Differentiable Neural Architecture Search through a Hypernetwork For Image Recognition

Built with

- Python
- PyTorch

Background and Description

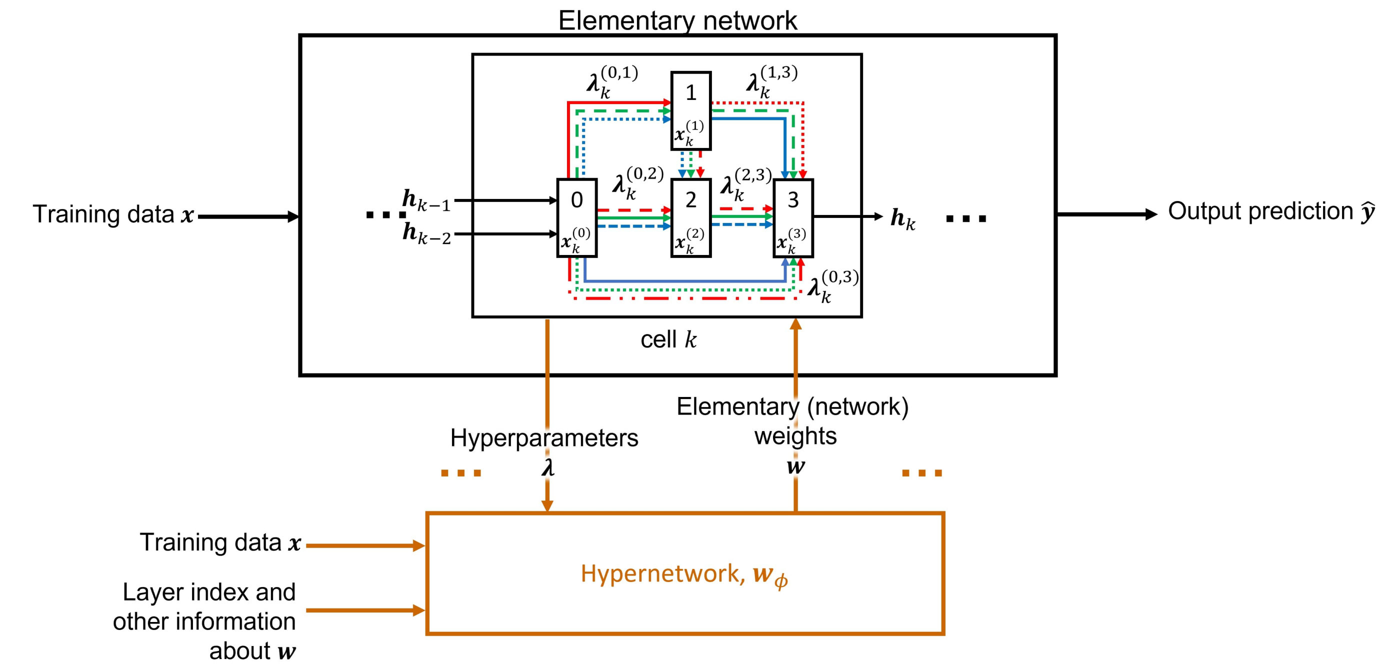
Neural Architecture Search (NAS) automates the search for optimal neural network architectures, a process traditionally requiring human expertise and trial-and-error. NAS employs machine learning algorithms to explore a predefined search space, evaluating and comparing architectures based on task performance. It utilizes strategies like reinforcement learning, evolutionary algorithms, or gradient-based optimization to efficiently identify architectures with superior performance. By automating architecture design, NAS accelerates the development of advanced neural networks, propelling AI research and applications.

DARTS (Differentiable Architecture Search) is a gradient-based approach to (NAS). It uses a continuous relaxation of the architecture space, allowing for efficient optimization through backpropagation. By introducing differentiable variables and leveraging the softmax relaxation, DARTS enables joint optimization of architecture and model weights. This approach has proven effective in discovering competitive neural architectures while reducing computational costs.

DARTS uses unrolling method to approach the bilevel optimization search problem to approximate hypergradient. This project proposes developing a DARTS variant by replacing the unrolling method with a hypernetwork to eliminate the approximation loss and further improve its performance in searching architectures.

