Research Article

Alexandr Udeneev, Petr Babkin, and Oleg Bahteev

Surrogate assisted diversity estimation in NES

https://doi.org/10.1515/sample-YYYY-XXXX
Received Month DD, YYYY; revised Month DD, YYYY; accepted Month DD, YYYY

Abstract: The automated search for optimal neural network architectures (NAS) is a challenging computational problem, and Neural Ensemble Search (NES) is even more complex. In this work, we propose a surrogate-based approach for ensebmle creation. Neural architectures are represented as graphs, and their predictions on a dataset serve as training data for the surrogate function. Using this function, we develop an efficient NES framework that enables the selection of diverse and high-performing architectures. The resulting ensemble achieves superior predictive accuracy on CIFAR-10 compared to other one-shot NES methods, demonstrating the effectiveness of our approach.

Keywords: NES, GCN, triplet loss, surrogate function.

1 Introduction

Neural network ensembles often demonstrate better accuracy compared to single models, especially in classification and regression tasks [1, 2]. This fact gives rise to the problem of constructing an efficient ensemble of models (NES) [3]. NES, in turn, relies on Neural Architecture Search (NAS) methods, which are extensively studied and applied to search for individual neural network architectures, such as evolutionary algorithms [4, 5], reinforcement learning [6–8], and Bayesian optimization [9, 10]. Selecting an optimal architecture for even a single model is a challenging task, particularly when considering data-specific constraints and computational limitations [11].

The simplest approach for ensemble construction is the use of DeepEns [12], implemented through DARTS [13]. It involves a random search for several architectures, which are then combined into an ensemble. Despite its simplicity in implementation and hyperparameter tuning, this method is computationally expensive. More sophisticated adaptation techniques are presented in some recent works [3, 14, 15], which are designed to efficiently combine multiple networks into an ensemble.

Our research also adapts ideas from NAS for NES, specifically using a surrogate function. Some modern NAS methods widely use surrogate functions to estimate architecture quality without requiring full model training [16–18]. These functions significantly reduce computational costs, expanding the applicability of such methods. For example, in [16], evolutionary algorithms were proposed in combination with surrogate models for real-time semantic segmentation. In [18], a Surrogate-assisted Multiobjective Evolutionary-based Algorithm (SaMEA) is used for 3D medical image segmentation.

In this work, we propose a method for constructing neural network ensembles using a surrogate function that accounts for both model classification accuracy and architectural diversity. Diversity is crucial because ensembles consisting of similar models often fail to provide a significant performance gain. The surrogate function is used to encode the architecture into a latent space [19], which reflects both the diversity and predictive ability of the architectures. Since a neural network architecture is represented as a graph, using a Graph Neural Network (GNN) [20] as a surrogate function [21] seems natural. To train it to predict model diversity, we use Triplet Loss [22], similar to [19]. We validate this approach on CIFAR-10, demonstrating the effectiveness of the surrogate function for predicting diversity and constructing ensembles. We claim that ensembles constructed in this manner achieve state-of-the-art accuracy compared to one-shot NES algorithms, such as DeepEns [12].

Main Contributions:

1) We propose a method for encoding the DARTS [13] search space into a representation suitable for training a Graph Neural Network (GNN), where graph nodes correspond to operations within the network.

2) We propose a way for training the surrogate function to predict the diversity of architectures.

3) We adapt surrogate functions for ensemble construction, taking into account both predictive performance and architectural diversity.

2 Problem statement

2.1 Neural Architecture Search

Let $\mathcal{V} = 1, \dots, N$ be the set of vertices, where N is the number of vertices, and let $\mathcal{E} = \{(i,j) \in \mathcal{V} \times \mathcal{V} \mid i < j\}$ be the set of edges connecting them. Furthermore, let $\mathcal{O}$ denote the set of possible operations between vertices (e.g., pooling, convolutions, etc.).For each edge there is an operation $o \in \mathcal{O}$ that transits information from one node to another. The neural architecture search (NAS) problem can be formulated as finding an operation $o^{(i,j)} \in \mathcal{O}$ for each edge (i, j).

