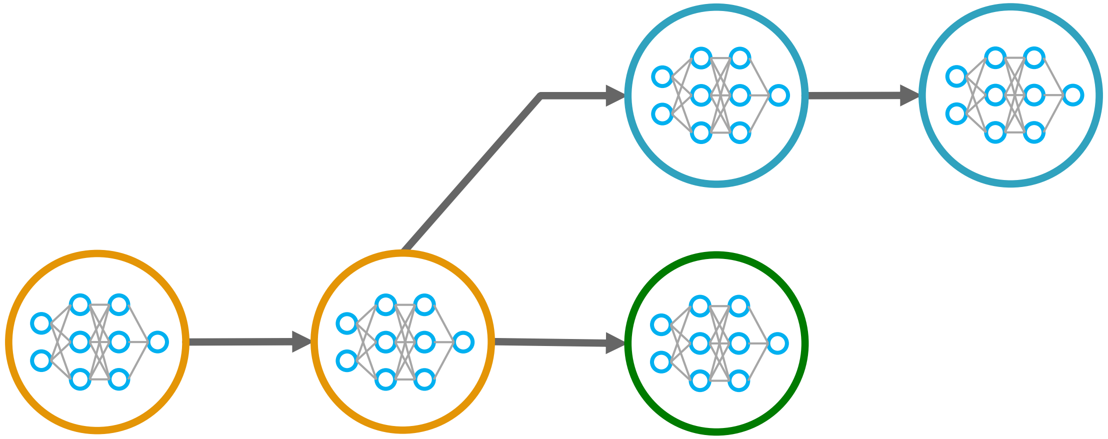


Model Zoos: A Dataset of Diverse Populations of Neural Network Models

Konstantin Schürholt[1], Diyar Taskiran[1], Boris Knyazev[2],
Xavier Giró-i-Nieto[3], Damian Borth[1]

[1] AI:ML Lab, School of Computer Science, University of St. Gallen
[2] Samsung - SAIT AI Lab, Montreal
[3] Institut de Robòtica i Informàtica Industrial, Universitat Politècnica de Catalunya

Introduction

Neural Networks are successfully applied on multiple domains

Loss surface and optimization problem of Neural Networks are highly non-convex

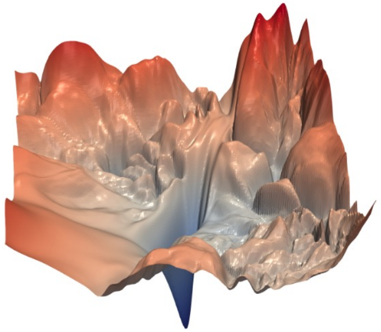
[Dauphin et al.,2014; Goodfellow et al., 2015; LeCun et al., 2015]

Neural Network training optimization is high dimensional

[Brown et al., 2020; Larsen et al., 2021]

Neural Network training is sensitive to hyperparameters and random initialization

[Hanin et al., 2018]



Li et al.; NeurIPS 2018; *Visualizing the Loss Landscape of Neural Nets*

Questions:

- **Do individual models in populations have something in common?**
- **Do they form meaningful structures in weight space?**
- **Can we learn representations of them?**
- **Can such structures be exploited to generate new models?**

Related Work

These questions have been partially addressed in previous work

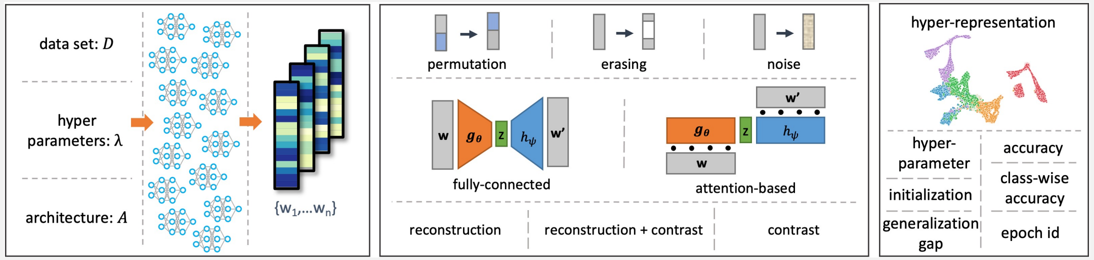
Discriminative: predict model properties

