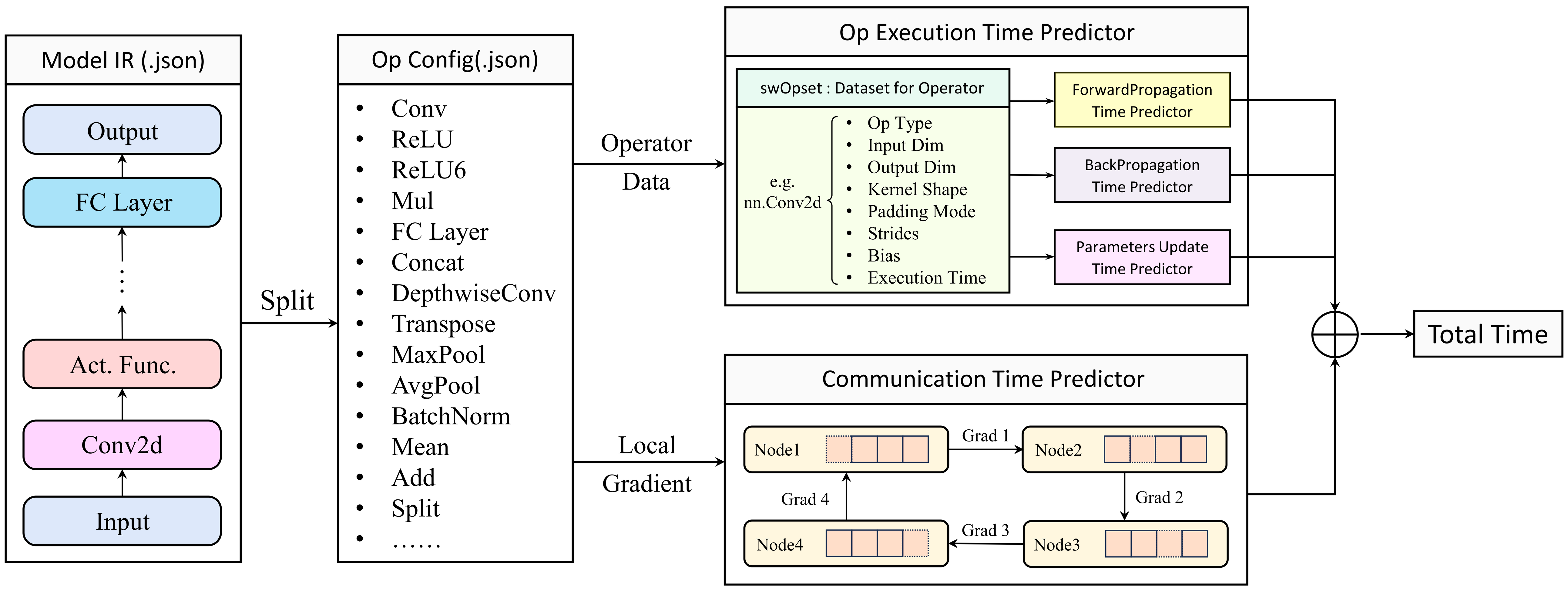
swPredictor: A data-driven performance model for distributed data parallelism training on large-scale HPC clusters

Xianyu Zhu, Ruohan Wu, Junshi Chen, Hong An
Mail: zhuxy@mail.ustc.edu.cn Department: School of Computer Science and Technology, University of Science and Technology of China
Journal: Performance Evaluation: An International Journal (PEVA)

1. Abstract

Given the complexity of heterogeneous architectures and multi-node collaboration, large-scale HPC (High-Performance Computing) clusters pose challenges in resource utilization and performance optimization during distributed data parallelism (DDP) training. Performance modeling aims to identify application bottlenecks and guide algorithm design, but existing performance models rarely consider the impact of system architecture on communication performance or provide a systematic analysis of distributed training. To address these issues, this paper proposes swPredictor, a data-driven performance model devised for accurately predicting the performance of DDP training. First, an original performance dataset is developed based on various communication patterns at runtime to avoid systematic errors. Subsequently, a novel multi-branch module FNO-Inception is proposed, combining FNO (Fourier Neural Operator) layer with Inception structure to simultaneously utilize various frequency features. Finally, by introducing the FNO-Inception module, a novel regression model FI-Net is constructed to fit complex nonlinear relationships. The experimental results demonstrate that FI-Net can accurately predict the performance of DDP training on the Sunway OceanLight supercomputer with an overall MAPE of 0.93%, which outperforms the other baseline models.