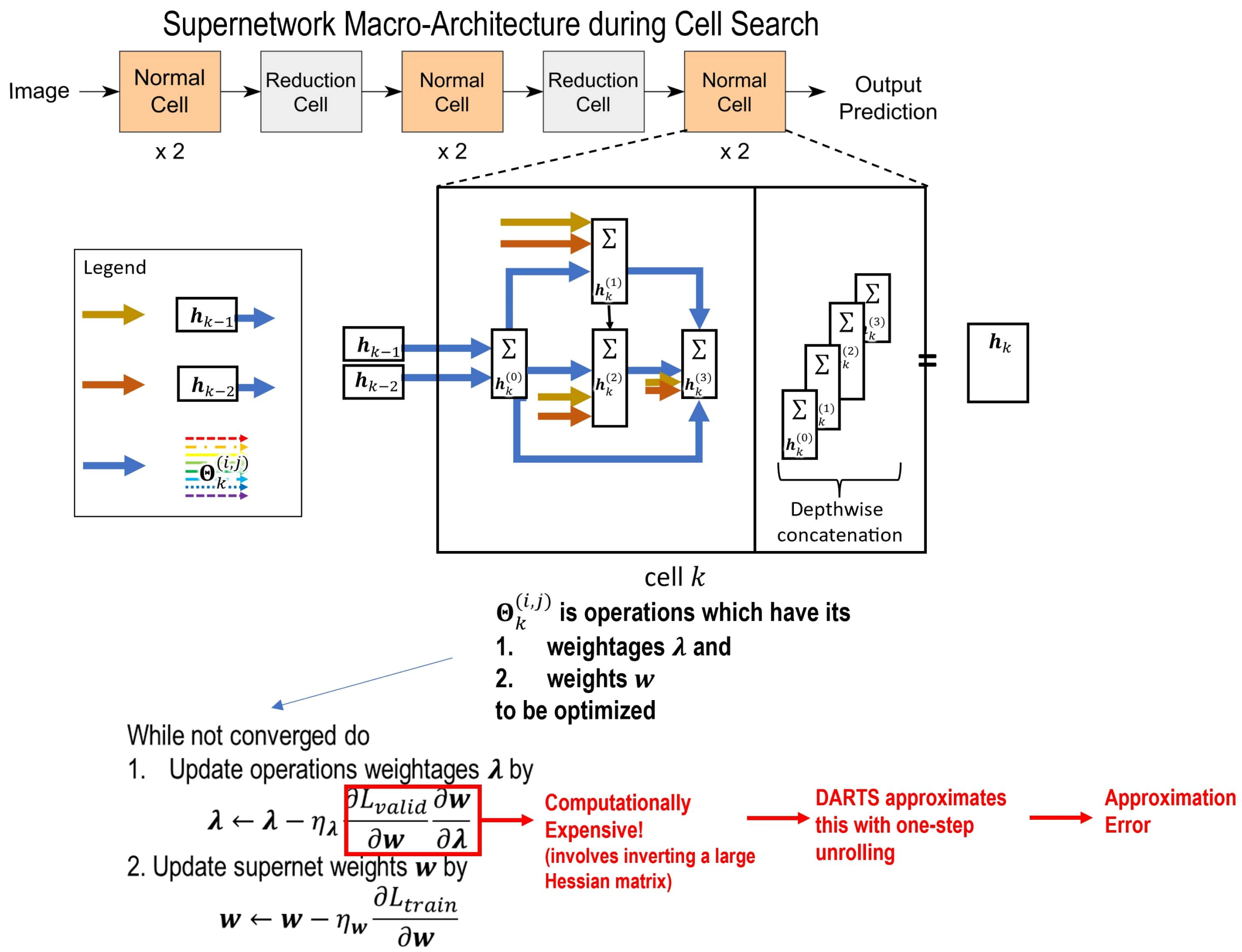
Reference Code: https://github.com/khanrc/pt.darts/tree/master (DARTS implementation)