- Predict: accuracy, generalization gap, hyperparameters
- Features: weights *[Unterthiner et al., 2020; Martin et al., 2021]*, activations *[Jiang et al., 2019]*, graph-metrics *[Corneanu et al., 2020]*

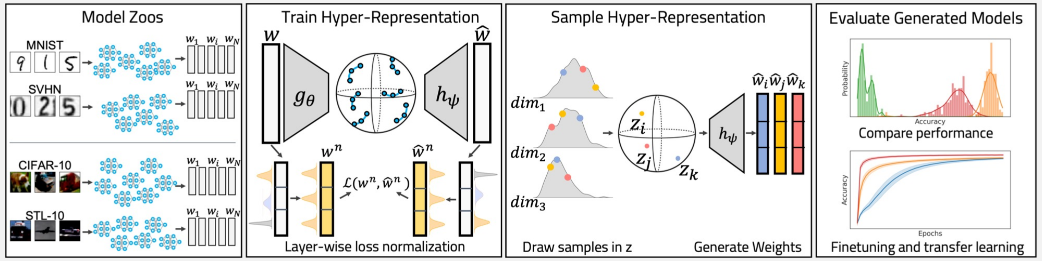
Generative: generate new models

- HyperNetworks *[Ha et al.. 2016; Deutsch, 2018; Zhang et al., 2020; Knyazev et al., 2021; Zhmoginov et al., 2022; Ratzlaff and Fuxin, 2019.]*
- Transfer Learning, Knowledge Distillation *[Shu et al., 2021; Liu et al., 2019.]*

Hyper-Representations: SSL representations of NN weights *[Schürholt et al., NeurIPS 2021]*



Generative Hyper-Representations *[Schürholt et al., NeurIPS 2022]*



Existing populations:
- small
- unstructured
- low diversity

This work:
- **large, structured, diverse populations**
- **open source: replicate, change, extend**

Model Zoo Generation

Overall:
- **27 model zoos**
- **50’360 unique NN models**
- **3’844’360 model states**

8 standard Image Classifcation Datasets:
MNIST, F-MNIST, SVHN, USPS, STL,
CIFAR10, CIFAR100, Tiny ImageNet

3 Architectures:
small CNN, medium CNN, ResNet-18

Varied Hyperparameters:
initalization method, activation function, optimizer,
learning rate, weight decay, dropout, seed

3 Zoo configurations:
- Variation of random seed
- Variation of hyperparamters (with 10 random/fixed seeds)

Sparse Model Zoo Twins

Table 1: Generating factors of the model zoos. Several values for each parameter define the grid. `Arch` denotes the architecture: CNN (s) - small CNN architecture, CNN (m) - medium CNN architecture, RN-18 - ResNet-18. `Init` denotes the initalization methods: U - uniform, N - normal, KU - Kaiming Uniform, KN - Kaiming Normal. `Activation` denotes the activation function: T - Tanh, S - Sigmoid, R - ReLU, G - GeLU. `Optim` denotes the optimizer: AD - Adam, SGD - Stochastic Gradient Descent. Models with learning rates denoted with * have been trained with a one-cycle LR scheduler, the listed LR is the maximum value.

