2023 IEEE Symposium Series on Computational Intelligence (SSCI)
Mexico City, Mexico. December 5-8, 2023

# Examination of the Multimodal Nature of Multi-objective Neural Architecture Search

Cheng Gong*†‡, Yang Nan*, Lie Meng Pang*, Hisao Ishibuchi*, Qingfu Zhang†‡

* Guangdong Provincial Key Laboratory of Brain-inspired Intelligent Computation, Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China

†Department of Computer Science, City University of Hong Kong, Kowloon Tong, Hong Kong

‡The City University of Hong Kong Shenzhen Research Institute, Shenzhen, China

chenggong6-c@my.cityu.edu.hk, nany@mail.sustech.edu.cn, panglm@sustech.edu.cn, hisao@sustech.edu.cn, qingfu.zhang@cityu.edu.hk

***Abstract*—Remarkable successes in deep learning have spurred significant growth in the field of neural architecture search (NAS), which is rapidly advancing as a promising technique for automating the design of network architecture. From an optimization standpoint, a NAS task for a given search space can be viewed as a multi-objective optimization problem (MOP) when considering multiple design criteria simultaneously (e.g., prediction accuracy, architecture complexity, hardware efficiency). However, whether a NAS problem is a multimodal multi-objective optimization problem or not (i.e., whether a single non-dominated solution in the objective space has multiple different neural network architectures or not) has not been examined in the literature. This presents an intriguing research question that merits further investigation. To fill this gap, we examine the multimodal nature of seven multi-objective NAS problems. By doing so, this work aims to help MOP researchers to better understand the characteristics of the multi-objective NAS problems.**

***Index Terms*—Neural architecture search, Multi-objective optimization, Multimodal multi-objective optimization.**

## I. Introduction

Recent years have witnessed the emergence of novel neural network architecture designs such as ResNet [1], Inception [2], Transformer [3] and GPT-3 [4]. These new architectures have greatly promoted the successes of deep learning in many real-world scenarios. However, manual design of an appropriate deep neural network architecture for a given task typically requires the involvement of human experts and a number of trials and errors [5]. That is, this process is time-consuming and prone to errors.

In contrast, neural architecture search (NAS) is a promising technique that automates the design of neural network architecture, which decreases the heavy involvement of human expert. The common goal of NAS is to find an appropriate neural network architecture in terms of prediction accuracy (or prediction error), which is the most important objective. However, in real applications, we also need to consider some other objectives (which are usually conflicting with the main objective: accuracy maximization). For instance, in certain scenarios such as hardware-dependent NAS tasks targeting mobile or embedded devices [6], it is crucial to strike a balance between accuracy and speed (computation time). Designing a neural network solely for high accuracy may lead to a large-scale computationally expensive model, which is difficult to implement in low-power devices. Therefore, from the optimization point of view, NAS tasks can be treated as multi-objective optimization problems (i.e., multi-objective NAS [7]).

Although multi-objective NAS can fall under the field of multi-objective optimization (MOO) [8], it is very challenging to employ conventional optimization methods, such as gradient-based methods [9], to solve them. First, due to the blackbox nature of NAS, it is difficult to mathematically formulate each objective function. Second, the decision-maker may have unknown preference among objectives. This means that the use of scalarizing approaches is not always appropriate (e.g., we cannot use the weighted sum with the pre-specified weight values for all decision makers since each decision maker has different preferences). One promising approach to search for a wide variety of non-dominated solutions along the trade-off surface among multiple objectives is the use of population-based evolutionary multi-objective optimization (EMO) algorithms [10]–[12]. We can find a number of different non-dominated solutions by a singe run of an EMO algorithm.

Recently, evolutionary computation (EC) methods have gained significant attention in the field of NAS due to their impressive performance in discovering optimal or near-optimal neural network architecture [13]–[15]. However, compared to the overall progress in the NAS field, utilizing EMO algorithms to tackle multi-objective NAS still falls behind [16]. To better design advanced EMO algorithms targeting multi-objective NAS, it is not trivial to study whether there exists some specific characteristics of fitness landscapes in multi-objective NAS from the optimization point of view, such as the multimodal nature [17]. In fact, many real-world problems can be classified as multimodal multi-objective problems. It is essential to identify whether there exists multimodal nature in

Corresponding Authors: Hisao Ishibuchi and Qingfu Zhang.

This work was supported by National Natural Science Foundation of China (Grant No. 62250710163, 62250710682), Guangdong Provincial Key Laboratory (Grant No. 2020B121201001), the Program for Guangdong Introducing Innovative and Enterpreneurial Teams (Grant No. 2017ZT07X386), the Stable Support Plan Program of Shenzhen Natural Science Fund (Grant No. 20200925174447003), Shenzhen Science and Technology Program (Grant No. KQTD2016112514355531), the Research Grants Council of the Hong Kong Special Administrative Region, China (GRF Project No. CityU11215622), and Natural Science Foundation of China (Project No: 62276223).

multi-objective NAS. On one hand, it is very likely that different model architectures can have the same performance on a given task. On the other hand, specific designed algorithms for multimodal multi-objective optimization may be needed once identifying the multimodal nature of multi-objective NAS. This is because compared to the common EMO algorithms, multimodal multi-objective evolutionary algorithms (MMEAs) have the advantage of maintaining the solution space diversity [18]. In this paper, to address the above issue, we examine the multimodal nature of multi-objective NAS by using a recently introduced multi-objective NAS benchmark called *EvoXBench* [16]. To be specific, we focus on examining the multimodal nature of specific test problems with the NAS-Bench-101 (NB101) [19], NAS-Bench-201 (NB201) [20], and NATS-Bench search space [21]. These tabular benchmark problems can support our exhaustive evaluation of all unique model architectures over the entire search space.