Overview of HyperDARTS that integrates hypernetworks into DARTS

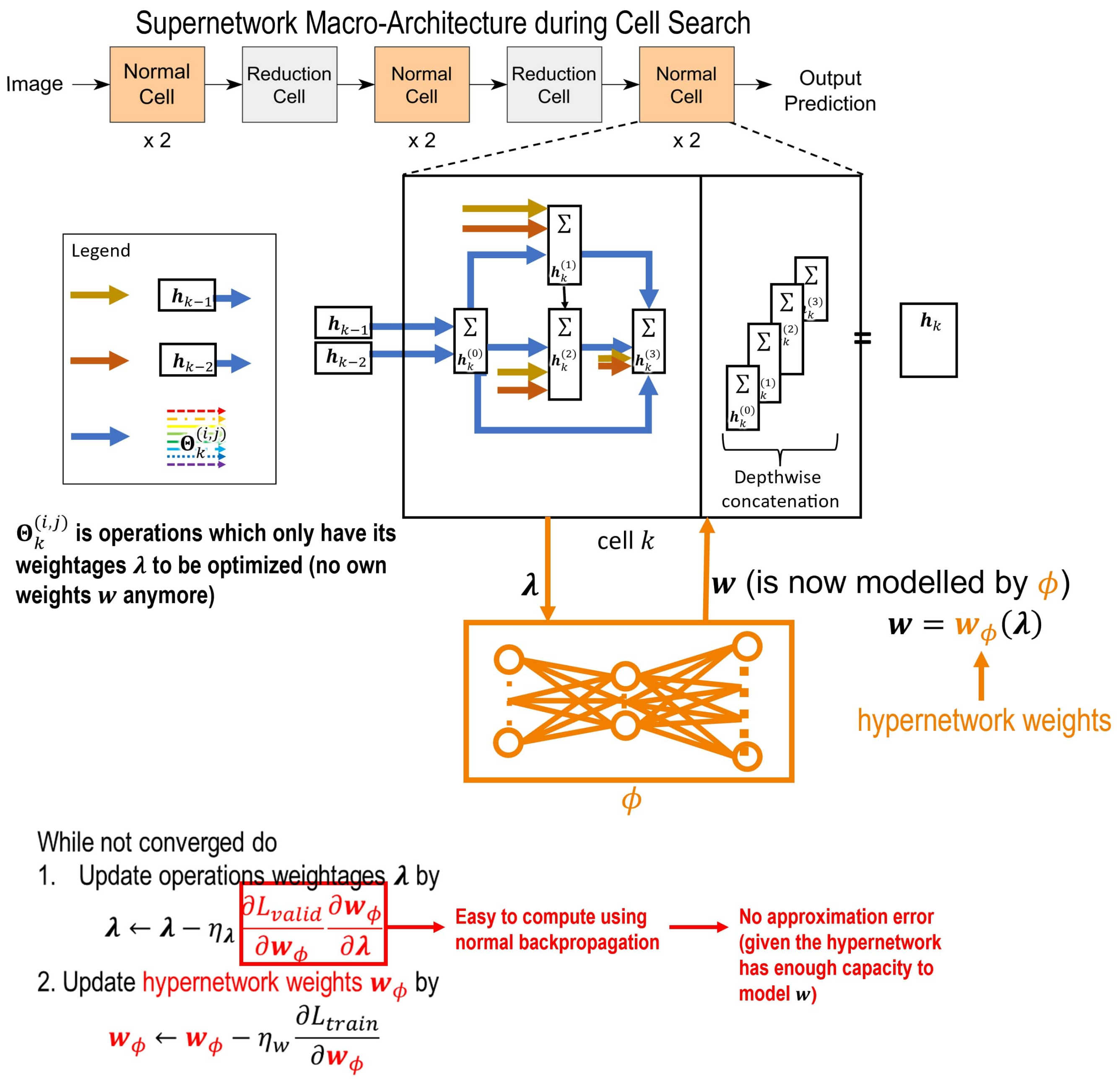


DARTS vs. HypereDARTS

Darts



HyperDARTS



Poster

MONASH University

PROJECT ID
2022S2-1148

MONASH ENGINEERING

ELECTRICAL AND COMPUTER SYSTEMS ENGINEERING
FINAL YEAR PROJECT

Mah Yuen Yee
Dr. Mohamed Hisham Jaward, Dr. Liang Shiuan-Ni

Differentiable Neural Architecture Search through Hypernetworks for Image Recognition

INTRODUCTION

Deep Learning has demonstrated exceptional performance in solving perceptual tasks, particularly in Computer Vision, e.g., image recognition and segmentation. Traditionally, neural network architectures require manual design, heavily relying on human expertise and trial-and-error, leading to limited exploration and difficult task-specific adaptation.

