Multi-Objective Hyperparameter Optimization with Multi-Armed Bandits

Fangyu Luo [1] Fatemeh Nargesian [1]

[1]Department of Computer Science, University of Rochester

HAJIM SCHOOL OF ENGINEERING & APPLIED SCIENCES
UNIVERSITY of ROCHESTER

Abstract

Our study dives into the application of Multi-Objective Multi-Armed Bandit (**MO-MAB**) algorithms in hyperparameter optimization (**HPO**) challenges. We identify areas for improvement within MAB-based algorithms in HPO, such as Hyperband, and demonstrate that incorporating fairness metrics can enhance model performance.

Introduction

- **Multi-Objective Multi-Armed Bandits (MO-MABs) [2]:** In a standard MAB, the algorithm selects an action (arm) at each round from all options without prior knowledge of the outcomes and receives a stochastic reward. The goal of the agent is often to maximize the total reward.

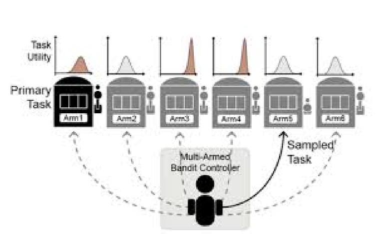


Figure 1. Multi-armed bandit [6]

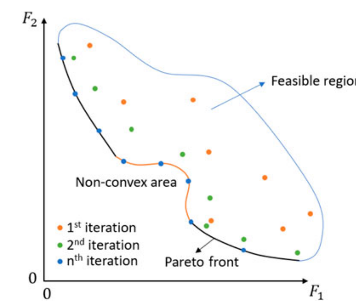


Figure 2. Pareto front [4]

In **MO-MABs**, rewards become multi-dimensional. Actions could be equally important, excelling in different objectives. The solutions that are optimal across multiple objectives where no single objective can be improved without compromising another are known as the **Pareto front**. This approach uncovers actions that single-objective analysis might miss (where rewards from each objective are combined through scalarization).

Figure 3. Hyper-parameters of a chocolate cookie [5]

- **Hyperparameter Optimization (HPO):** In Machine Learning, HPO plays a critical role in refining the parameters that shape the learning process. Unlike model parameters that are derived from data, hyperparameters are external configurations whose calibration profoundly influences model performance (e.g. learning rate, batch size, etc.).
- **Hyperband [3]:** Hyperband is a HPO algorithm based on a MAB variant, the infinitely-armed bandit.
  - Each 'arm' represents a different set of hyperparameters
  - The goal is to minimize validation loss—the 'reward' in this context
  - Samples a manageable number from the infinite configurations
  - Balances exploration of the search space with the exploitation of high-performing configurations by dynamically allocating resources to the most promising configurations at intervals and getting rid of the others

Method

We use the CIFAR10 dataset and a model that can reach 94% accuracy in just 140s on a V100 GPU [1] in our experiments. While the original model employs a dynamic learning rate schedule, starting at 0, peaking at 0.6 by the 8th epoch, and tapering back to 0 by the 30th, with a batch size of 768, momentum of 0.9, and weight decay calibrated to the batch size, we expand on the above hyperparameter as detailed in Table 1.

1. **Hyperband:**
   - We switch Hyperband's configuration ranking strategy from accuracy to fairness to examine if it can uncover better configurations than the standard approach focused on accuracy.
   - Hyperband's resource allocation and bracket structure are detailed in Table 2.
2. **Fully Trained Configurations:**
   - We altered the CIFAR10 to be imbalanced by reducing several classes to just 100 images. 50 configurations are thoroughly trained across 10 trials and 3 imbalanced scenarios.

Type	Range
learning rate	UniformFloat, 5×10^{-1} to 8×10^{-1}
batch size	UniformInteger, 50 to 1000 (log)
weight decay ratio	UniformFloat, 1×10^{-10} to 1×10^{-2} (log)
momentum	UniformFloat, 1×10^{-9} to 1×10^{-1}

Table 1. The upper and lower bound of the Hyperparameter search space

	s = 3		s = 2		s = 1		s = 0	
i	n_i	r_i	n_i	r_i	n_i	r_i	n_i	r_i
0	27	1	12	3	6	10	4	30
1	9	3	4	9	2	20		
2	3	9	1	18				
3	1	17						

Table 2. The values of configs (n) vs. resources (r)

Results of Hyperband

- **Benchmarking Fairness vs. Accuracy:** Fairness-oriented Hyperband identified equivalent configurations in 75% of trials compared to accuracy-focused methods, while the remainder difference seemed to be influenced by initialization variability as shown in Figure 4.
- **Dynamic Learning Rate Impacting the Best bracket:** Contrary to typical Hyperband, the most conservative bracket $(s = 0)$ proved most effective, uncovering 95% of optimal configurations. This suggests that early performance boosts from learning rate changes can mislead the algorithm, emphasizing the need for a balance between early achievements and consistent long-term performance. This finding opens new avenues for investigating how learning rate schedules can enhance future MAB-based HPO strategies.