The remainder of this paper is organized as follows. In Section II, the preliminaries on multi-objective NAS and multimodal nature are introduced. The details of the examined multi-objective NAS benchmark problems are described in Section III. Experimental results are reported in Section IV to illustrate whether there exists multimodal nature on the general multi-objective NAS. Finally, the conclusion is given in Section V.

## II. Preliminaries

### A. Multi-objective NAS

In general, a multi-objective NAS task is inherently a bilevel problem, where the upper-level task is to optimize architecture and the lower-level task is to optimize the parameter of a given model architecture. Mathematically, it can be formulated as a multi-objective problem as follows:

$$\begin{aligned} &\underset{\boldsymbol{x}\in\boldsymbol{\Omega}_x}{\text{Minimize}} \;\; \boldsymbol{F}(\boldsymbol{x}) = [f_1(\boldsymbol{x};\boldsymbol{\omega}^*(\boldsymbol{x})), f_2(\boldsymbol{x}), ..., f_m(\boldsymbol{x})] \\ &\text{Subject to} \;\; \boldsymbol{\omega}^*(\boldsymbol{x}) \in \underset{\boldsymbol{\omega}\in\boldsymbol{\Omega}_\omega}{\arg\min}\, \mathcal{L}(\boldsymbol{\omega};\boldsymbol{x}) \end{aligned}$$

where $\boldsymbol{\Omega}_x$ and $\boldsymbol{\Omega}_\omega$ are the architecture space and the associated weight space, respectively. $\boldsymbol{F}(\boldsymbol{x})$ is a $m$-dimensional vector that represents $m$ objectives. $f_1(\cdot)$ denotes the prediction error objective that conditions on the trained optimal weight $\boldsymbol{\omega}^*$; $f_2(\cdot), ..., f_m(\cdot)$ denote the other objectives that only condition on the model architecture (e.g., model complexity [22] and hardware latency [23]).

### B. Multimodal Nature

According to a recent review [18], the definition of multimodal multi-objective optimization problems (MMOPs) is still controversial. One common consensus on defining MMOPs is that multiple clearly separated solutions in the decision space map to the same or very similar points on the Pareto front in the objective space. These solutions are considered equivalent. In this paper, we consider a multi-objective NAS problem that exhibits multimodal nature (i.e., MMOPs) when there exist multiple equivalent optimal solutions that correspond to different model architectures, but achieve the same or similar level of performance on each objective.

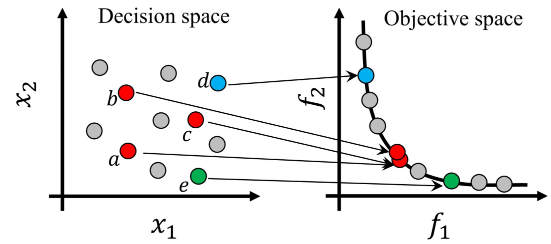



Fig. 1: Illustration of multimodal nature.

We define the equivalency as follows: formally, for two given separated solutions $\boldsymbol{x}$ and $\boldsymbol{y}$ in the decision space, they are defined as equivalent solutions if $\|\boldsymbol{F}(\boldsymbol{x}) - \boldsymbol{F}(\boldsymbol{y})\| \leq \varepsilon$, where $\|\cdot\|$ is an arbitrary norm operator and $\varepsilon$ is a predefined non-negative real number that controls the relaxation on equivalency. As illustrated in Fig. 1, when $\varepsilon = 0$, only the solutions $a$ and $c$ that map to the exact same point in the objective space are defined as equivalent. When we allow a relaxation on equivalency, i.e., $\varepsilon > 0$, the solution $a$, $b$, and $c$ can be equivalent. Both solutions $d$ and $e$ have no equivalent solution.

In real-world scenarios, many MOPs have proved to exhibit multimodal natures, such as architecture layout design [24], multi-objective knapsack optimization problem [25] and rock engine design [26]. When we solve MOPs, it is worth to finding different equivalent solutions that are separated in the decision space but close in the objective space. This because multiple alternative solutions will benefit the decision makers by providing more choices, especially when some solutions are hard to attain in reality.

In the context of multi-objective NAS, it is very likely that there exists multimodal nature on this real-world problem. This is because we can expect that different model architectures show the same or very similar performances on each objective. Note that the study [16] showed that NAS with NB101 and NB201 search spaces exhibit multimodal natures, as they contain a number of different optimal architectures that has very close performances in terms of prediction error (i.e., single-objective NAS). However, whether there exists multimodal nature on multi-objective NAS is still unclear. To examine the multimodal nature of multi-objective NAS, we consider both the exact equivalency and relaxed equivalency (i.e., $\varepsilon = 0$ and $\varepsilon > 0$). This is because obtaining the exact equivalent solutions may be hard on the continuous objective space of multi-objective NAS. Additionally, the stochastic training process for optimal weights will add noise to the evaluation of a model architecture (i.e., prediction error).

## III. Examined Multi-objective NAS Benchmark

In this section, we will provide a thorough introduction to the multi-objective NAS benchmark problems that were examined to determine whether they exhibit multimodal natures.

