1. On the consistency of hyper-parameter selection in value-based deep reinforcement learning, RLC 2024, Raghu: The paper introduces a new score to quantify the consistency and reliability of various hyperparameters and studies the reliability of hyperparameter selection for value-based deep reinforcement learning agents. They show which hyperparameters are most critical and which tunings remain consistent across different training regimes.
2. The Dormant Neuron Phenomenon in Deep Reinforcement Learning, ICML 2023, Raghu: The paper defines dormancy of neurons and studies the phenomenon of dormant neurons increasing during training. They propose a method to *recycle* these neurons and make them useful again.
3. ARLBench: Flexible and Efficient Benchmarking for Hyperparameter Optimization in Reinforcement Learning, arxiv 2024, Raghu: The paper creates a new and efficient benchmark for RL algorithms. They implement agents in JAX and perform subset selection using the technique from the paper *Atari-5* to obtain speed-ups for evaluation. They implement an environment for AutoRL for easy use with various environment frameworks and release a public dataset of multiple training runs.
4. Genie: Generative Interactive Environments, ICML 2024, Raghu: The paper introduces an approach to generate interactive environments from an image or text prompt. They train on internet scale videos in an unsupervised manner (with no labelled data) to obtain a world model with a learnt latent action model which allows them to transfer the learnt action model to unseen prompts while maintaining similar semantics in the environment generated from it.
5. Colored Noise in PPO: Improved Exploration and Performance through Correlated Action Sampling, AAAI 2024, Baohe: This paper evaluates different colored noise and their impact on on-policy algorithms, highlighting the essence of AutoRL for exploration and challenges of applying colored noise to RL.
6. Adaptive Horizon Actor-Critic for Policy Learning in Contact-Rich Differentiable Simulation, ICML 2024, Baohe: This paper proposes a new angle for RL/AutoRL researchers. By ultilizing a differentialable simulator, the authors suggest to adaptively change the rollout horizon of the model-based RL algorithm by looking into the gradient and the physics property (stiff contact). This is different from what AutoRL usually does, such as using PBT/BOHB, that are gradient-free and no prior knowledge assumed.
7. Combining Automated Optimisation of Hyperparameters and Reward Shape, RLC2024, Baohe: This paper shows the dependencies between reward shaping and the hyperparameters. An empirical study has been performed on PPO, showing a combined optimisation works great with little extra computational cost.
8. EMORL: Effective multi-objective reinforcement learning method for hyperparameter optimization, EAAI 2021, Hasan: This paper presents EMORL, that balances both accuracy and latency, addressing real-world concerns like CPU utilization. By framing the optimization problem within a reinforcement learning framework, an agent sequentially selects hyperparameters using a reward function that combines accuracy and latency. The method also improves efficiency by reusing successful configurations.
9. Hyperparameter Optimization for Multi-Objective Reinforcement Learning, arXiv 2023, Hasan : This paper explores the challenge of hyperparameter optimization in multi-objective reinforcement learning (MORL). While MORL expands the scope of reinforcement learning by allowing agents to balance multiple objectives, its success

depends heavily on tuning hyperparameters, a task that can be difficult in practice. The authors formalize this problem and propose a systematic methodology to address it.

10. Population Based Training of Neural Networks, arXiv 2017, André/Noor: Population Based Training (PBT) is one of the most used solution approaches for automating reinforcement learning. PBT maintains an ensemble of RL agents that each maintain their own set of neural network weights and hyperparameters. The hyperparameters are initially sampled randomly and, after fixed intervals, the worst performing agents will copy the weights of the best performing ones. To continue training and exploring the hyperparameters, they are slightly perturbed or even resampled. PBT thus enables dynamic tuning of RL agents.
11. Sample-Efficient Automated Deep Reinforcement Learning, ICLR 2021, André: SEARL combines multiple improvements for off-policy reinforcement learning algorithms and advances PBT to also adapt the neural architectures of the RL agents. Overall, SEARL can provide more efficient, dynamic tuning of RL agents by leveraging shared experiences.
12. Provably Efficient Online Hyperparameter Optimization with Population-Based Bandits, NeurIPS 2020, André: PB2 extends PBT by introducing a time-varying Gaussian Process to sample new hyperparameter configurations from. This combination of PBT with BO enables smarter exploration of the configuration space and reduces the compute demand of PBT.
13. On the Importance of Hyperparameter Optimization for Model-based Reinforcement Learning, AISTATS 2021, André: This work studies hyperparameter optimization for model-based RL both using classical methods from AutoML but also using dynamic ones. This study uncovers advantages and disadvantages of either approach and proposes when to use which. The work further demonstrates that some hyperparameters are best tuned dynamically in MBRL.
14. Hyperparameters in Reinforcement Learning and How To Tune Them, ICML 2023, André/Noor: This work proposes one of the most comprehensive studies of hyperparameter tuning of deep RL agents, to date. Prior works typically compare only a small set of tuning approaches, on a limited set of agent-environment pairs. The work shows that multi-fidelity HPO is an underutilised avenue for deep RL and shows that one of the most popular approaches to dynamic tuning (PBT) often underperforms.
15. AutoRL Hyperparameter Landscapes, AutoML 2023, André: While the (Auto)RL community broadly understood that many hyperparameters are best tuned dynamically, it is often not well understood how often and how strong to adapt them. Further, it was not well understood how the interplay of such hyperparameters looks like. This work proposes a method to carefully analyze the hyperparameter landscapes to try and answer these open questions.
16. Generalized Population-Based Training for Hyperparameter Optimization in Reinforcement Learning, arXiv 2024, Noor: this work propose a generalized version of PBT that incorporates adaptive strategies to maintain diversity within the population, leading to more efficient exploration of hyperparameter spaces. The framework allows for dynamic adjustments of hyperparameters based on the performance of agents, thus improving sample efficiency and convergence in complex RL environments. The empirical results demonstrate that the proposed

approach significantly outperforms standard HPO methods across various RL benchmarks.