Information Fusion 118 (2025) 102925

Contents lists available at ScienceDirect

Information Fusion

journal homepage: www.elsevier.com/locate/inffus

Full length article

Rethinking neural architecture representation for predictors: Topological encoding in pixel space

Caiyang Yu, Jian Wang, Yifan Wang, Wei Ju *, Chenwei Tang *, Jiancheng Lv

College of Computer Science, Sichuan University, Chengdu, 610065, Sichuan, China

ARTICLE INFO

Keywords:
Neural architecture representation
Neural predictor
Pixel space
Information fusion

ABSTRACT

Neural predictors (NPs) aim to swiftly evaluate architectures during the neural architecture search (NAS) process. Precise evaluations with NPs heavily depend on the representation of training samples (*i.e.*, the architectures), as the representation determines how well the NP captures the intrinsic properties and intricate dependencies of the architecture. Existing methods, which represent neural architectures as graph structures or sequences, are inherently limited in their expressive capabilities. In this study, we explore the image representation of neural architecture, describing the architecture in pixel space and using the long-range modeling capability of attention mechanisms to construct connections among pixels and extract tangible (tractable) architecture topology representation from them. Our attempt provides an efficient architecture representation for NPs, combined with today's powerful pre-training models, showing promising prospects. Furthermore, recognizing that images alone may fall short in capturing configuration specifics, we design a corresponding text representation to provide a more accurate complement. Our experimental analysis reveals that the existing visual language model can efficiently identify the topological information in the pixel space. Additionally, we propose a Dual-Input Multichannel Neural Predictor (DIMNP) that simultaneously accepts multiple representations of architectures, facilitating information complementarity and accelerating convergence of the NP. Extensive experiments on NAS-Bench-101, NAS-Bench-201, and DARTS datasets demonstrate the superiority of DIMNP compared to the state-of-the-art NPs. In particular, on the NAS-Bench-101 and NAS-Bench-201 search spaces, DIMNP achieves performance improvements of 0.01 and 0.52, respectively, compared to the second-best algorithm on average.

1. Introduction

Rapid advancements in deep learning [1–5] have heightened the complexity and diversity of neural network architectures, making traditional manual design methods inadequate for efficient optimization. While Neural Architecture Search (NAS) [6–8] addresses these challenges, it is still limited by high computational costs and time constraints. As a result, Neural predictor (NP) has become one of the most sought-after tools in NAS with its ability to evaluate architectures rapidly [9]. The capability of the NP to make precise evaluations is crucial for identifying the optimal architectures. Therefore, a well-crafted representation of the neural architecture is indispensable, which can substantially enhance the NP's comprehension of the internal dependencies within the architecture, thus improving its performance and reliability. Current mainstream research on architectural representations can be classified into two categories: sequence-based and graph-based methods. The sequence-based methods encode architectures as sequences or strings, where each character or symbol represents specific network components or configurations [10,11]. For the graph-based method, neural architectures are represented as directed acyclic graphs (DAGs), where nodes and edges represent network operations and connections, closely mirroring the actual structural layout of neural networks [12,13].

However, the aforementioned encoding methods significantly limit the expressive power of architectural representations and struggle to establish global relationships within an architecture. Sequence-based methods often fail to capture tight connections in complex structures over long sequences, while graph-based methods can result in representations that miss important internal dependencies, especially when dealing with architectures that have many unconnected nodes, leading to sparsity issues. Thus, developing a more comprehensive representation of neural architectures remains a crucial challenge. Inspired by the success of pre-training visual models [16,17], an intuitive approach

* Corresponding authors.
E-mail addresses: yucy324@gmail.com (C. Yu), jianwang.scu@gmail.com (J. Wang), wangyifan5217@stu.scu.edu.cn (Y. Wang), juwei@scu.edu.cn (W. Ju), tangchenwei@scu.edu.cn (C. Tang), lvjiancheng@scu.edu.cn (J. Lv).

https://doi.org/10.1016/j.inffus.2024.102925
Received 1 December 2024; Received in revised form 24 December 2024; Accepted 29 December 2024
Available online 10 January 2025
1566-2535/© 2025 Elsevier B.V. All rights are reserved, including those for text and data mining, AI training, and similar technologies.