TABLE I: The detailed information of the examined multi-objective NAS benchmark problems in this paper.

| Problem | $\Omega$ | $\lvert\Omega\rvert$ | $D$ | $M$ | Objectives |
|---|---|---|---|---|---|
| C-10/MOP1 | NB101 | 423,624 | 26 | 2 | $f^e$, $f_1^c$ |
| C-10/MOP2 | NB101 | 423,624 | 26 | 3 | $f^e$, $f_1^c$, $f_2^c$ |
| C-10/MOP3 | NATS | 32,768 | 5 | 3 | $f^e$, $f_1^c$, $f_2^c$ |
| C-10/MOP4 | NATS | 32,768 | 5 | 4 | $f^e$, $f_1^c$, $f_2^c$, $f_1^{h_1}$ |
| C-10/MOP5 | NB201 | 15,625 | 6 | 5 | $f^e$, $f_1^c$, $f_2^c$, $f_1^{h_1}$, $f_2^{h_1}$ |
| C-10/MOP6 | NB201 | 15,625 | 6 | 6 | $f^e$, $f_1^c$, $f_2^c$, $f_1^{h_2}$, $f_2^{h_2}$, $f_3^{h_2}$ |
| C-10/MOP7 | NB201 | 15,625 | 6 | 8 | $f^e$, $f_1^c$, $f_2^c$, $f_1^{h_1}$, $f_2^{h_1}$, $f_1^{h_2}$, $f_2^{h_2}$, $f_3^{h_2}$ |

A recent study [16] proposed a unified and comprehensive multi-objective NAS benchmark, dubbed *EvoXBench*, which covers seven search spaces and up to six objectives (i.e., prediction error, model complexity related performances, and hardware devices related performances). We conduct experiments to examine the multimodal nature of the specific test problems with NAS-Bench-101, NAS-Bench-210 and NATS-Bench search space. Note that only the channel size search space (i.e., configuration for the channel size in each layers) is considered as the search space of model architecture for multi-objective NAS benchmark problems. Thus, all these three search spaces have a limited size of search space, enabling exhaustive examination of their multimodal natures over the entire search space. In Table I, we list the details of these examined multi-objective NAS benchmark problems.

In Table I, $\boldsymbol{\Omega}$ denotes the model architecture search space and $|\boldsymbol{\Omega}|$ is the number of totally unique models in terms of the encoding in their original paper; $D$ and $M$ are the number of decision variables and objectives, respectively. The number of objectives of the examined test problems ranges from two to eight, including the mean prediction error $f^e$, the model complexity related performances $\boldsymbol{f}^c$, and the hardware devices related performances $\boldsymbol{f}^H$. In the above test problems, the model complexity related performances include the number of weights and floating point operations. The hardware devices set $H$ of these test problems includes GPUs ($h_1$) and Eyeriss ($h_2$) [27], with the related performances of hardware latency, energy consumption and arithmetic intensity.

For NB101 search space, each model architecture is encoded by a 7-vertex directed acyclic graph (DAG) with 21 possible edges, which can be represented by a $7 \times 7$ upper-triangular binary matrix. Three operations including $3 \times 3$ convolution, $1 \times 1$ convolution and $3 \times 3$ max-pooling can be chosen for each of the 5 vertices (removing the fixed input and output vertices). Thus, there are a total $2^{21} \times 3^5 \approx 510M$ possible unique graphs in this encoding. However, a large number of encodings correspond to invalid graphs, since the maximum number of edges is limited to nine for exhaustive enumeration. Also, some different encodings decode isomorphic graphs that exhibit equivalent model architecture. After de-duplication and filtering invalid encodings, we identify 423,624 total unique model architectures for NB101 search space.

For NB201 search space, each model architecture is represented as a directed acyclic graph with 4 nodes and 6 edges, where the nodes are densely connected. Each edge is associated with 5 representative operation candidates, including $zeroize$, $skip\ connect$, $1\times1$ $convolution$, $3\times3$ $convolution$ and $3\times3$ $average\ pooling$, where $zeroize$ is the operation of dropping the edge that eliminates the restriction on the search topology of DAG. Thus, for NB201 search space, there are $5^6 = 15625$ total unique model architecture by using the encoding of 6-dimensional vectors, where the $i$-$th$ element denotes the operation in the $i$-$th$ edge.

NATS is a newly proposed NAS benchmark for both the topology search and channel size search, where the size refers to the number of channels configured to each layer. For the multi-objective NAS test problems with NATS search space listed in Table 1, only the channel size search space is considered. In the channel size search space of NATS, 8 candidates for the number of channels are pre-defined for each of the 6 layers. Thus, there are $8^5 = 32767$ total unique models that can be encoded by a 5-dimensional vector, where the $i$-$th$ elements denotes the number of channels in the $i$-$th$ layer.

## IV. Experimental Study

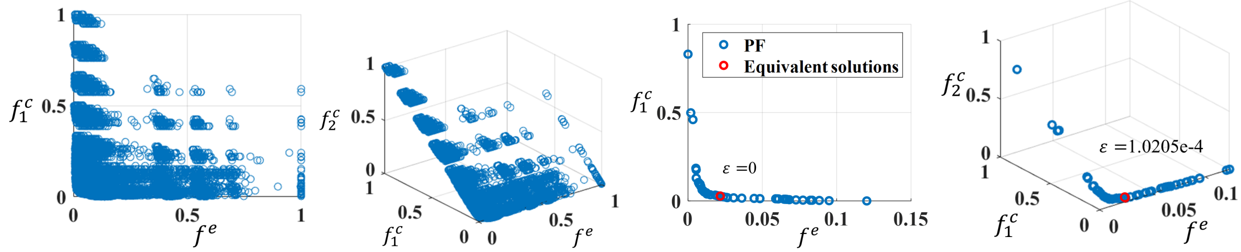
In this section, we show our experimental study on examining the multimodal natures of the multi-objective NAS benchmark problems introduced in Section III. Specially, we examine the benchmark problems with NB101, NB201, and NATS search space, respectively.