2. Contributions

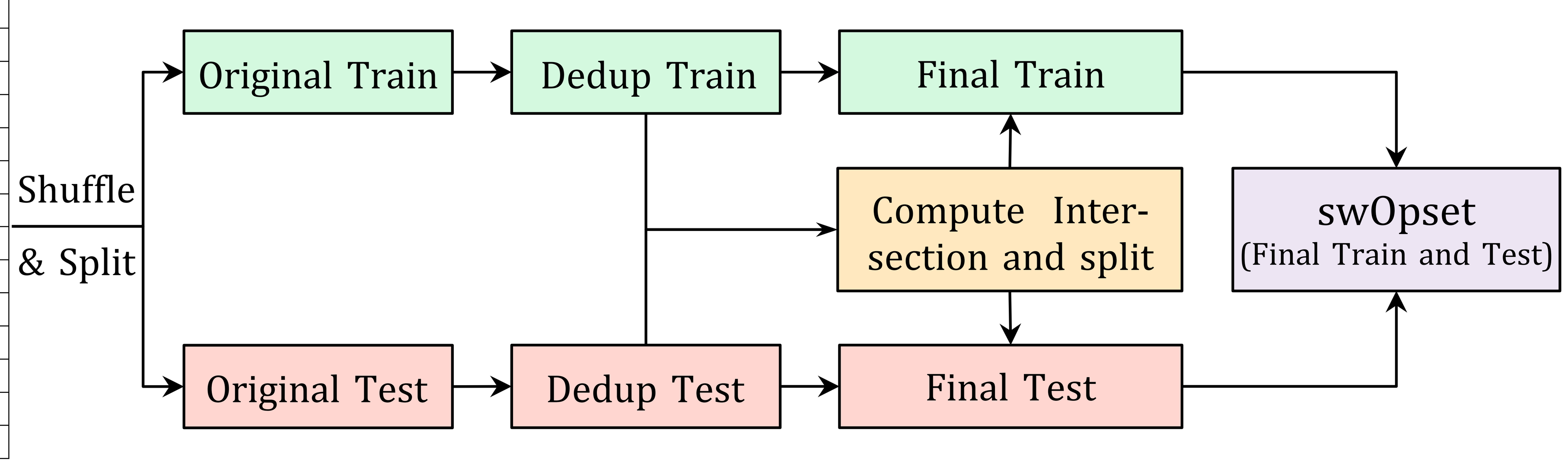
1. Construct an original performance dataset, eliminating the systematic predicted error and adapting to different network topologies on large-scale HPC clusters.
2. Propose a novel multi-branch module namely FNO-Inception to avoid experimental selection of optimal FNO mode and capture various frequency features.
3. Build a new regression model FI-Net, fitting the complex nonlinear relationships between algorithm parameters and perfor-mance within the heterogeneous system.

3. Performance dataset construction

This study adopts a convenient dataset partitioning method to help mitigate data imbalance problems. The specific partitioning process is illustrated in Figure 4. Initially, the original dataset is randomly divided into $Original\_Train$ and $Original\_Test$ in a $4:1$ ratio. Subsequently, both $Original\_Train$ and $Original\_Test$ are deduplicated to form $Dedup\_Train$ and $Dedup\_Test$. This deduplication relies on an operator's key parameters, such as input/output size, kernel size, stride, and bias. Since operator configurations may be identical across different types of networks, the intersection of the training and test sets, denoted as I, is calculated to ensure that the training set and test set do not overlap. Subsequently, the intersection set I is randomly split in a ratio of $4:1$ and assigned to the training and test sets. Thus the $Final\_Train$ is computed as $Final\_Train = (Dedup\_Train \backslash I) \cup I_{0.8}$. In the same way, the $Final\_Test$ is calculated as $Final\_Test = (Dedup\_Test \backslash I) \cup I_{0.2}$. Finally, the swOpset is composed of $Final\_Train$ and $Final\_Test$.

DNN's Operator Configuration Table from nn-Meter Dataset

Model	Conv	ReLU	ReLU6	Mul	FC Layer	MaxPool	AvgPool	BatchNorm	Mean	Add	Concat	DepthConv	Transpose	Split
AlexNet	√	√			√	√			√					
VGG	√	√			√	√		√	√					
GoogLeNet	√	√			√	√		√	√		√			
ResNet	√	√			√	√		√	√	√				
MobileNet-v1	√	√			√			√	√			√		
MobileNet-v2	√	√	√		√			√	√	√		√		
MobileNet-v3	√	√	√	√	√		√	√	√	√		√		
SqueezeNet	√	√			√	√			√		√			
MnasNet	√	√	√	√	√		√	√	√	√		√		
NASBench	√	√			√		√	√	√	√				
ProxyLessNas	√	√	√		√			√	√	√		√		
DenseNet	√	√			√	√	√	√	√		√			
ShuffleNet-v2	√	√			√	√		√	√		√	√	√	√



4. Conclusion

This paper presents a novel data-driven performance model swPredictor designed for predicting the performance of DDP training on large-scale HPC clusters. To fit the complex nonlinear relationships between algorithm parameters and performance within the heterogeneous environment, we construct a new regression prediction model namely FI-Net, which includes a novel multi-branch structure FNO-Inception. By introducing the Inception module to the FNO layer, the FNO-Inception can utilize various frequency features and improve the models accuracy. Meanwhile, it avoids additional experiments to select the optimal FNO mode. Furthermore, in order to make the model adapt to different network topologies and eliminate systematic errors, we develop an original performance dataset by considering the impact of runtime system topology on algorithm performance. The experimental results demonstrate that the proposed model FI-Net can accurately predict the time of DDP training on Sunway OceanLight supercomputer with an overall MAPE of 0.93%.

Building on swPredictors solid foundation, future work will extend our methodology. We aim to generalize the model to encompass other distributed parallelism strategies and communication algorithms, and to validate its predictive accuracy across diverse large-scale system architectures.