Technische Universität München
School of Computation, Information and Technology
Department of Informatics

Exposé — Bachelor's Thesis

Adaptation Techniques for using NAS Methods in the FL Setting

Author:	Max Coetzee
Supervisor:	M.Sc. Nick Henze
Advisor:	Dr.-Ing. Niclas Kannengießer
Date:	tbd.

Contents

1 Problem Statement

In the wake of the deep learning revolution, deep learning is being applied to an increasingly diverse set of tasks. Applying deep learning traditionally involves a team of deep learning and domain experts who tailor a neural network architecture to a specific task through a lengthy process of trial and error. To automate this process, researchers invented *Neural Architecture Search* (NAS) [3] methods. NAS methods employ diverse strategies to automatically search for a neural network architecture for a given deep learning task within an architecture search space.

Independently, but in parallel to NAS, researchers developed a machine learning approach called *Federated Learning* (FL) [8], which enables training an ML model on data distributed across a set of *clients* without sharing their data. The model is disseminated from the server to clients and trained in *communication rounds* by clients using their local data, after which model weight updates are sent to a *server* for aggregation and re-dissemination in the next round. FL enhances the privacy of client data and enables training even when clients' data is too large to transfer effectively to a central training site or cannot be shared due to regulatory or privacy constraints.

When practitioners[1] make use of deep learning in FL, they typically use a predefined architecture. However, clients' data distributions and hardware capabilities are highly heterogeneous in FL. Since practitioners cannot directly observe these client characteristics [6], it is hard to choose a single architecture that generalises well across all clients and that meets inference latency constraints on every client. Using NAS methods in FL provides an alternative: the server can coordinate an architecture search across clients, guided by performance feedback from candidate architectures evaluated directly on clients — making clients' characteristics observable to the NAS method. Thus, the NAS method can better guide the architecture search than a practitioner could.

Despite the benefit, using NAS in FL is not straightforward. Most NAS methods were developed for a centralised setting and rely on assumptions that often do not hold for FL. These assumption discrepancies create several *challenges* for using NAS in FL, and practitioners need to *adapt* the NAS method or the FL pipeline to overcome them — creating a *FedNAS* [6] method. For example, in a centralised setting, practitioners can expect a NAS method to run on worker nodes connected via low-latency, high-bandwidth links. However, in FL, a NAS method may run on clients connected via an unreliable, high-latency, low-bandwidth network. This creates a challenge when transferring the weights of large candidate models between clients and the server to coordinate an architecture search. Practitioners have developed various *adaptation techniques* to overcome this challenge, such as only transferring an encoding of a

[1] In this thesis, *practitioners* refers to both users and developers of FL and FedNAS methods.

candidate architecture for clients to evaluate [9].

Practitioners stand to benefit from comparing and re-using existing adaptation techniques when creating FedNAS methods. To this end, practitioners need to comb the increasingly large body of FedNAS literature for suitable adaptation techniques — i.e. techniques that overcome the challenges relevant to their targeted FL system characteristics. The relevance of a challenge depends on the extent to which the targeted FL system characteristics violate centralised NAS assumptions. FL system characteristics include the average hardware capabilities of clients, the average network latency of clients, the number of participating clients, etc. Additionally, practitioners need to weigh which adaptation techniques to use to overcome their relevant set of challenges, since an adaptation technique for overcoming one challenge may inadvertently make it harder to overcome another.

However, comparing and re-using adaptation techniques is difficult. Prior surveys of the FedNAS literature categorise FedNAS methods as a whole rather than analyse the individual adaptation techniques they use. [10] categorises FedNAS methods into offline versus online architecture search and single-objective versus multi-objective methods. [7] gives a high-level overview of the FedNAS landscape at the time of its publication as part of a larger survey into combining NAS and Hyperparameter Optimisation. [5] investigates how multi-objective optimisation can be integrated into FL in general and includes only parts that discuss how this is done specifically for FedNAS methods. All of the surveys analyse only a small fraction of the entire FedNAS literature. This means knowledge on adaptation techniques and potential challenges when adapting NAS methods to FL remains fragmented, leading to our research question:

What challenges arise when adapting centralised NAS methods to FL, and how do adaptation techniques in the literature overcome these challenges?