### A. Multimodal Nature on NB101 Search Space

We first examine whether there exists multimodal natures on the multi-objective NAS benchmark problems with NB101 search space, i.e., C-10/MOP1-2 test problems.

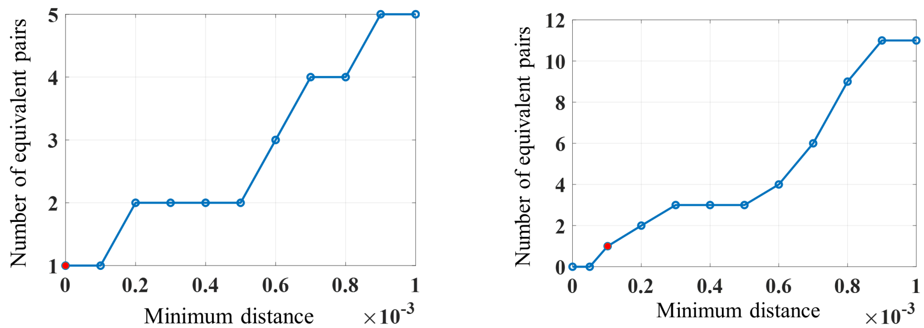
As introduced in Section III, in NB101 search space, a large number of encodings will just decode invalid graphs or isomorphic graphs. However, the analysis of multimodal nature in NAS should focus on the uniqueness of the resulting models, rather than just the uniqueness of the graphs themselves. Therefore, we first filter out invalid graphs and employ an iterative graph hashing algorithm [28] to identify the model-level unique solutions. Then, exhaustive enumeration can be conducted in this resulting search space to ensure we are considering the full range of possible unique models. The resulting normalized model-level unique solutions in the objective space are shown in Fig. 2 (a) and 2 (b).

From Fig. 2 (a) and 2 (b), we can see that in many regions multiple solutions are overlapped or very close to each other, which indicates that there may exist multimodal nature on




(a) 2-objective NAS ($\Omega$ = NB101) (b) 3-objective NAS ($\Omega$ = NB101) (c) 2-objective NAS ($\Omega$ = NB101) (d) 3-objective NAS ($\Omega$ = NB101)

Fig. 2: Obtained Unique solutions for the multi-objective NAS with NB101 search space.




(a) 2-objective NAS ($\Omega$ = NB101) (b) 3-objective NAS ($\Omega$ = NB101)

Fig. 3: The number of equivalent paired solutions with a distance measure that falls within the target minimum distance for NB101 search space.

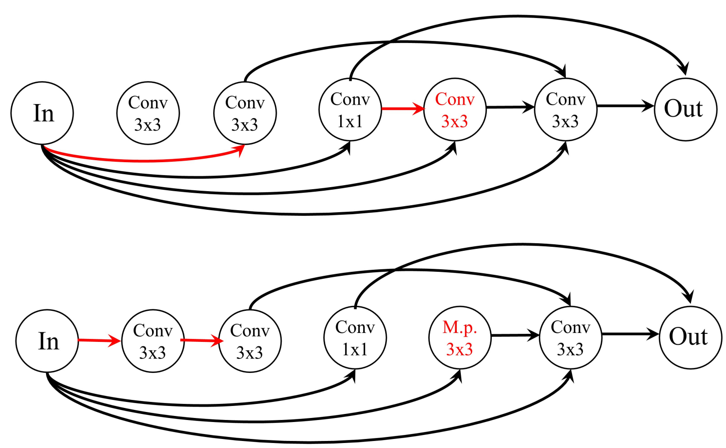



Fig. 4: The equivalent architectures that have exactly the same performance on each objective for 2-objective NAS with NB101 search space on the validation set.