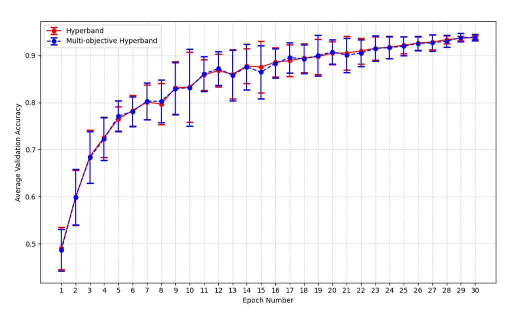


Figure 4. Average Validation Accuracy of the Best Configuration vs. Epoch Number

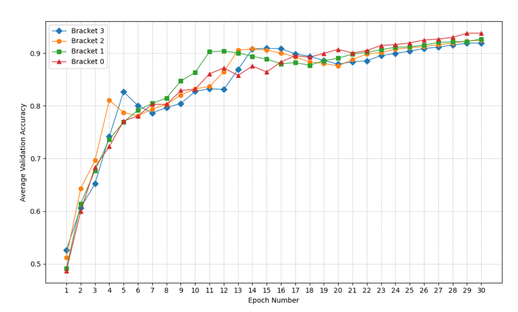
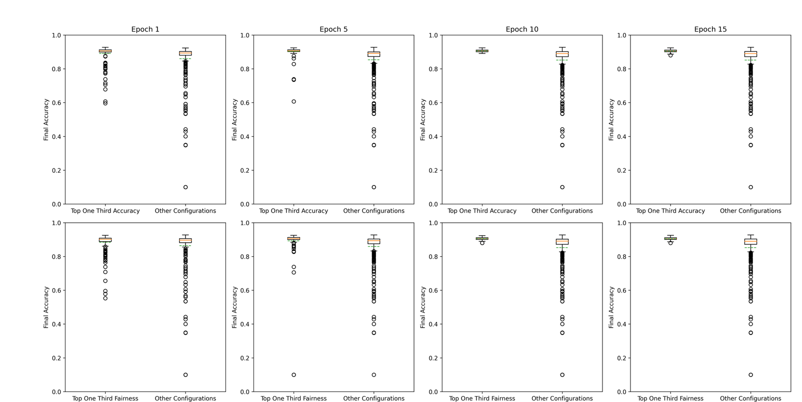


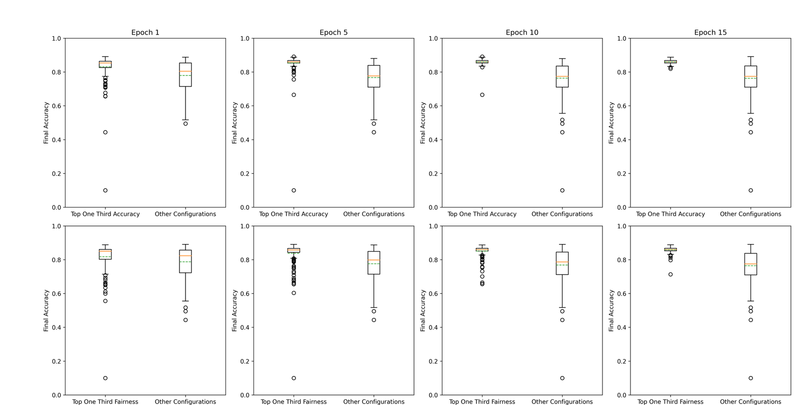
Figure 5. Average Validation Accuracy of the Best Configuration of each bracket vs. Epoch Number

