

Sanity-Checking Pruning Methods: Random Tickets can Win the Jackpot

Jingtong Su^{*1}, Yihang Chen^{*1}, Tianle Cai^{*2 3},
Tianhao Wu¹, Ruiqi Gao^{2 3}, Liwei Wang¹ and Jason D. Lee²

¹Peking University

²Princeton University

³Zhongguancun Haihua Institute for Frontier Information Technology

*Equal Contribution, Reverse Alphabetical Order.

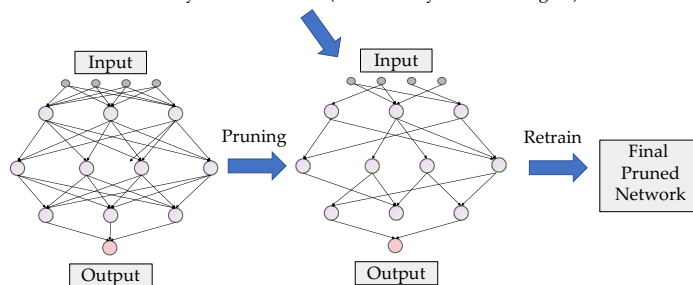
Background Introduction

Network pruning has been proposed as an effective technique to reduce the resource requirements with minimal performance degradation.

Pruning Pipeline

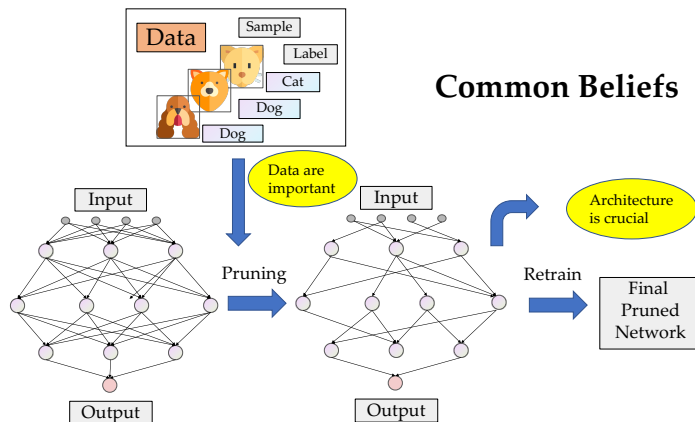
Initial Tickets (Back-To-Initialization)

Partially-Trained Tickets (Use Partially-Trained Weights)



Common Beliefs

- The architecture of the subnetwork is crucial.
- The data used for pruning are important.



Sanity Checks

Use sanity-checks to unravel the information encoded by the pruning method. Two check methods: Corrupted data & Layerwise rearrange.

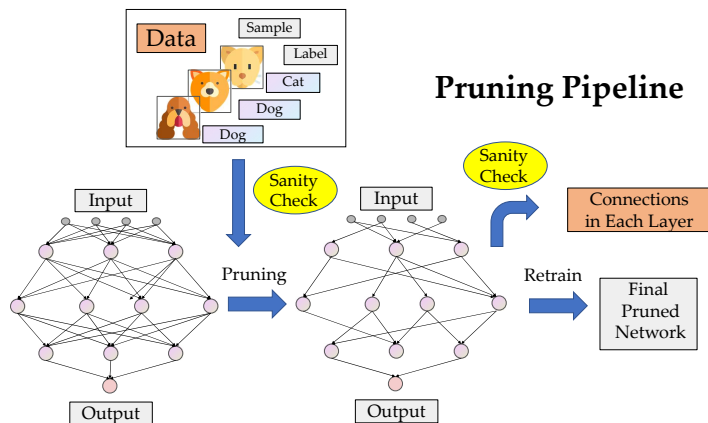
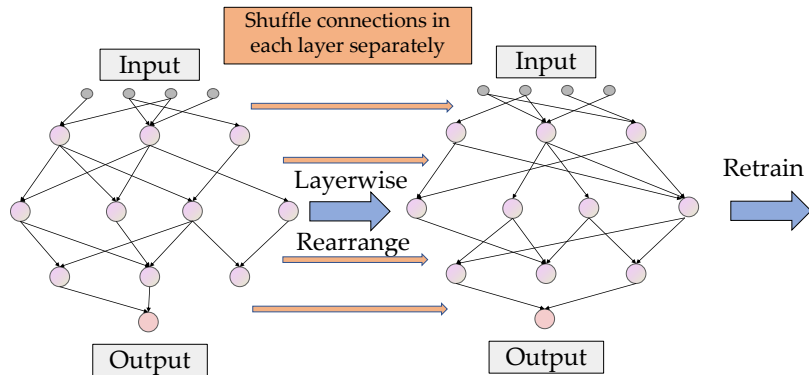
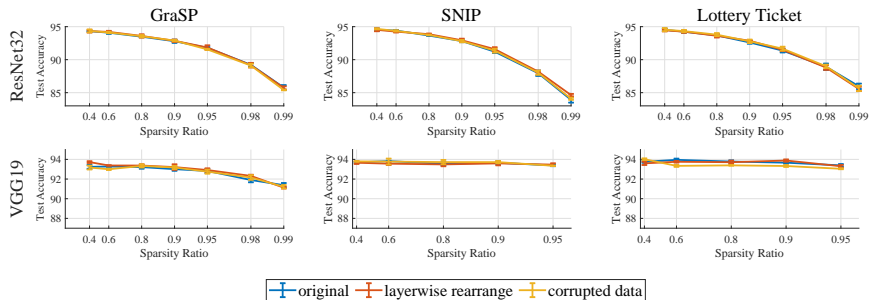


Illustration of Layerwise Rearrange



Results: Sanity Checks on Initial Tickets



Results: Sanity Checks on Partially-Trained Tickets

Table: Sanity-check on partially-trained tickets on CIFAR-10 dataset.