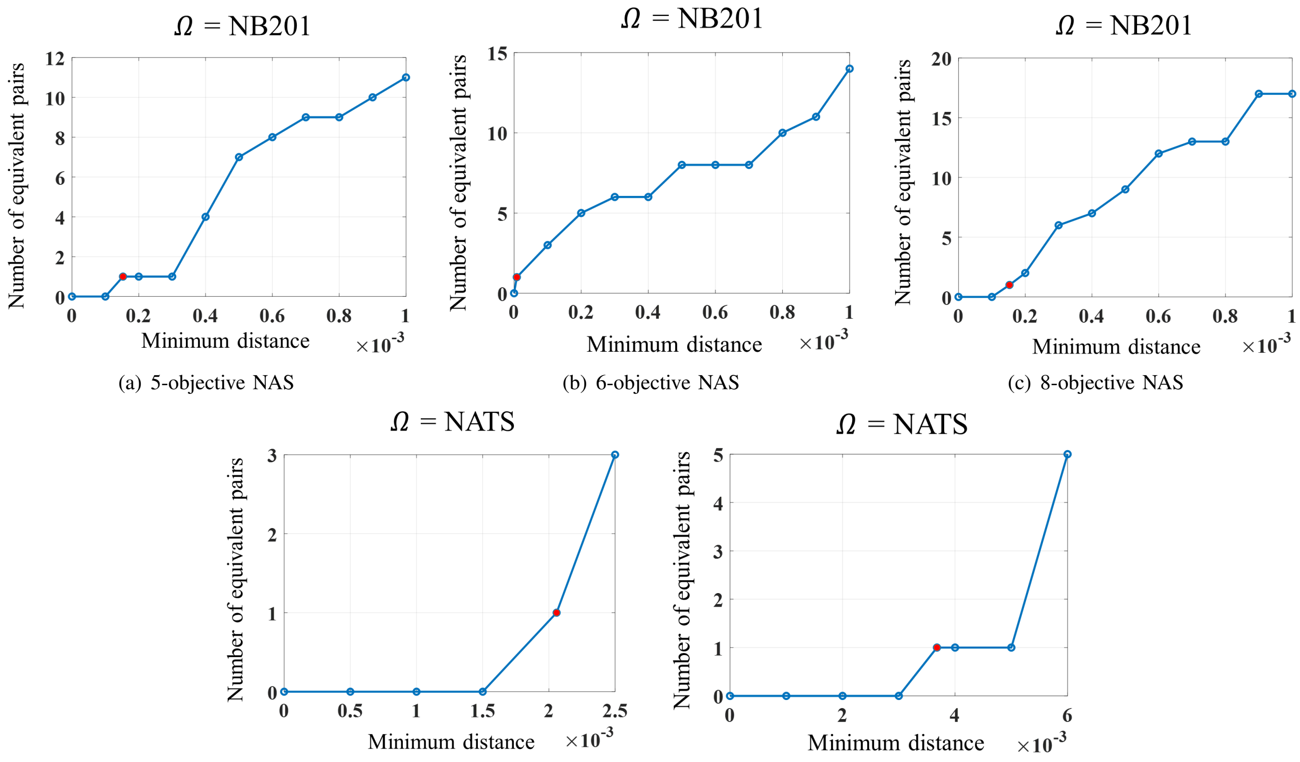
the multi-objective NAS with NB101 search space. Then, we perform non-dominated sorting by using an efficient non-dominated sort method described in [29]. The achieved Pareto fronts are shown in Fig. 2 (c) and 2 (d).

To examine whether there exists multimodal nature (i.e., whether there exist equivalent solutions in the Pareto front), we calculate the minimum distance of each solution to other solutions. As we have clarified in Section 2.2, two solutions with their distance within a pre-defined acceptable threshold value can be considered as equivalent solutions. Thus, as shown in Fig. 3, we calculate the number of paired solutions with a distance equal to or smaller than a specific range of threshold values $D_{TH}$, respectively (i.e., the number of paired solutions that are equivalent with a relaxation $\varepsilon \in D_{TH}$). The red dots are the starting point of minimum distance that indicates the existence of equivalent paired solutions.

In Fig. 3 (a), we can observe that in 2-objective NAS with NB101 search space, there exists one pair of equivalent solutions with a distance of zero between them. It means that these two different model architectures show the exact

TABLE II: Summary on the number of groups where different specific number of equivalent solutions map to the exact same point in the objective space for 2-objective NAS with NB101 search space.

| Group size | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| # Groups | 38876 | 15503 | 7517 | 3992 | 2228 | 1239 | 721 | 404 | 224 | 121 | 54 | 32 |
| Group size | 14 | 15 | 16 | 17 | 18 | 21 | 27 | 30 | 48 | 54 | 61 | 83 |
| # Groups | 17 | 6 | 5 | 3 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 1 |



(a) 5-objective NAS (b) 6-objective NAS (c) 8-objective NAS

(d) 3-objective NAS (e) 4-objective NAS

Fig. 5: The number of equivalent paired solutions with a distance measure that falls within the target minimum distance for NB201 and NATS search space.

same performance, which indicates the multimodal nature of multi-objective NAS with NB101 search space. The obtained equivalent paired solutions in the objective space are shown in red points in Fig. 2 (c). The model architectures of these two equivalent solutions are illustrated in Fig. 4. We can observe that the two model architectures in Fig. 4 are totally different, but they achieve the exact same performance on each objective.

However, in 3-objective NAS with NB101 search space as shown in Fig. 3 (b), no equivalent paired solutions with a distance of zero can be found (i.e., no multimodal nature). We can only find equivalent paired solutions when we increase the value of acceptable minimum distance to identify equivalent solutions (e.g., $\varepsilon \geq 1.0205 \times 10^{-4}$). The obtained equivalent paired solutions in the objective space are shown in Fig. 2 (d).

Considering that there exists multimodal nature on the 2-objective NAS with NB101 search space, we examine whether there exists multimodal nature in other regions of the search space besides Pareto front (i.e., including the dominated solutions) on this test problem. Our experimental results show that a group of equivalent solutions will map to the exact same point in the objective space. Also, there exist multiple such groups that consist of a different numbers of equivalent solutions (i.e., with different group sizes). In Table II, we record the number of such groups where a specific number of different solutions are equivalent with $\varepsilon = 0$.

In Table II, the group size denotes the number of equivalent solutions that map to the same point in the objective space, and # group denotes the number of such groups with the same group size. Table II shows that the 2-objective NAS exhibits very clear multimodal nature on the whole NB101 search space. Besides, we can find different numbers of equivalent solutions ranging from 2 to 83. It suggests that we should consider the multimodal nature when designing EMO algorithms to solve multi-objective NAS (e.g., NB101 search space), since many solutions have exactly the same objective values but represent completely different model architectures.

