Surrogate-Assisted Diversity Estimation in Neural Ensemble Search

Udeneev Alexandr Vladimirovich

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

Moscow Institute of Physics and Technology
My First Research Paper

April 17, 2025

Research Goals

Objective:

To develop an effective surrogate-assisted method for Neural Ensemble Search (NES).

Motivation:

Accurate diversity estimation in ensembles is crucial for improving performance.

Proposed Approach:

Project network architectures into a latent space equipped with the Euclidean norm.

Related Work

This study builds upon the following research:

- Similarity surrogate-assisted evolutionary neural architecture search with dual encoding strategy (Xue, Zhang, Neri, 2024)
- Neural Predictor for Neural Architecture Search (Wen et al., 2019)

Problem Statement

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

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

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

Challenges:

- A vast search space of possible architectures
- Exponentially many ensemble combinations

Surrogate Functions

Overview

We propose using two surrogate functions:

- $f_{\text{acc}} : \mathcal{A} \to \mathbb{R}$ — a regression model that predicts the accuracy of a network given its architecture.
- $f_{\text{div}} : \mathcal{A} \to \mathbb{R}^d$ — a function that maps architectures to a latent space. Trained using Triplet Loss based on a similarity matrix derived from model outputs.

Ensemble Construction

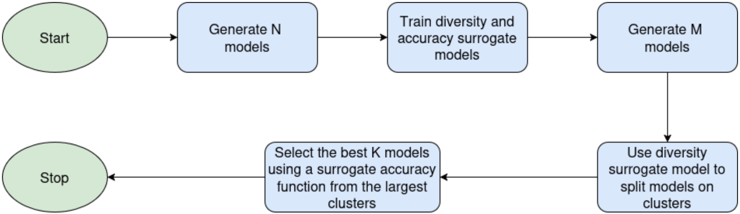


Figure: The proposed ensemble construction algorithm.

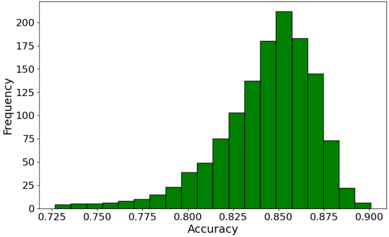
Goals of the Computational Experiment

1. Construct a training dataset of architectures and their evaluations
2. Evaluate the surrogate models' ability to predict accuracy and diversity
3. Assess the effectiveness of the ensemble construction strategy

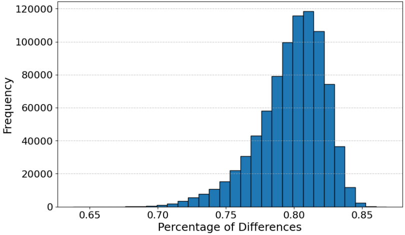
Dataset Collection

We generate architectures from the DARTS search space.

- A total of 1,000 architectures were trained to obtain ground truth accuracy and diversity values.
- For inference, we generated 250,000 candidate architectures.



Distribution of model accuracies



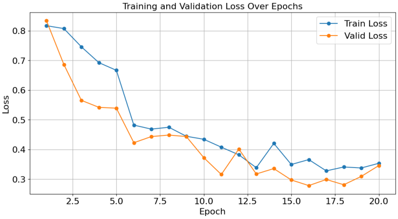
Distribution of model diversity

Surrogate Model Performance

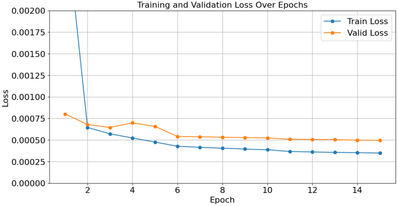
We trained:

- the accuracy surrogate using the Mean Squared Error (MSE) loss,
- the diversity surrogate using the Triplet Loss.

To evaluate the learned diversity representation, we computed the Pearson correlation between distances in the latent space and output similarity, obtaining approximately –0.4, indicating that the surrogate captures the inverse relationship.



Training process of the diversity surrogate

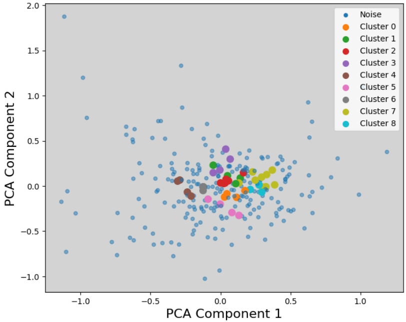


Training process of the accuracy surrogate

Ensemble Effectiveness

In our main experiment:

- We generated 250,000 candidate architectures from the DARTS space.
- The top five models achieved an average test accuracy of approximately 79% on CIFAR-10.
- The ensemble of these five models reached approximately 84% accuracy.



Conclusion

1. Proposed a method to encode architectures from the DARTS search space for training a Graph Convolutional Network (GCN), with nodes representing network operations.

2. Developed a surrogate model for predicting architectural diversity based on latent representations.

3. Demonstrated the effectiveness of using surrogate models to guide ensemble construction, balancing performance and diversity.