Consider $\alpha \in \mathcal{A}$ as a parameter vector representing the operations assigned to edges. Then, the NAS problem can be formulated as:

$$\begin{aligned} &\min_{\alpha \in \mathcal{A}} \mathcal{L}_{val}(\omega^* \alpha, \alpha) \\ &\text{s.t.} \quad \omega^*_\alpha = \arg\min_{\omega \in \mathcal{W}} \mathcal{L}_{train}(\omega, \alpha) \end{aligned} \tag{1}$$

where $\mathcal{W}$ is the set of all possible weights associated with operations for all potential edges in the architecture. The main challenge is the vast architecture search space (e.g., in DARTS [13], it is approximately 10^{25}).

References

[1] P.N. Suganthan Ye Ren, Le Zhang. Ensemble classification and regression-recent developments, applications and future directions [review article]. *IEEE Computational Intelligence Magazine*, 11(1):41–53, Feb 2016. ISSN 1556-603X. 10.1109/mci.2015.2471235. URL https://doi.org/10.1109/mci.2015.2471235.

[2] Lars Kai Hansen and Peter Salamon. Neural network ensembles. *IEEE Trans. Pattern Anal. Mach. Intell.*, 12(10): 993–1001, 1990. 10.1109/34.58871.

[3] Sheheryar Zaidi, Arber Zela, Thomas Elsken, Chris C. Holmes, Frank Hutter, and Yee Whye Teh. Neural ensemble search for uncertainty estimation and dataset shift. In Marc'Aurelio Ranzato, Alina Beygelzimer, Yann N. Dauphin, Percy Liang, and Jennifer Wortman Vaughan, editors, *Advances in Neural Information Processing Systems 34: Annual Conference on Neural Information Processing Systems 2021, NeurIPS 2021, December 6-14, 2021, virtual*, pages 7898–7911, 2021. URL https://proceedings.neurips.cc/paper/2021/hash/41a6fd31aa2e75c3c6d427db3d17ea80-Abstract.html.

[4] Esteban Real, Sherry Moore, Andrew Selle, Saurabh Saxena, Yutaka Leon Suematsu, Jie Tan, Quoc V Le, and Alexey Kurakin. Large-scale evolution of image classifiers. In *International conference on machine learning*, pages 2902–2911. PMLR, 2017.

[5] Esteban Real, Alok Aggarwal, Yanping Huang, and Quoc V Le. Regularized evolution for image classifier architecture search. In *Proceedings of the aaai conference on artificial intelligence*, volume 33, pages 4780–4789, 2019.

[6] Barret Zoph and Quoc V. Le. Neural architecture search with reinforcement learning. In *5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017, Conference Track Proceedings*. OpenReview.net, 2017. URL https://openreview.net/forum?id=r1Ue8Hcxg.

[7] Sirui Xie, Hehui Zheng, Chunxiao Liu, and Liang Lin. Snas: stochastic neural architecture search. *arXiv preprint arXiv:1812.09926*, 2018.

[8] Yuqiao Liu, Yanan Sun, Bing Xue, Mengjie Zhang, Gary G. Yen, and Kay Chen Tan. A survey on evolutionary neural architecture search. *IEEE Trans. Neural Networks Learn. Syst.*, 34(2):550–570, 2023. 10.1109/TNNLS.2021.3100554.

[9] Haifeng Jin, Qingquan Song, and Xia Hu. Auto-keras: An efficient neural architecture search system. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*, pages 1946–1956, 2019.

[10] Kirthevasan Kandasamy, Willie Neiswanger, Jeff Schneider, Barnabas Poczos, and Eric P Xing. Neural architecture search with bayesian optimisation and optimal transport. *Advances in neural information processing systems*, 31, 2018.

