonomous driving

1070-986X © 2022 IEEE
Digital Object Identifier 10.1109/MMUL.2022.3196479
Date of current version 23 September 2022.

industry.[12] However, how to ensure the reliability of the produced decisions remains challenging for the existing systems. Further research in effective multimedia information extraction and reliable decision-making should be conducted.

On the other hand, it is critical to transfer the knowledge and generalize the existing models. While variants of policy distillation and network adaptation methods[13–15] have been proposed to transfer knowledge between DRL models for new environments and problems, these methods are usually specific to the DRL frameworks and types of problem formulation. More general and flexible ways, which can work for various models and multimedia data, are under study.

Meanwhile, in many applications, it is necessary for automated agents to collaborate with other automated and human agents to accomplish the tasks. This requires DRL to work in a multiagent environment and take human factors into consideration based on the multimedia data inputs. DRL models have shown superior performance in cooperating with other automated agents to play video games, such as Multiplayer Online Battle Arena.[16] However, when it comes to human agents, aligning the target and intentions becomes a much more difficult task. Recently, Peschi et al. proposed a multiobjective reinforced active learning framework to help the model understand social norms and human intentions.[17] Lv et al. proposed to utilize user interaction to guide the training of the DRL model and learn the personalized image aesthetic accordingly.[18] These models are all designated for specific applications, and thus more generalizable DRL methods for interacting with humans based on multimedia data inputs should be investigated.

DRL technique is a promising approach to bridge the gaps between analyzing multimedia data and making decisions for real-world applications upon data. Many opportunities are open for the multimedia community to build multimedia frameworks and systems to support important decision making for various real-world problems. However, there remain many existing and emerging challenges and technical problems to be addressed to develop the DRL model effectively and efficiently, which calls for the need for continuous research efforts when multimedia meets DRL.

## REFERENCES


1. S. Pouyanfar et al., "A survey on deep learning: Algorithms, techniques, and applications," *ACM Comput. Surv.*, vol. 51, no. 5, pp. 1–36, 2018, doi: 10.1145/3234150.
2. D. Silver et al., "Mastering the game of Go with deep neural networks and tree search," *Nature*, vol. 529, pp. 484–489, 2016, doi: 10.1038/nature16961.
3. M. Moravčík et al., "Deepstack: Expert-level artificial intelligence in heads-up no-limit poker," *Science*, vol. 356, no. 6337, pp. 508–513, 2017, doi: 10.1126/science.aam6960.
4. Ł. Kaiser et al., "Model based reinforcement learning for atari," in *Proc. Int. Conf. Learn. Representations*, 2020, pp. 1–28.
5. O. Vinyals et al., "Grandmaster level in StarCraft II using multi-agent reinforcement learning," *Nature*, vol. 575, no. 7782, pp. 350–354, 2019, doi: 10.1038/s41586-019-1724-z.
6. X. Wang et al., "SCC: An efficient deep reinforcement learning agent mastering the game of StarCraft II," in *Proc. Int. Conf. Mach. Learn.*, 2021, pp. 10905–10915. [Online]. Available: http://proceedings.mlr.press/v139/wang21v.html
7. P. Jin, K. Keutzer, and S. Levine, "Regret minimization for partially observable deep reinforcement learning," in *Proc. Int. Conf. Mach. Learn.*, 2018, pp. 2342–2351. [Online]. Available: http://proceedings.mlr.press/v80/jin18c.html
8. C. Mo, K. Hu, S. Mei, Z. Chen, and Z. Wang, "Keyframe extraction from motion capture sequences with graph based deep reinforcement learning," in *Proc. 29th ACM Int. Conf. Multimedia*, 2021, pp. 5194–5202, doi: 10.1145/3474085.3475635.
9. R. Zhang, L. Guo, S. Huang, and B. Wen, "ReLLIE: Deep reinforcement learning for customized low-light image enhancement," in *Proc. 29th ACM Int. Conf. Multimedia*, 2021, pp. 2429–2437, doi: 10.1145/3474085.3475410.
10. M. Xu, Y. Song, J. Wang, M. Qiao, L. Huo, and Z. Wang, "Predicting head movement in panoramic video: A deep reinforcement learning approach," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 11, pp. 2693–2708, Nov. 2019, doi: 10.1109/TPAMI.2018.2858783.
11. Y. Dai, G. Wang, K. Muhammad, and S. Liu, "A closed-loop healthcare processing approach based on deep reinforcement learning," *Multimedia Tools Appl.*, vol. 81, pp. 3107–3129, 2022, doi: 10.1007/s11042-020-08896-5.
12. A. Sallab, M. Abdou, E. Perot, and S. Yogamani, "Deep reinforcement learning framework for autonomous driving," *Electron. Imag.*, no. 19, pp. 70–76, 2017.
13. A. A. Rusu et al., "Policy distillation," in *Proc. Int. Conf. Learn. Representations*, 2016, pp. 1–13. [Online]. Available: https://openreview.net/forum?id=9nHQRKrWAMt
14. C. Tessler, S. Givony, T. Zahavy, D. Mankowitz, and S. Mannor, "A deep hierarchical approach to lifelong learning in Minecraft," in *Proc. AAAI Conf. Artif. Intell.*, vol. 31, 2017, Art. no. 1, doi: 10.1609/aaai.v31i1.10744.

15. S. James et al., "Sim-to-real via sim-to-sim: Data-efficient robotic grasping via randomized-to-canonical adaptation networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 12619–12629. [Online]. Available: https://openaccess.thecvf.com/content_CVPR_2019/html/James_Sim-To-Real_via_Sim-To-Sim_Data-Efficient_Robotic_Grasping_via_Randomized-To-Canonical_Adaptation_Networks_CVPR_2019_paper.html
16. D. Ye et al., "Towards playing full MOBA games with deep reinforcement learning," *Adv. Neural Inf. Process. Syst.*, vol. 33, pp. 621–632, 2020. [Online]. Available: https://proceedings.neurips.cc/paper/2020/file/06d5ae105ea1bea4d800bc96491876e9-Paper.pdf
17. M. Peschl, A. Zgonnikov, F. A. Oliehoek, and L. C. Siebert, "MORAL: Aligning AI with human norms through multi-objective reinforced active learning," in *Proc. 21st Int. Conf. Auton. Agents Multiagent Syst.*, 2022, pp. 1038–1046, doi: 10.5555/3535850.3535966.
18. P. Lv et al., "User-Guided personalized image aesthetic assessment based on deep reinforcement learning," *IEEE Trans. Multimedia*, early access, Nov. 25, 2021, doi: 10.1109/TMM.2021.3130752.


**SHU-CHING CHEN** is a professor with the University of Missouri-Kansas City, Kansas City, MO, 64110, USA. Contact him at s.chen@umkc.edu.