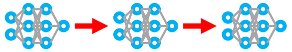
Dataset	Arch	Config	Init	Activation	Otpim	LR	WD	Dropout	Seed
MNIST	CNN (s)	Seed	U	T	AD	3e-4	0	0	1-1000
	CNN (s)	Hyp-10-r	U, N, KU, KN	T, S, R, G	AD, SGD	1e-3,1e-4	1e-3, 1e-4	0, 0.5	$\sim$ 10
	CNN (s)	Hyp-10-f	U, N, KU, KN	T, S, R, G	AD, SGD	1e-3,1e-4	1e-3, 1e-4	0, 0.5	1-10
F-MNIST	CNN (s)	Seed	U	T	AD	3e-4	0	0	1-1000
	CNN (s)	Hyp-10-r	U, N, KU, KN	T, S, R, G	AD, SGD	1e-3,1e-4	1e-3, 1e-4	0, 0.5	$\sim$ 10
	CNN (s)	Hyp-10-f	U, N, KU, KN	T, S, R, G	AD, SGD	1e-3,1e-4	1e-3, 1e-4	0, 0.5	1-10
SVHN	CNN (s)	Seed	U	T	AD	3e-3	0	0	1-1000
	CNN (s)	Hyp-10-r	U, N, KU, KN	T, S, R, G	AD, SGD	1e-3,1e-4	1e-3, 1e-4, 0	0, 0.3, 0.5	$\sim$ 10
	CNN (s)	Hyp-10-f	U, N, KU, KN	T, S, R, G	AD, SGD	1e-3,1e-4	1e-3, 1e-4, 0	0, 0.3, 0.5	1-10
USPS	CNN (s)	Seed	U	T	AD	3e-4	1e-3	0	1-1000
	CNN (s)	Hyp-10-r	U, N, KU, KN	T, S, R, G	AD, SGD	1e-3,1e-4	1e-2, 1e-3	0, 0.5	$\sim$ 10
	CNN (s)	Hyp-10-f	U, N, KU, KN	T, S, R, G	AD, SGD	1e-3,1e-4	1e-2, 1e-3	0, 0.5	1-10
CIFAR10	CNN (s)	Seed	KU	G	AD	1e-4	1e-2	0	1-1000
	CNN (s)	Hyp-10-r	U, N, KU, KN	T, S, R, G	AD, SGD	1e-3	1e-2, 1e-3	0, 0.5	$\sim$ 10
	CNN (s)	Hyp-10-f	U, N, KU, KN	T, S, R, G	AD, SGD	1e-3	1e-2, 1e-3	0, 0.5	1-10
CIFAR10	CNN (m)	Seed	KU	G	AD	1e-4	1e-2	0	1-1000
	CNN (m)	Hyp-10-r	U, N, KU, KN	T, S, R, G	AD, SGD	1e-3	1e-2, 1e-3	0, 0.5	$\sim$ 10
	CNN (m)	Hyp-10-f	U, N, KU, KN	T, S, R, G	AD, SGD	1e-3	1e-2, 1e-3	0, 0.5	1-10
STL (s)	CNN (s)	Seed	KU	T	AD	1e-4	1e-3	0	1-1000
	CNN (s)	Hyp-10-r	U, N, KU, KN	T, S, R, G	AD, SGD	1e-3,1e-4	1e-2, 1e-3	0, 0.5	$\sim$ 10
	CNN (s)	Hyp-10-f	U, N, KU, KN	T, S, R, G	AD, SGD	1e-3,1e-4	1e-2, 1e-3	0, 0.5	1-10
STL	CNN (m)	Seed	KU	T	AD	1e-4	1e-3	0	1-1000
	CNN (m)	Hyp-10-r	U, N, KU, KN	T, S, R, G	AD, SGD	1e-3,1e-4	1e-2, 1e-3	0, 0.5	$\sim$ 10
	CNN (m)	Hyp-10-f	U, N, KU, KN	T, S, R, G	AD, SGD	1e-3,1e-4	1e-2, 1e-3	0, 0.5	1-10
CIFAR10	RN-18	Seed	KU	R	SGD	1e-4*	5e-4	0	1-1000
CIFAR100	RN-18	Seed	KU	R	SGD	1e-4*	5e-4	0	1-1000
t-Imagenet	RN-18	Seed	KU	R	SGD	1e-4*	5e-4	0	1-1000

Potential Use-Cases

Model Analysis

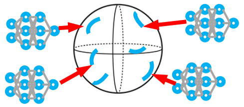
- Predict model properties w/o need of test data:
    - Accuracy
    - Generalization gap
    - Transfer-learning performance
- Diagnose model bias
- Derive model identity