To tackle our research question, we conduct an extensive systematic literature review. We collect an initial set of relevant literature with a Scopus [2] search and extend it with a snowball search. We first document the challenges that arise when adapting a centralised NAS method to FL, as well as the influence of FL system characteristics on the relevance of these challenges, in preparation for the next step. We then extract adaptation techniques from the literature through open coding and iteratively refine them to synthesise a coherent set of mutually exclusive and collectively exhaustive adaptation techniques. Next, we discuss how each adaptation technique works towards, against, or does not affect overcoming each challenge. Finally, we summarise our results in a table that allows practitioners to quickly find adaptation techniques to overcome a specific challenge.

Our review aims to synthesise adaptation techniques from the literature to support their re-use and comparison in practice. We make the following three contributions. First, by providing a consolidated overview of the challenges of adapting centralised NAS methods to FL, practitioners can grasp which issues they can expect to encounter at a glance. Second, by describing how FL system characteristics influence the relevance of challenges, we enable practitioners to focus on the challenges relevant to them. Third, by extracting adaptation techniques and highlighting the challenges they address, we enable practitioners to compare existing adaptation techniques for adapting centralised NAS methods to FL for their specific set of challenges.

2 Objectives

The primary objective of this thesis is to provide an overview of techniques used to adapt centralised NAS to FL, spawning the following sub-objectives:

1. Catalogue the assumptions that hold for centralised NAS, but are violated in FedNAS.
2. Describe the challenges arising from the mismatch in assumptions between centralised NAS and FedNAS.
3. Catalogue the adaptation techniques used to overcome these challenges.
4. Describe the influence of FL system characteristics on the relevance of challenges.

3 Explanation of Terms

3.1 Neural Architecture Search

Neural Architecture Search (NAS) enables the automatic search for neural network architecture as opposed to manual experimentation. A NAS method consists of a *search space* for potential candidate architectures, a *search strategy* that guides the architecture search, and a *performance estimation strategy* used by the search strategy to decide which candidates to favour or explore next.

3.2 Federated Learning

FL is a machine learning approach in which multiple *clients* collaboratively train a shared model while keeping their training data local. A central *server* coordinates *communication rounds* with the clients, wherein each client trains the model for several epochs locally and finally sends gradient updates to the server for aggregation into the shared model. After aggregation, the server sends the updates to clients for the next round of training.

3.3 FedNAS method

A FedNAS method is a NAS method that runs on an FL system.

3.3.1 Adapted vs Novel FedNAS method

We consider a FedNAS method to be an adaptation of a centralised NAS method if a large share of its three components (search space, search strategy, performance estimation strategy) are adaptations of components of an existing centralised NAS method. If this is not the case, i.e., a large share of the three components is a novel creation, we consider the FedNAS method to be novel. Using this definition, we observe that novel FedNAS methods are scarce in the literature.

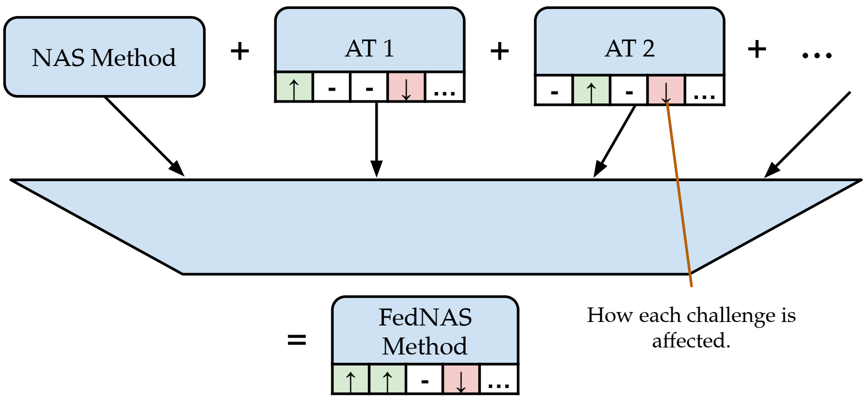

Figure 1: Conceptual overview of an adapted FedNAS method.

3.4 Adaptation technique

An adaptation technique is a substantial modification to the FL pipeline or a centralised NAS method that:

1. allows running a centralised NAS method on an FL system, and
2. is explicitly motivated by the need to overcome a challenge that arises as a result of an assumption mismatch between a centralised and an FL setting.