TABLE III: Summary of the experimental results on examining the multimodal nature of different multi-objective NAS benchmark problems using the validation set.

| Results on validation set | | | | | | |
|---|---|---|---|---|---|---|
| Problem | Ω | M | multimodal nature | # POS | # Equivalent pairs | Minimum Distance |
| C-10/MOP1 | NB101 | 2 | ✓ | 45 | 1 | 0 |
| C-10/MOP2 | NB101 | 3 | X | 60 | 0 | 1.0205e-4 |
| C-10/MOP3 | NATS | 3 | X | 380 | 0 | 2.1e-3 |
| C-10/MOP4 | NATS | 4 | X | 962 | 0 | 3.7e-3 |
| C-10/MOP5 | NB201 | 5 | X | 123 | 0 | 1.5281e-4 |
| C-10/MOP6 | NB201 | 6 | X | 105 | 0 | 8.1405e-6 |
| C-10/MOP7 | NB201 | 8 | X | 455 | 0 | 1.5281e-4 |

TABLE IV: Summary of the experimental results on examining the multimodal nature of different multi-objective NAS benchmark problems using the testing set.

| Results on testing set | | | | | | |
|---|---|---|---|---|---|---|
| Problem | Ω | M | multimodal nature | # POS | # Equivalent pairs | Minimum Distance |
| C-10/MOP1 | NB101 | 2 | ✓ | 44 | 1 | 0 |
| C-10/MOP2 | NB101 | 3 | X | 66 | 0 | 4.1317e-5 |
| C-10/MOP3 | NATS | 3 | X | 347 | 0 | 8.0601e-4 |
| C-10/MOP4 | NATS | 4 | X | 891 | 0 | 4.8e-3 |
| C-10/MOP5 | NB201 | 5 | X | 122 | 0 | 4.8176e-4 |
| C-10/MOP6 | NB201 | 6 | ✓ | 96 | 3 | 0 |
| C-10/MOP7 | NB201 | 8 | X | 434 | 0 | 3.9036e-4 |

### *B. Multimodal Nature on NB201 and NATS Search Space*

Similar to the previous subsection, we examine the multimodal nature of multi-objective NAS on benchmark problems with NB201 and NATS search spaces, i.e., C-10/MOP3-7 test problems. In Fig. 5, we illustrate the plots of the number of equivalent paired solutions over a range of minimum distances for different test problems.

Fig. 5 (a)-(c) show that on all three test problems, no equivalent paired solutions can be found with a distance of zero. It indicates that the multi-objective NAS benchmark problems with NB201 search space exhibit no multimodal nature. However, in Fig. 5 (b), we observe that the red point is very close to (almost overlapping with) the minimum distance of zero (i.e., $\varepsilon = 8.1405 \times 10^{-6}$). This suggests that there exist solutions that are very similar to each other in the objective space, which can be considered equivalent with a very small relaxation in the definition of equivalency (i.e., $\varepsilon$ could be set to a very small value). Actually, in the upcoming subsection, we will demonstrate that the performance evaluation of model architectures on the testing set for 6-objective NAS with the NB201 search space shows the evidence of multimodal nature. Based on our results, we can conclude that the 6-objective NAS with the NB201 search space has the potential to demonstrate multimodal nature.

For multi-objective NAS with the NATS search space, as shown in Fig. 5, we can observe that there exists no multimodal nature on these test problems. In addition, the red points are far away from the original point. It suggests that, even when we relax the definition of equivalency by setting a small threshold value, identifying the multimodal nature of multi-objective NAS with the NATS search space can be challenging (i.e., multimodal nature can only be observable when using a larger value of $\varepsilon$).

