

Auto-tuning of Hyperparameters by Parallel Search Using Xcrypt

Tatsuro Hanyu† Masatoshi Kawai‡ Takahiro Katagiri‡ Tasuku Hiraishi†3 Tetsuya Hoshino‡ Toru Nagai‡
Graduate School of Informatics, Nagoya University† Information Technology Center, Nagoya University‡ Department of
Information and Computer Science, Faculty of Engineering, Kyoto Tachibana University†3

Introduction

Optimizing hyperparameters is essential for ensuring quality and performance in AI models, particularly in CNNs. Parallel processing, both at program and job levels, maximizes computing resource utilization. Xcrypt[1][2], a Perl-based scripting language, facilitates job-level parallel programming across various supercomputing environments. Our research focuses on developing an automated hyperparameter search system using Xcrypt, emphasizing auto-tuning functions[3].

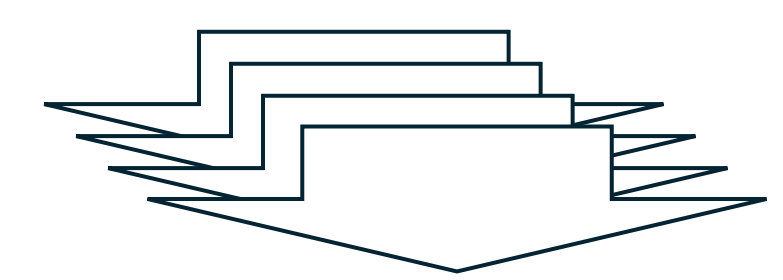
Xcrypt

Xcrypt is a Perl-based scripting language designed for managing multiple jobs on supercomputers, either sequentially or in parallel. Its unique feature is the ability to execute a single script across various environments, supporting custom extensions like limiting simultaneous job submissions.

Main Usage Scenarios

- Repeated execution of a single program.
- Multiple input variations to try.
- Performance evaluation while varying problem sizes and the number of parallel executions.

Xcrypt



job

```
#!/bin/sh
#PJM -L elapse=0:30:00
#PJM -L
jobenv=singularity
#PJM -L rscgrp=cx-share
...
resnet.py
--batch_size=700
--num_iter=300
...
```

Result

Auto-Tuning (AT) for batch size optimization on ResNet50 using Xcrypt

- Exploration covered batch sizes ranging from 1 to 700.
- Conducted a comprehensive search involving a total of 700 configurations.

Table1: Top Five Batch Size Optimization Results by Smallest Loss

	Batch size	Loss	Execution time [s.]
1 st	210	8.4305	71.3
2 nd	223	8.4332	74.8
3 rd	202	8.4385	68.1
4 th	192	8.4415	80.5
5 th	231	8.4458	89.6
default	256	8.5044	95