For example, The FedNAS method *FedorAS* [1] makes use of the centralised NAS method *SPOS* [4], but runs into the following challenge caused by the fact that centralised NAS can assume worker nodes are connected via low latency, high-bandwidth links, whereas clients in FL are not: sending the parameters of the entire supernet used in SPOS to each client would take an infeasible amount of time. Instead, *FedorAS* adapts *SPOS* such that only a small, sampled subspace of the supernet is sent to each client for training and evaluation per communication round.

4 Research Approach

4.1 Data Collection

We obtain the set of literature relevant to our review by searching the abstract, title and keywords of literature in Scopus [2] with the search query *"federated learning" AND "neural architecture search"* and include literature that describes a FedNAS method as defined in 3.3. FedNAS method. We extend this initial set of literature by a snowball search. We exclude literature that does not provide sufficient methodological detail to extract adaptation techniques.

4.2 Data Analysis

We employ a qualitative research approach, specifically a systematic literature review. In our review, we develop a conceptual model that links violated centralised NAS assumptions to resulting challenges for adapting a centralised NAS method to FL, and the adaptation techniques used to overcome these challenges as follows:

1. **Challenges arising from Assumption Discrepancies:** We first document challenges that arise from using centralised NAS methods in FL, because of assumptions that hold for centralised NAS, but not for FL.
2. **Influence of FL System Parameters on Challenges:** We then classify the influence of FL system characteristics on the relevance of challenges as either "increasing" or "decreasing" the relevance of each challenge.
3. **Adaptation Technique Conceptualisation:** Next, we extract adaptation techniques from the literature in two steps.
   3.1 First, we perform open coding on each FedNAS method to extract unrefined adaptation techniques. Any modification to a NAS method explicitly motivated by the federated setting is initially coded as a single unrefined adaptation technique.

3.2 Then, we iteratively refine and merge the unrefined adaptation techniques by merging unrefined techniques with conceptually highly similar mechanisms into a single representative adaptation technique. Thus, we obtain a coherent set of mutually exclusive and collectively exhaustive adaptation techniques.

4. **Discuss FedNAS Challenges for Adaptation Techniques:** Afterwards, we discuss how each adaptation technique works towards, against, or does not affect overcoming each of the FedNAS challenges.

5. **Table Overview:** Finally, we create a table that contains a coded vector of effects over the challenges for each adaptation technique based on the prior discussion.

Practitioners can use our classification from 2. to first identify challenges relevant to their targeted FL system parameters and then refer to our table from 5. to gain an overview at a glance of which adaptation techniques could be helpful to them.

5 Structure

1. **Introduction** *(3 pages)*

   What is the motivation behind this thesis? What is the relevance of this thesis? What problem does this thesis try to solve? What kind of approach does this thesis take to solve that problem? How is the thesis structured?

2. **Background** *(6 pages)*

   What background knowledge is required to understand this thesis?

   2.1 Neural Architecture Search (2.5 pages)

   What is NAS used for, and how does it work? What is a NAS method? What are the origins of NAS? What environment are NAS methods typically developed in?

   2.2 Federated Learning (2.5 pages)

   What is Federated Learning? What was it designed for and how does it work?

   2.3 Federated Neural Architecture Search (1 page)

   How do FedNAS methods relate to NAS and FL? What are the origins of FedNAS methods, and how have they been developed over the last couple of years?

3. **Method** *(5 pages)*

   How does this thesis aim to achieve the objectives described in Section 2?

   3.1 Reviewed Literature *(2 pages)* Will include visualisations that break down the included literature for the reader.

   Which literature is included in the review?

   3.2 Methodology *(3 pages)*

   How is the literature analysed? Which methodology is employed and how?

4. **Challenges with Adapting Centralised NAS methods to FL** (10 pages):

   4.1 Assumption Discrepancies between NAS and FedNAS *(3 pages)*

   How do centralised NAS and FedNAS differ? How does this difference manifest in different assumptions?

   4.2 Adaptation Challenges *(4 pages)*

   How do FL system characteristics influence challenges faced in adapting centralised NAS methods to FL?

   4.3 Influence of FL system characteristics on Challenges *(3 pages)*

   Which FL system characteristics influence the degree to which a challenge is relevant? Which FL system characteristics are relevant and which ones are not?