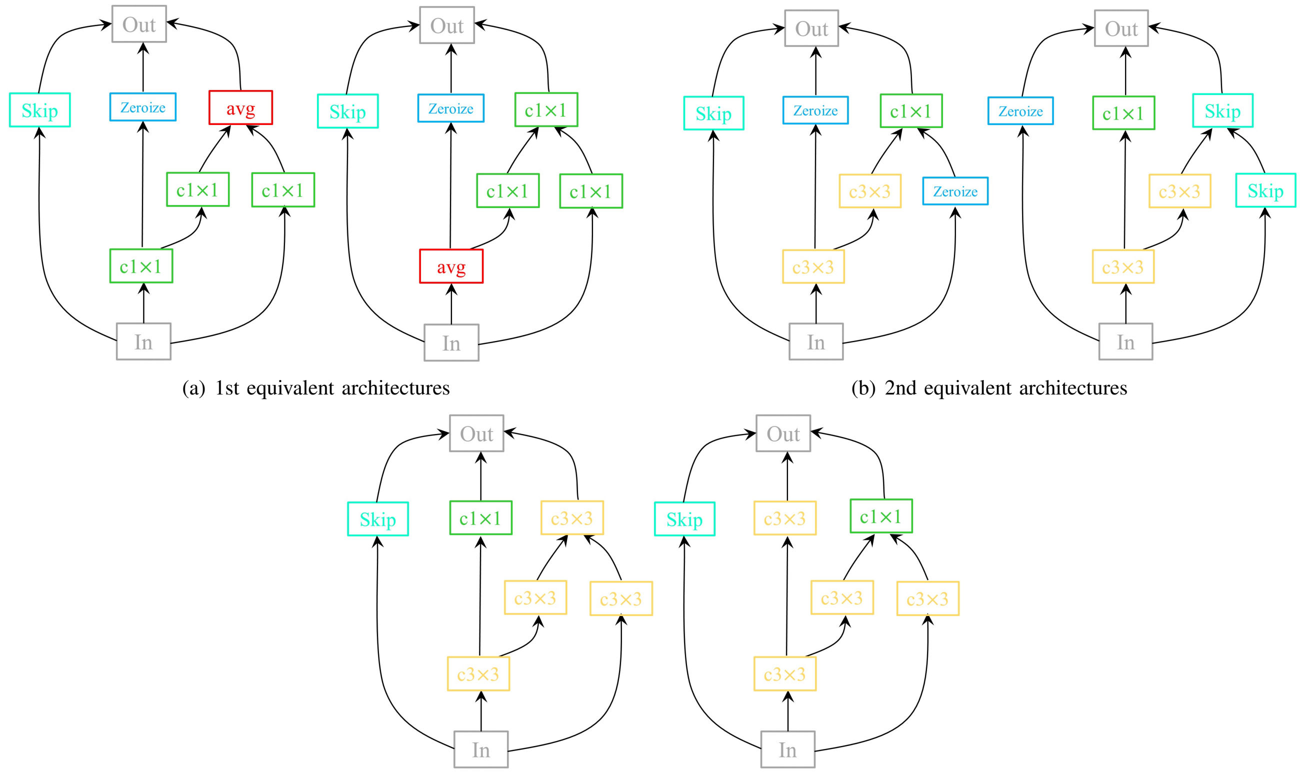
### *C. Multimodal Nature of Multi-objective NAS on Testing Set*

In the previous subsections, we discussed the multimodal nature of various multi-objective NAS benchmark problems. However, the experimental results we presented were based on the evaluation of model architectures by using the validation set, where the first objective is associated with the validation prediction error. In this subsection, we will examine the multimodal nature of multi-objective NAS by evaluating the performance of model architectures on the testing set. It can help us better understand the generalization of the multimodal nature of model architectures to novel data.

We summarize the experimental results on the examination of the multimodal nature of various multi-objective NAS benchmark problems on the validation set and testing set in Tables III and IV.

In Tables III and IV, we introduce several key metrics that help us better understand the examination of multimodal nature of different multi-objective NAS problems. Specifically, we use $M$ to represent the number of objectives, and #POS to represent the number of Pareto optimal solutions on the Pareto front that can be obtained through exhaustive enumeration. Additionally, #Equivalent pairs denotes the number of paired solutions that have the exact same performance on each objective. The minimum distance represents the smallest distance between two paired solutions that we can find. This quantity helps to determine the smallest threshold value of $\varepsilon$, which defines a relaxation in equivalency that allows for finding equivalent solutions.

We can observe that 2-objective NAS with the NB101 search space (C-10/MOP1) exhibit multimodal nature on both




(a) 1st equivalent architectures

(b) 2nd equivalent architectures

(c) 3rd equivalent architectures

Fig. 6: Three equivalent architectures that have the exact same performance on each objective for 6-objective NAS with the NB201 search space on the testing set.

validation and testing set. For 6-objective NAS with the NB201 search space (C-10/MOP6), it exhibits multimodal nature on the testing set. We illustrated the model architectures of the found equivalent pairs in Fig. 6. On the validation set, no two solutions have the exact same objective values for C-10/MOP6. However, as we have explained, its minimum distance (i.e., $8.1405 \times 10^{-6}$) is very small on the validation set. It means that we can easily find its multimodal nature when slightly increasing the equivalency relaxation $\varepsilon$ to a very small value.

On other multi-objective NAS problems with the NB201 search space (i.e., C-10/MOP5 and 7), we find no multimodal nature when considering the equivalency based on $\varepsilon = 0$. However, they can exhibit multimodal natures when we adjust the equivalency relaxation $\varepsilon$ to a slightly larger value. Multi-objective NAS problems with the NATS search space also exhibit no multimodal nature when considering $\varepsilon = 0$. Besides, their minimum distances between any two solutions we can find are large. Thus, it seems that it is challenging to find multimodal nature on the multi-objective NAS with the NATS search space.

## V. Conclusion

In this paper, we examined the multimodal nature of multi-objective neural architecture search. More specifically, we searched for different architectures which have (almost) the same performance in the objective space. Our experimental results showed that on some multi-objective NAS benchmark problems, there exists multimodal nature. Some other problems exhibit no multimodal nature when we consider the exact equivalency (i.e., $\varepsilon = 0$). However, these problems can exhibit multimodal nature if we allow for a small relaxation on defining the equivalency. There are still a number of problems that exhibit no multimodal nature even if we consider a small relaxation on equivalency. Interestingly, we observed that on the other objective space besides the Pareto front, there exist a large number of equivalent solutions that achieve the exact same performance on each objective for many problems. All these results suggest that it is worth paying attention to the multimodal nature when we design EMO algorithms to solve multi-objective NAS problems.

## References

[1] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[2] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1–9.

[3] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.