Network	VGG19			ResNet32		
	90%	95%	98%	90%	95%	98%
Pruning ratio						
VGG19/ResNet32 (Full Network)	93.70	-	-	94.62	-	-
Lottery Tickets	93.66±0.08	93.39±0.13	10.00±0.00	92.61±0.19	91.37±0.28	88.92±0.49
Shuffle Weights	93.54±0.04	93.33±0.10	10.00±0.00	92.38±0.36	91.29±0.28	88.52±0.47
Half Dataset	93.79±0.14	93.53±0.13	10.00±0.00	93.01±0.18	92.03±0.21	89.95±0.08
Learning Rate Rewinding	94.14±0.17	93.99±0.15	10.00±0.00	94.14±0.10	93.02±0.28	90.83±0.22

Results: Random Tickets

Table: Test accuracy of pruned VGG19 and ResNet32 on CIFAR-10 and CIFAR-100 datasets. In the full paper, the bold number indicates the average accuracy is within the best confidence interval.

Dataset	CIFAR-10			CIFAR-100		
	90%	95%	98%	90%	95%	98%
VGG19 (Full Network)	93.70	-	-	72.60	-	-
Lottery Tickets	93.66±0.08	93.39±0.13	10.00±0.00	72.58±0.27	70.47±0.19	1.00±0.00
SNIP	93.65±0.14	93.39±0.08	81.37±18.17	72.83±0.16	71.81±0.11	10.83±6.74
GraSP	93.01±0.16	92.82±0.20	91.90±0.23	71.07±0.31	70.14±0.21	68.34±0.20
Random Tickets (Ours)	93.77±0.10	93.42±0.22	92.45±0.22	72.55±0.14	71.37±0.09	68.98±0.34
ResNet32 (Full Network)	94.62	-	-	74.57	-	-
Lottery Tickets	92.61±0.19	91.37±0.28	88.92±0.49	69.63±0.26	66.48±0.13	60.22±0.60
SNIP	92.81±0.17	91.20±0.19	87.94±0.40	69.97±0.17	64.81±0.44	47.97±0.82
GraSP	92.79±0.24	91.80±0.11	89.21±0.26	70.12±0.15	67.05±0.39	59.25±0.33
Random Tickets (Ours)	92.97±0.05	91.60±0.26	89.10±0.33	69.70±0.48	66.82±0.12	60.11±0.16

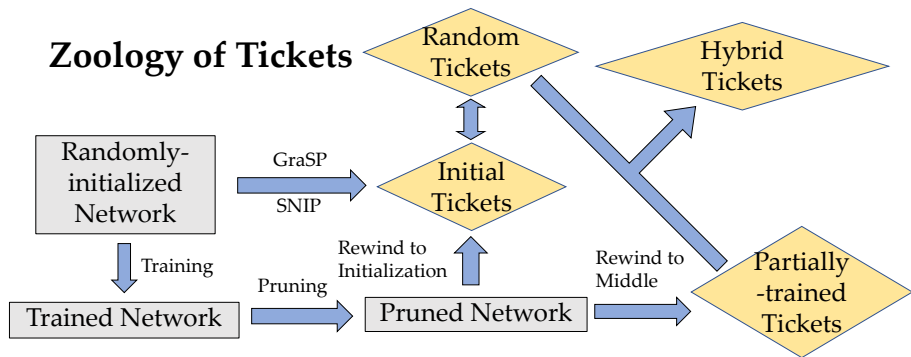
Results: Hybrid Tickets

Table: Test accuracy of partially-trained tickets and our hybrid tickets of VGG19 and ResNet32 on CIFAR-10 and CIFAR-100 datasets.

Dataset	CIFAR-10			CIFAR-100		
	90%	95%	98%	90%	95%	98%
Pruning ratio	90%	95%	98%	90%	95%	98%
VGG19 (Full Network)	93.70	-	-	72.60	-	-
OBD	93.74	93.58	93.49	73.83	71.98	67.79
Learning Rate Rewinding	94.14±0.17	93.99±0.15	<i>10.00±0.00</i>	73.73±0.18	72.39±0.40	<i>1.00±0.00</i>
Hybrid Tickets (Ours)	94.00±0.12	93.83±0.10	93.52±0.28	73.53±0.20	73.10±0.11	71.61±0.46
ResNet32 (Full Network)	94.62	-	-	74.57	-	-
OBD	94.17	93.29	90.31	71.96	68.73	60.65
Learning Rate Rewinding	94.14±0.10	93.02±0.28	90.83±0.22	72.41±0.49	67.22±3.42	59.22±1.15
Hybrid Tickets (Ours)	93.98±0.15	92.96±0.13	90.85±0.06	71.47±0.26	69.28±0.40	63.44±0.34

Overview of our Terminology

Zoology of Tickets



Thanks For Watching!

Contact

Jingtong Su (jtsu@pku.edu.cn), Yihang Chen (yihang.chen@pku.edu.cn) and Tianhao Wu (1700010607@pku.edu.cn) are applying for Ph.D. this year!

Please contact if you are interested!