[11] Ankit Kumar Saroj Kumar Pandey Neeraj varshney Teekam Singh Chetan Swarup, Kamred Udham Singh. Brain tumor detection using cnn, alexnet amp;amp; googlenet ensembling learning approaches. *Electronic Research Archive*, 31(5): 2900–2924, 2023. ISSN 2688-1594. 10.3934/era.2023146. URL https://doi.org/10.3934/era.2023146.

[12] Balaji Lakshminarayanan, Alexander Pritzel, and Charles Blundell. Simple and scalable predictive uncertainty estimation using deep ensembles. *Advances in neural information processing systems*, 30, 2017.

[13] Hanxiao Liu, Karen Simonyan, and Yiming Yang. DARTS: differentiable architecture search. *CoRR*, abs/1806.09055, 2018. 10.48550/arxiv.1806.09055. URL http://arxiv.org/abs/1806.09055.

[14] Yao Shu, Yizhou Chen, Zhongxiang Dai, and Bryan Kian Hsiang Low. Neural ensemble search via bayesian sampling. In James Cussens and Kun Zhang, editors, *Proceedings of the Thirty-Eighth Conference on Uncertainty in Artificial Intelligence*, volume 180 of *Proceedings of Machine Learning Research*, pages 1803–1812. PMLR, 01–05 Aug 2022. URL https://proceedings.mlr.press/v180/shu22a.html.

[15] Haibin Ling Minghao Chen, Jianlong Fu. One-shot neural ensemble architecture search by diversity-guided search space shrinking. *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 16525–16534, Jun 2021. 10.1109/cvpr46437.2021.01626. URL https://doi.org/10.1109/cvpr46437.2021.01626.

[16] Zhichao Lu, Ran Cheng, Shihua Huang, Haoming Zhang, Changxiao Qiu, and Fan Yang. Surrogate-assisted multi-objective neural architecture search for real-time semantic segmentation. *CoRR*, abs/2208.06820, 2022. 10.48550/ARXIV.2208.06820.

[17] Zhichao Lu, Kalyanmoy Deb, Erik D. Goodman, Wolfgang Banzhaf, and Vishnu Naresh Boddeti. Nsganetv2: Evolutionary multi-objective surrogate-assisted neural architecture search. In Andrea Vedaldi, Horst Bischof, Thomas Brox, and Jan-Michael Frahm, editors, *Computer Vision - ECCV 2020 - 16th European Conference, Glasgow, UK, August 23-28, 2020, Proceedings, Part I*, volume 12346 of *Lecture Notes in Computer Science*, pages 35–51. Springer, 2020. 10.1007/978-3-030-58452-8_3.

[18] Maria G. Baldeon Calisto and Susana K. Lai-Yuen. Emonas-net: Efficient multiobjective neural architecture search using surrogate-assisted evolutionary algorithm for 3d medical image segmentation. *Artif. Intell. Medicine*, 119:102154, 2021. 10.1016/J.ARTMED.2021.102154.

[19] Ferrante Neri Yu Xue, Zhenman Zhang. Similarity surrogate-assisted evolutionary neural architecture search with dual encoding strategy. *Electronic Research Archive*, 32(2):1017–1043, 2024. ISSN 2688-1594. 10.3934/era.2024050. URL https://doi.org/10.3934/era.2024050.

[20] Thomas N. Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. In *5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017, Conference Track Proceedings*. OpenReview.net, 2017. URL https://openreview.net/forum?id=SJU4ayYgl.

[21] Wei Wen, Hanxiao Liu, Yiran Chen, Hai Li, Gabriel Bender, and Pieter-Jan Kindermans. Neural predictor for neural architecture search. In *European conference on computer vision*, pages 660–676. Springer, 2020.

[22] Florian Schroff, Dmitry Kalenichenko, and James Philbin. Facenet: A unified embedding for face recognition and clustering. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 815–823, 2015.