[4] T. Brown, B. Mann, N. Ryder, M. Subbiah, J. D. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell *et al.*, "Language models are few-shot learners," *Advances in neural information processing systems*, vol. 33, pp. 1877–1901, 2020.
[5] B. Zoph and Q. V. Le, "Neural architecture search with reinforcement learning," *arXiv preprint arXiv:1611.01578*, 2016.
[6] M. Tan, B. Chen, R. Pang, V. Vasudevan, M. Sandler, A. Howard, and Q. V. Le, "Mnasnet: Platform-aware neural architecture search for mobile," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 2820–2828.
[7] Z. Lu, R. Cheng, S. Huang, H. Zhang, C. Qiu, and F. Yang, "Surrogate-assisted multiobjective neural architecture search for real-time semantic segmentation," *IEEE Transactions on Artificial Intelligence*, 2022.
[8] K. Miettinen, *Nonlinear multiobjective optimization*. Springer Science & Business Media, 1999, vol. 12.
[9] L. Bottou, F. E. Curtis, and J. Nocedal, "Optimization methods for large-scale machine learning," *SIAM review*, vol. 60, no. 2, pp. 223–311, 2018.
[10] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: Nsga-ii," *IEEE transactions on evolutionary computation*, vol. 6, no. 2, pp. 182–197, 2002.
[11] Q. Zhang and H. Li, "MOEA/D: A multiobjective evolutionary algorithm based on decomposition," *IEEE Transactions on evolutionary computation*, vol. 11, no. 6, pp. 712–731, 2007.
[12] N. Beume, B. Naujoks, and M. Emmerich, "Sms-emoa: Multiobjective selection based on dominated hypervolume," *European Journal of Operational Research*, vol. 181, no. 3, pp. 1653–1669, 2007.
[13] E. Real, A. Aggarwal, Y. Huang, and Q. V. Le, "Regularized evolution for image classifier architecture search," in *Proceedings of the aaai conference on artificial intelligence*, vol. 33, no. 01, 2019, pp. 4780–4789.
[14] Y. Sun, B. Xue, M. Zhang, and G. G. Yen, "Evolving deep convolutional neural networks for image classification," *IEEE Transactions on Evolutionary Computation*, vol. 24, no. 2, pp. 394–407, 2019.
[15] T. Zhang, C. Lei, Z. Zhang, X.-B. Meng, and C. P. Chen, "As-nas: Adaptive scalable neural architecture search with reinforced evolutionary algorithm for deep learning," *IEEE Transactions on Evolutionary Computation*, vol. 25, no. 5, pp. 830–841, 2021.
[16] Z. Lu, R. Cheng, Y. Jin, K. C. Tan, and K. Deb, "Neural architecture search as multiobjective optimization benchmarks: Problem formulation and performance assessment," *IEEE Transactions on Evolutionary Computation*, 2023.
[17] K. Deb and S. Tiwari, "Omni-optimizer: A generic evolutionary algorithm for single and multi-objective optimization," *European Journal of Operational Research*, vol. 185, no. 3, pp. 1062–1087, 2008.
[18] R. Tanabe and H. Ishibuchi, "A review of evolutionary multimodal multiobjective optimization," *IEEE Transactions on Evolutionary Computation*, vol. 24, no. 1, pp. 193–200, 2019.
[19] C. Ying, A. Klein, E. Christiansen, E. Real, K. Murphy, and F. Hutter, "Nas-bench-101: Towards reproducible neural architecture search," in *International Conference on Machine Learning*. PMLR, 2019, pp. 7105–7114.
[20] X. Dong and Y. Yang, "Nas-bench-201: Extending the scope of reproducible neural architecture search," *arXiv preprint arXiv:2001.00326*, 2020.
[21] X. Dong, L. Liu, K. Musial, and B. Gabrys, "Nats-bench: Benchmarking nas algorithms for architecture topology and size," *IEEE transactions on pattern analysis and machine intelligence*, vol. 44, no. 7, pp. 3634–3646, 2021.
[22] B. Zoph, V. Vasudevan, J. Shlens, and Q. V. Le, "Learning transferable architectures for scalable image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 8697–8710.
[23] X. Dai, P. Zhang, B. Wu, H. Yin, F. Sun, Y. Wang, M. Dukhan, Y. Hu, Y. Wu, Y. Jia *et al.*, "Chamnet: Towards efficient network design through platform-aware model adaptation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 11 398–11 407.
[24] J. Michalek, R. Choudhary, and P. Papalambros, "Architectural layout design optimization," *Engineering optimization*, vol. 34, no. 5, pp. 461–484, 2002.
[25] A. Jaszkiewicz, "On the performance of multiple-objective genetic local search on the 0/1 knapsack problem-a comparative experiment," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 4, pp. 402–412, 2002.
[26] F. Kudo, T. Yoshikawa, and T. Furuhashi, "A study on analysis of design variables in pareto solutions for conceptual design optimization problem of hybrid rocket engine," in *2011 IEEE Congress of Evolutionary Computation (CEC)*. IEEE, 2011, pp. 2558–2562.
[27] Y.-H. Chen, T. Krishna, J. S. Emer, and V. Sze, "Eyeriss: An energy-efficient reconfigurable accelerator for deep convolutional neural networks," *IEEE journal of solid-state circuits*, vol. 52, no. 1, pp. 127–138, 2016.
[28] C. Ying, "Enumerating unique computational graphs via an iterative graph invariant," *arXiv preprint arXiv:1902.06192*, 2019.
[29] X. Zhang, Y. Tian, R. Cheng, and Y. Jin, "An efficient approach to nondominated sorting for evolutionary multiobjective optimization," *IEEE Transactions on Evolutionary Computation*, vol. 19, no. 2, pp. 201–213, 2014.