Learning Dynamics



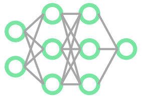
- Evaluate model potential
    - Early stopping
    - Population Based Training
    - Neural Architecture Search
- Improved understanding of learning dynamics
- Study sparsification trajectories

Representation Learning



- Analyze weight space
    - Low-loss regions
    - Git Re-Basin
- Weight Manifold Learning (Hyper-Representations)
- Diffusion models on learning trajectories
- Learning to optimize

Generating New Models



- Initialization
- Transfer Learning
    - Single pretrained models
    - Population based
- Ideal Ensemble Composition
- Generative Hyper-Representations

Acknowledgements

Model zoos available at **www.modelzoos.cc**

More info: **hsg.ai/neurips22**

Funding:

- Google Research Scholar Award (Damian Borth)
- HSG Basic Research Fund
- MCIN/ AEI /10.13039/501100011033

References

[1] Goodfellow, Vinyals, Saxe; ICLR 2015; *Qualitatively characterizing neural network optimization problems*

[2] Dauphin et al.; NeurIPS 2014; *Identifying and attacking the saddle point problem in high-dimensional non-convex optimization*

[3] LeCun, Bengion, Hinton; Nature 2015; *Deep Learning*

[4] Brown et al.; 2020; *Language Models are Few-Shot Learners*

[5] Larsen et al.; ICML 2021; *How many degrees of freedom do we need to train deep networks: a loss landscape perspective*

[6] Hanin, Rolnick; NeurIPS 2018; *How to Start Training: The Effect of Initialization and Architecture*

[7] Unterthiner, Keysers, Gelly, Bousquet, and Tolstikhin. Predicting Neural Network Accuracy from Weights. arXiv:2002.11448, February 2020.

[8] Martin, Peng, and Mahoney. Predicting trends in the quality of state-of-the-art neural networks without access to training or testing data. Nature Communications, 12(1):1–13, 2021.

[9] Jiang, Krishnan, Mobahi, Bengio. Predicting the Generalization Gap in Deep Networks with Margin Distributions. ICLR 2019.

[10] Corneanu, Madadi, Sergio Escalera, Aleix Martinez. Computing the Testing Error Without a Testing Set. CVPR 2020.

[11] Ha, ai, and Le. HyperNetworks. In arXiv:1609.09106 [Cs], 2016.

[12] Deutsch. Generating Neural Networks with Neural Networks. arXiv:1801.01952 [cs, stat], April 2018.

[13] Zhang, Ren, and Urtasun. Graph HyperNetworks for Neural Architecture Search. arXiv:1810.05749 [cs, stat], December 2020.

[14] Knyazev, Drozdzal, Taylor, and Romero-Soriano. Parameter Prediction for Unseen Deep Architectures. In Conference on Neural Information Processing Systems (NeurIPS), 2021.

[15] Zhmoginov, Sandler, and Vladymyrov. HyperTransformer: Model Generation for Supervised and Semi-Supervised Few-Shot Learning. arXiv:2201.04182 [cs], January 2022.

[16] Ratzlaff and Fuxin. HyperGAN: A Generative Model for Diverse, Performant Neural Networks. In Proceedings of the 36th International Conference on Machine Learning, May 2019.

[17] Shu, Kou, Cao, Wang, and Long. Zoo-tuning: Adaptive transfer from a zoo of models. In International Conference on Machine Learning, pages 457 9626–9637. PMLR, 2021.

[18] Liu, Peng, and Schwing. Knowledge Flow: Improve Upon Your Teachers. In International Conference on Learning Representations (ICLR), April 2019.

[19] Schürholt, Kostadinov, and Borth. Self-Supervised Representation Learning on Neural Network Weights for Model Characteristic Prediction. In Conference on Neural Information Processing Systems (NeurIPS), volume 35, 2021.

[20] Schürholt, Knyazev, Giro-i-Nieto, and Borth. Hyper-Representations as Generative Models: Sampling Unseen Neural Network Weights. Conference on Neural Information Processing Systems (NeurIPS), 2022.