5. **Adaptation Techniques** *(25 pages)*

   What adaptation techniques are described in the literature? What can we learn from these adaptation techniques that we can use for adapting NAS methods to FL in the future? For which FL system characteristics was each adaptation technique developed? How do adaptation techniques work towards, against, or do not affect overcoming each of the challenges relevant to that FL system characteristics?

   5.1 Adaptation Technique 1 *(1 page)*

   Description of the refined adaptation technique, the FedNAS methods that use it and a discussion on how it overcomes challenges.

   5.2 Adaptation Technique 2 *(1 page)*

   5.3 ...

   5.4 Adaptation Technique 20 *(1 page)*

5.5 Overview *(5 pages)* A large overview table that makes it easy to see which adaptation techniques are beneficial for overcoming which challenges at a glance.

6. **Discussion** *(2 pages)*

   What do our findings indicate for the future of adapting centralised NAS methods to FL? How can researchers and practitioners use these contributions when adapting centralised NAS methods to targeted FL system characteristics?

7. **Conclusion** *(1 page)*

   What is the answer to the thesis's research question, and what are the main contributions?

52 pages in total.

6 Expected Results

We expect to produce the following:

1. A consolidated catalogue of challenges that arise when adapting centralised NAS methods to FL.
2. Classification of the influence of FL system characteristics on the relevance of a challenge.
3. A catalogue of adaptation techniques extracted from FedNAS methods with an in-depth discussion of the challenges they overcome.
4. A tabular overview of adaptation techniques that shows which challenges they overcome at a glance.

References

[1] L. Dudziak, S. Laskaridis, and J. Fernandez-Marques. *FedorAS: Federated architecture search under system heterogeneity*. 2022. arXiv: `2206.11239 [cs.LG]`.

[2] Elsevier. *Scopus*. Abstract and citation database. Accessed 2026-01-17. 2026.

[3] T. Elsken, J. H. Metzen, and F. Hutter. "Neural architecture search: A survey." In: *Journal of Machine Learning Research* 20.55 (2019), pp. 1–21.

[4] Z. Guo, X. Zhang, H. Mu, W. Heng, Z. Liu, Y. Wei, and J. Sun. “Single path one-shot neural architecture search with uniform sampling.” In: *Computer vision – ECCV 2020*. Ed. by A. Vedaldi, H. Bischof, T. Brox, and J.-M. Frahm. Cham: Springer International Publishing, 2020, pp. 544–560. ISBN: 978-3-030-58517-4.

[5] M. Hartmann, G. Danoy, and P. Bouvry. *Multi-objective methods in Federated Learning: A survey and taxonomy*. 2025. arXiv: 2502.03108 [cs.LG].

[6] C. He, M. Annavaram, and S. Avestimehr. *Towards Non-I.I.D. and invisible data with FedNAS: Federated deep learning via neural architecture search*. 2021. arXiv: 2004.08546 [cs.LG].

[7] S. Khan, A. Rizwan, A. N. Khan, M. Ali, R. Ahmed, and D. H. Kim. “A multi-perspective revisit to the optimization methods of Neural Architecture Search and Hyper-parameter optimization for non-federated and federated learning environments.” In: *Computers and Electrical Engineering* 110 (2023), p. 108867. ISSN: 0045-7906. DOI: https://doi.org/10.1016/j.compeleceng.2023.108867.

[8] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas. “Communication-efficient learning of deep networks from decentralized data.” In: *Proceedings of the 20th international conference on artificial intelligence and statistics*. Ed. by S. Aarti and Z. Jerry. Vol. 54. Proceedings of machine learning research. PMLR, Apr. 2017, pp. 1273–1282.

[9] R. E. Shawi. “AgingFedNAS: Aging evolution federated deep learning for architecture and hyperparameter search.” In: *Advances in knowledge discovery and data mining*. Ed. by X. Wu, M. Spiliopoulou, C. Wang, V. Kumar, L. Cao, Y. Wu, Y. Yao, and Z. Wu. Singapore: Springer Nature Singapore, 2025, pp. 107–119. ISBN: 978-981-96-8173-0.

[10] H. Zhu, H. Zhang, and Y. Jin. “From federated learning to federated neural architecture search: a survey.” In: *Complex and Intelligent Systems* 7.2 (Apr. 2021), pp. 639–657. ISSN: 2198-6053. DOI: 10.1007/s40747-020-00247-z.