Results of Fully Trained Configurations

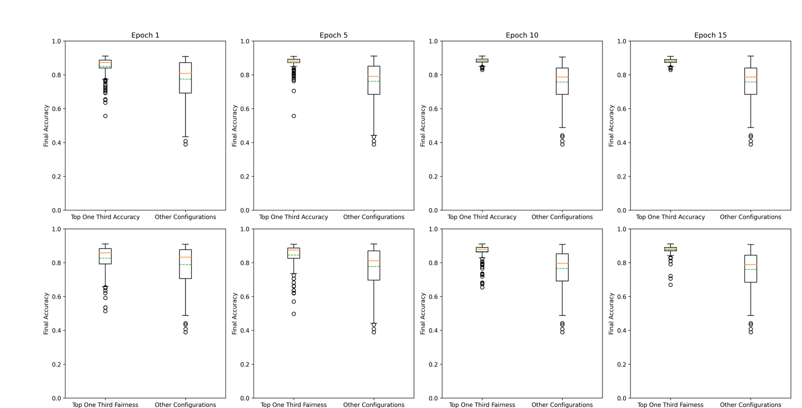
- **Stability in Accuracy vs. Fairness (Figure 6):** The reliability of accuracy metrics stays steady across various levels of class imbalance. In contrast, the effectiveness of fairness metrics decreases as more classes are altered.
- **Importance of Timing in Selection (Figure 6):** Configurations begin to show stable rankings around epoch 10. Early pruning often mistakenly cuts off configurations that could have been top performers if given more time to develop.
- **Potential in Top Configurations (Table 3):** In certain instances, focusing on fairness increases the likelihood of identifying the best configurations, suggesting that fairness metrics can sometimes outperform accuracy in selecting optimal models.



(a) One-class-reduced



(b) Five-class-reduced



(c) Nine-class-reduced

Figure 6. Comparison of final accuracy based on early results across different class imbalances.

	One class reduced		Five class reduced		Nine class reduced	
Epoch	Rank in Accuracy	Rank in Fairness	Rank in Accuracy	Rank in Fairness	Rank in Accuracy	Rank in Fairness
1	13.33%	3.33%	3.33%	**10.00%**	13.33%	10.00%
5	3.33%	**16.67%**	26.67%	13.33%	23.33%	13.33%
10	6.67%	**10.00%**	36.67%	30.00%	10.00%	3.33%
15	10.00%	10.00%	33.33%	20.00%	16.67%	6.67%

Table 3. Percentage of top 3 Configurations After Full Training Identified in the top 3 at Specified Epochs Across Varied Class Reduction Scenarios

Conclusion

Our studies indicate that though accuracy remains stable across different levels of imbalance, emphasizing fairness over accuracy in model training can still reveal superior configurations that might be missed by focusing solely on accuracy.

A critical observation from our use of the Hyperband algorithm reveals that its early pruning strategy, particularly when combined with dynamic learning rate schedules, risks discarding potentially high-performing configurations too soon. This observation calls for a thoughtful reassessment of both resource allocation and the timing of pruning decisions, to better predict and foster long-term model efficacy.

For future studies, we propose a recalibration of HPO strategies to strike a more effective balance between fairness and accuracy. Key initiatives will include optimizing resource allocation and fine-tuning pruning strategies to address the complex dynamics among fairness, accuracy, and operational efficiency within MAB-inspired HPO frameworks. Additionally, we aim to delve deeper into understanding how various hyperparameters interact and to refine the mechanism for selecting the most promising arms, advancing the efficacy of MO-HPO techniques with MABs.

Acknowledgements

This research was made possible through the generous support of the Schwartz Discover Grant. Special gratitude is extended to Professor Adam Purtee for their invaluable advice and guidance throughout this research.

References

[1] Davidcpage. cifar10-fast. `https://github.com/davidcpage/cifar10-fast`, Year.

[2] M. M. Drugan and A. Nowe. Designing multi-objective multi-armed bandits algorithms: A study. In *IJCNN*, pages 1–8, 2013.

[3] L. Li et al. Hyperband: A novel bandit-based approach to hyperparameter optimization. *Journal of Machine Learning Research*, 18(185):1–52, 2018.

[4] V. De Buck et al. Exploiting trade-off criteria to improve the efficiency of genetic multi-objective optimisation algorithms. *Frontiers in Engineering*, 2021.

[5] Google for Developers. Vizier: Black-box optimization and automl - pittsburgh ml summit '19. `https://www.youtube.com/watch?v=6aSG8SdvkoU`, 2019.

[6] S. N. Srihari. Multi-armed bandits. `https://cedar.buffalo.edu/~srihari/CSE574/Chap15/15.7-Multi-armed%20Bandits.pdf`, 2023.