Neural Lineage

Runpeng Yu, Xinchao Wang

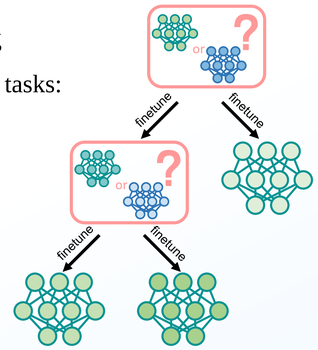
National University of Singapore

NUS National University of Singapore

CVPR SEATTLE, WA JUNE 17-21, 2024

Takeaways

- It is possible to detect with high accuracy from which pre-trained model a neural network was fine-tuned.
- Possible for: an entire neural network or a sub-network.
- We refer to the fine-tuning relationship between neural networks as **Neural Lineage** and term the task of identifying the parent model as the **Neural Lineage Detection**.
- Neural lineage detection can be achieved through learning-free methods conveniently and efficiently.
- Training a lineage detector network further improves the accuracy of neural lineage detection.
- Core intuition: the similarity between a true parent-child model pair should be higher.
- Supports cross-generational lineage detection.
- Supports various learning setups:
  - Common Supervised Learning (with regularization)
  - Few-Shot Learning
  - Imbalanced Learning
- Supports various vision tasks:
  - Classification
  - Segmentation
  - Detection



⇨ Fig. 1 Illustration of Neural Lineage Detection

Learning-Free Detection

- Approximate the finetuning process and then compare the similarity.
- Approximation is based on neural network linearization.
- Efficiency issue is addressed by Taylor Expansion.
- Supports commonly used norm similarity, like l_1 and l_2, and advanced representation similarity, like CKA and DC.

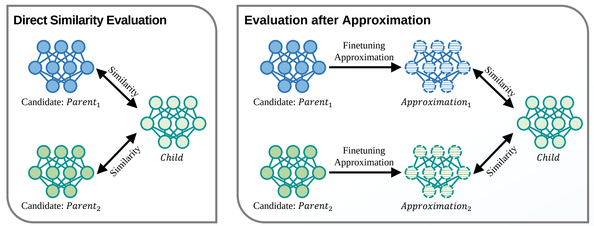


Fig. 2 Workflow: the learning-free similarity-based method.

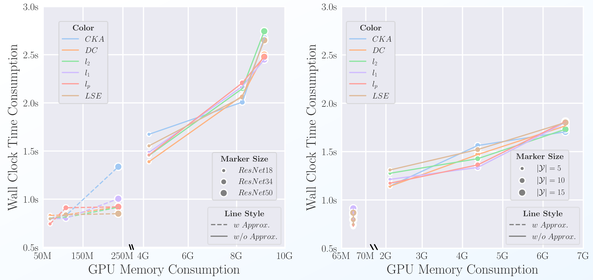


Fig. 3 Comparison of the execution time and memory consumption for methods with & without approximation. Left: different model size; Right: different output size.

Lineage Detector

- Comprise 2 CNN encoders and 1 transformer detector.
- Train with cross-entropy loss.
- Input with network weight and(or) network feature.

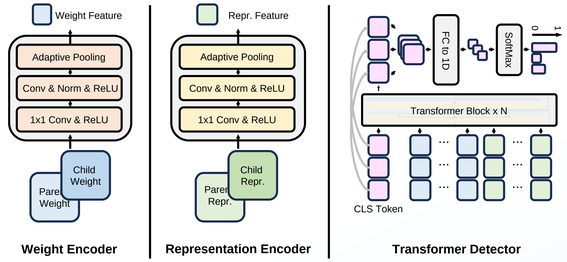


Fig. 4 Workflow: learning-based Lineage Detector.

Lineage Across Generations

- G1: ImageNet pre-trained ResNet18.
- Finetune through three generations, G2, G3, and G4.
- Neural lineage detection between any two generations.
- Larger generational gap, lower accuracy.

		G2	G3	G4
G1	l_1 w/o Appx.	90.47±0.84	89.07±1.72	86.31±2.47
	CKA w/o Appx.	75.59±2.19	44.44±2.89	38.09±1.15
	l_1 w Appx.	98.92±0.59	97.02±0.51	96.83±0.82
	CKA w Appx.	77.38±1.93	51.59±2.98	51.19±1.27
	Lineage Detector	99.11±0.36	98.81±0.13	97.75±1.53
G2	l_1 w/o Appx.	-	91.27±1.89	83.93±2.67
	CKA w/o Appx.	-	65.88±2.63	67.86±2.03
	l_1 w Appx.	-	98.41±0.53	95.24±1.71
	CKA w Appx.	-	75.41±2.81	72.03±3.17
	Lineage Detector	-	99.61±0.83	98.38±0.02
G3	l_1 w/o Appx.	-	-	94.05±1.49
	CKA w/o Appx.	-	-	66.67±1.42
	l_1 w Appx.	-	-	97.02±0.75
	CKA w Appx.	-	-	72.03±1.91
	Lineage Detector	-	-	98.41±0.98

Tab. 1 Cross-generational lineage detection performance.

Lineage Among 300 Models

- Lineage detector accurately identifies: all parent-child pairs and all cross-generational grandparent-child pairs and great-grandparent-child pairs.

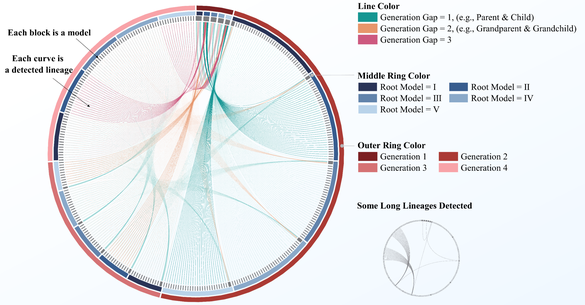


Fig. 5 Detected lineage among 300 deep models.

Lineage for Sub-Network

- A hybrid ViT-B model: concatenate transformer layers from 9 ViT-B models in timm and then finetune.
- Lineage detector accurately identifies: the origins of all transformer layers.

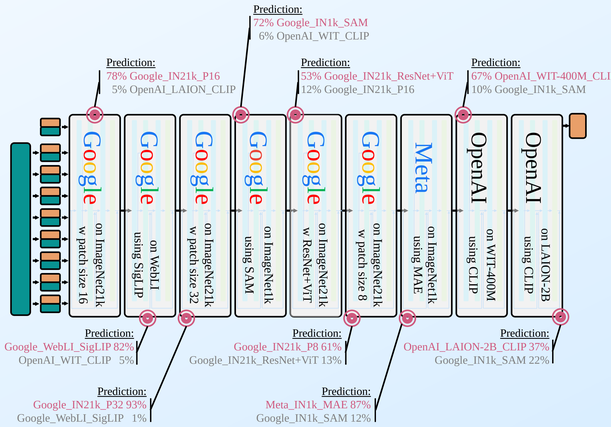


Fig. 6 Detected origin of transformer layers.