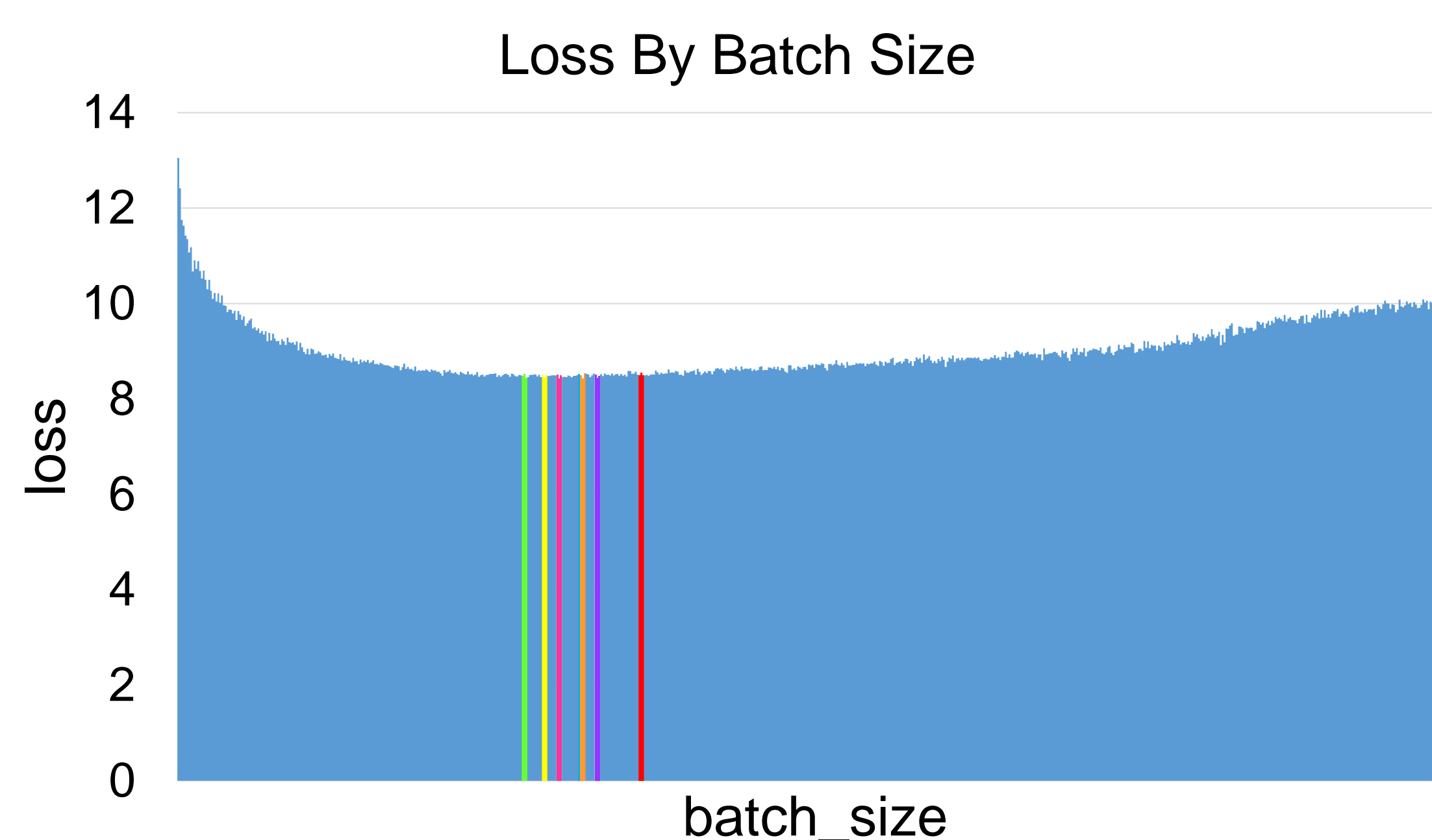


Figure1: Loss Magnitude vs. Batch Size in ResNet50 Optimization

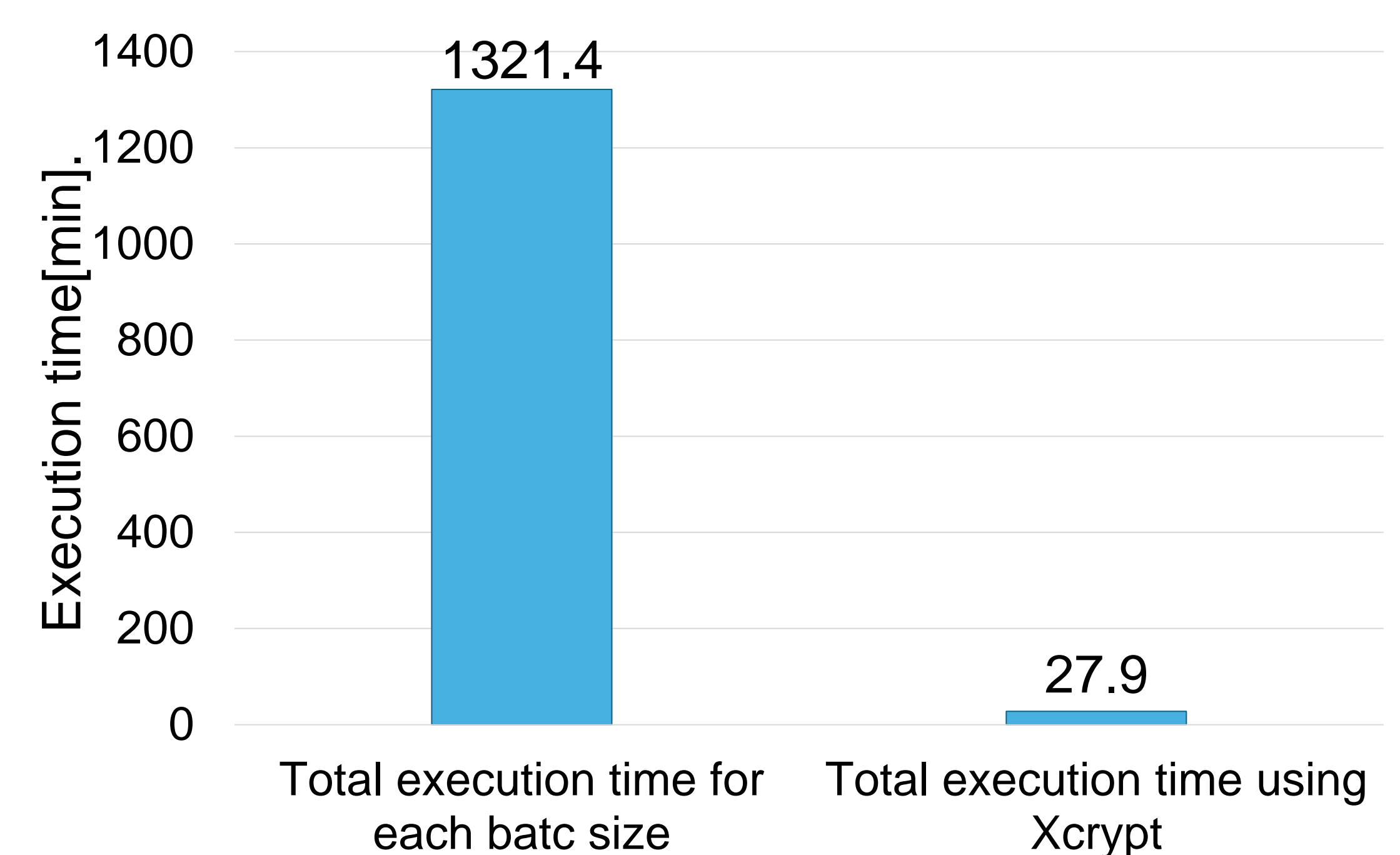


Figure2: Total execution time [min]. for each batch size vs. Total execution time [min]. using Xcrypt

Conclusion

In this study, Xcrypt has been integrated for hyperparameter optimization specifically applied to ResNet50. In future work, we plan to delve into the exploration of a sophisticated search methodology using Xcrypt.

Future Work

- Combining Xcrypt with libraries for hyperparameter optimization, such as Optuna.

Acknowledgments

This work was supported by JSPS KAKENHI Grant Number JP19H05662.

References

- [1] Tasuku Hiraishi, Tatsuya Abe, Takeshi Iwashita, Hiroshi Nakashima. Xcrypt: A Perl Extension for Job Level Parallel Programming. Second International Workshop on High-performance Infrastructure for Scalable Tools WHIST, 2012
- [2] Masaru Ueno, Tasuku Hiraishi, Motoharu Hibino, Takeshi Iwashita, Hiroshi Nakashima. Multilingualization Based on RPC for Job-Level Parallel Script Language, Xcrypt. IPSJ Transaction on Programming, Vol. 6, No. 2, pp. 55-68, 2013
- [3] Takahiro Katagiri, Daisuke Takahashi, Japanese Autotuning Research: Autotuning Languages and FFT, Proc. of the IEEE, Vol. 106, Issue 11, pp. 2056 – 2067, 2018