Surrogate-Assisted Diversity Estimation in Neural Ensemble Search

Udeneev Alexandr Vladimirovich

Scientific Advisor: Ph.D. in Physics and Mathematics, Oleg Yurievich Bakhteev

Moscow Institute of Physics and Technology

December 20, 2025

Motivation and Approach

Problem
Develop an efficient method for Neural Ensemble Search (NES)

Challenge
Neural architecture search is computationally expensive, and constructing optimal ensembles leads to a exponential explosion of candidate ensemble configurations.

Proposed Approach
Employ a surrogate-assisted method by embedding architectures into a latent space, where similar networks are close and dissimilar ones are far.

Motivation
Efficient estimation of ensemble diversity enables better ensemble selection without exhaustive training.

Problem Statement

Let $\boldsymbol{\alpha} \in \mathcal{A}$ denote a network architecture, and $\boldsymbol{\omega}(\boldsymbol{\alpha})$ its trained weights. The output of a network on input $\boldsymbol{x}$ is $f_{\boldsymbol{\alpha}}(\boldsymbol{x}, \boldsymbol{\omega}(\boldsymbol{\alpha}))$. Given a subset $S \subset \mathcal{A}$, the NES problem is formally defined as:

$$\min_{\mathcal{S}} \mathcal{L}_{val} \left(\frac{1}{|S|} \sum_{\boldsymbol{\alpha} \in S} f_{\boldsymbol{\alpha}}(\boldsymbol{x}, \boldsymbol{\omega}^*(\boldsymbol{\alpha})) \right)$$

$$\text{s.t.} \quad \forall \boldsymbol{\alpha} \in \mathcal{S} : \; \boldsymbol{\omega}^*(\boldsymbol{\alpha}) = \arg\min_{\boldsymbol{\omega}(\boldsymbol{\alpha})} \mathcal{L}_{train}(f_{\alpha}(\boldsymbol{x}, \boldsymbol{\omega}(\boldsymbol{\alpha})))$$

Challenges:

- Exponentially growing number of architectures
- Exponentially growing number of ensemble combinations

Surrogate Functions

Overview

We propose two surrogate functions:

- $f_{\text{acc}}^{\theta} : \mathcal{A} \to \mathbb{R}$ — predicts network accuracy. Trained by minimizing mean squared error (MSE):

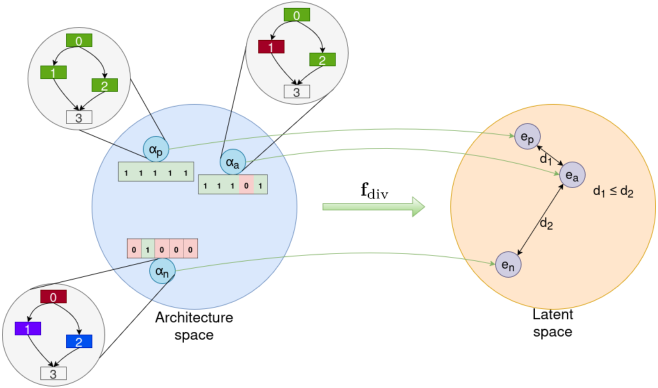
$$\theta^* = \arg\min_{\theta} \sum_i \left(f_{\text{acc}}^{\theta}(\alpha_i) - y_i \right)^2$$

- $\mathbf{f}_{\text{div}}^{\theta} : \mathcal{A} \to \mathbb{R}^d$ — maps architectures to a latent space. Trained with triplet loss to preserve diversity:

$$\theta^* = \arg\min_{\theta} \sum_{(\alpha_a, \alpha_p, \alpha_n)} \max \Big(\|\mathbf{f}_{\text{div}}^{\theta}(\alpha_a) - \mathbf{f}_{\text{div}}^{\theta}(\alpha_p)\|_2^2 - \|\mathbf{f}_{\text{div}}^{\theta}(\alpha_a) - \mathbf{f}_{\text{div}}^{\theta}(\alpha_n)\|_2^2 + m, 0 \Big)$$

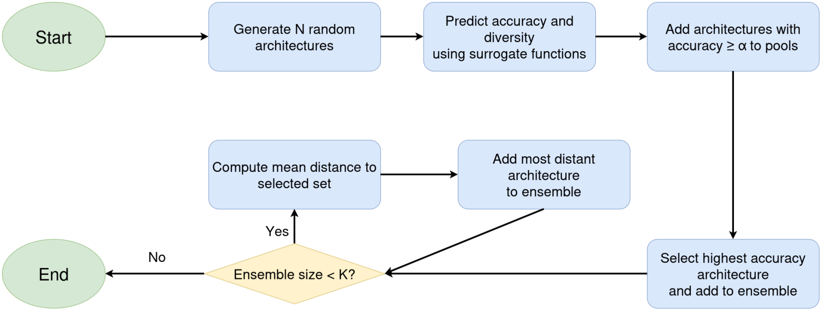
where $(\alpha_a, \alpha_p, \alpha_n)$ denote anchor, positive, and negative architectures, m $\in \mathbb{R}$ denote margin.

Surrogate Diversity Function



- Surrogate function maps discrete architectures into a continuous latent space.
- Similar architectures are embedded close together, while dissimilar ones are mapped farther apart (e.g., $\mathbf{e}_a$ vs $\mathbf{e}_p$ and $\mathbf{e}_a$ vs $\mathbf{e}_n$).
- Latent-space geometry allows efficient diversity evaluation without pre-training all architectures.

Surrogate assisted ensemble construction



The proposed ensemble construction algorithm.

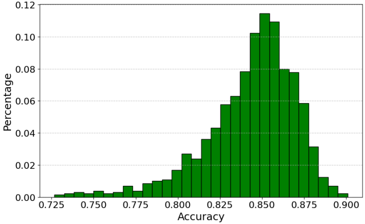
Goals of the Computational Experiment

On datasets FashionMNIST, CIFAR10, CIFAR100:

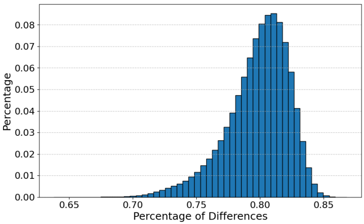
- Construct a training dataset of architectures and their evaluations
- Evaluate the surrogate models' ability to predict accuracy and diversity
- Assess the effectiveness of the ensemble construction strategy
- Compare surrogate-assisted method with DeepEns and RandomSearch

Dataset of Architectures

For each dataset, 3,000 neural architectures were partially trained.
The train/validation split was set to 0.8/0.2 for all datasets.
As an example, the resulting distributions for CIFAR10 are shown below:



Distribution of model accuracies



Distribution of model diversity

Comparison of Surrogate-Assisted NES and DeepEns

Metric	Surrogate-Assisted NES	DeepEns
FashionMNIST		
Ensemble size = 3		
Top-1 Accuracy	95.3%	**95.4**%
Average Model Accuracy	94.7%	**95.1**%
NLL	0.263	**0.256**
CIFAR10		
Ensemble size = 5		
Top-1 Accuracy	**97.80**%	97.61%
Average Model Accuracy	96.73%	96.73%
NLL	**0.208**	0.209
CIFAR100		
Ensemble size = 5		
Top-1 Accuracy	**85.17**%	84.16%
Average Model Accuracy	**80.50**%	79.54%
NLL	**0.692**	0.735

Conclusion

- We propose a surrogate function for predicting architectural diversity.
- We introduce a surrogate-assisted framework for neural ensemble construction.
- Our method matches or outperforms DeepEns and Random Search on FashionMNIST, CIFAR-10, and CIFAR-100 by explicitly accounting for architectural diversity during ensemble construction.
- The work is ready for submission to a conference or journal.