Neural Architecture Search (NAS) automates the search for optimal neural network architectures, by using a search strategy to explore a predefined architecture search space, evaluating and comparing candidate architectures based on task performance.

DARTS (Differentiable Architecture Search) [1] is a pioneering gradient-based NAS model, using a continuous relaxation of the architecture space allowing for efficient optimization through backpropagation. DARTS outperforms most discrete search strategy-based models, effectively discovering competitive neural architectures while reducing computational costs. DARTS tackles the search problem with bilevel optimization, jointly optimizing architecture parameters $\boldsymbol{\lambda}$ (hyperparameters) and candidate network weights $\boldsymbol{w}$ using gradient descent. In its search strategy, DARTS approximates the computationally expensive hyperparameter gradient (hypergradient) through a one-step unrolling scheme, incurring approximation error.

AIMS AND OBJECTIVES

Aim:

This project proposes to **develop a novel DARTS-based NAS model**, replacing the approximation-based search strategy with hypernetworks ϕ to improve the performance in finding optimal architectures. The optimality of discovered architectures is measured by their performance on image recognition tasks.

Objectives:

- To develop **hyperDARTS** (hypernetwork-DARTS), a new DARTS-based NAS model that uses hypernetworks as the search strategy to **find optimal architectures with higher overall performance in image recognition tasks than DARTS.**
- To evaluate the image recognition capability (test accuracy) of architectures discovered by hyperDARTS on CIFAR-10, the dataset used for searching.
- To evaluate the transferability (generalization capability) of architectures discovered by hyperDARTS, quantified by their test accuracy on CIFAR-100, a more challenging dataset not used for searching.

METHODOLOGY AND METHODS

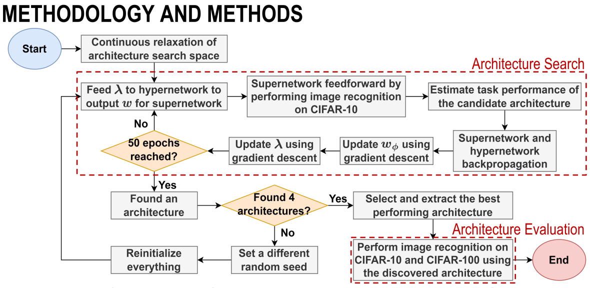


Figure 1. Flow of a hyperDARTS model with 2 main phases: Architecture Search and Architecture Evaluation.

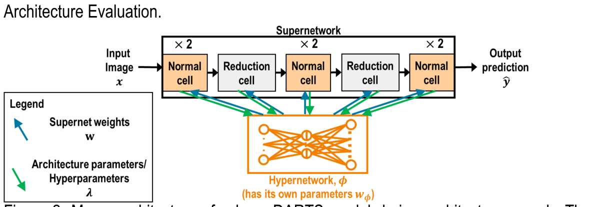


Figure 2. Macro-architecture of a hyperDARTS model during architecture search. The supernetwork contains all the candidate architectures. All candidate architectures are built using only two types of cells: Normal (preserve spatiality) and Reduction (reduce spatiality). Therefore, architecture search is effectively cell search.

Table 1: The two hyperDARTS models developed: hyperDARTS-1 and hyperDARTS-2. # denotes the number of or the size of.

	HyperDARTS-1	HyperDARTS-2
Hypernetwork used	Linear Regression	2-layer multilevel perceptron (MLP) - 1 hidden layer + non-linear (sigmoid) activation - #nodes in hidden layer = ½ #$\boldsymbol{\lambda}$
Model Capacity	Lower - Linear	Higher – Non-linear

RESULTS AND DISCUSSION

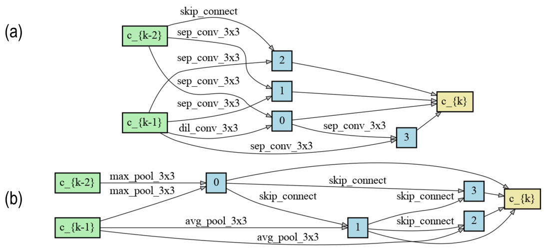


Figure 3. (a) Normal Cell and (b) Reduction Cell discovered by HyperDARTS-2

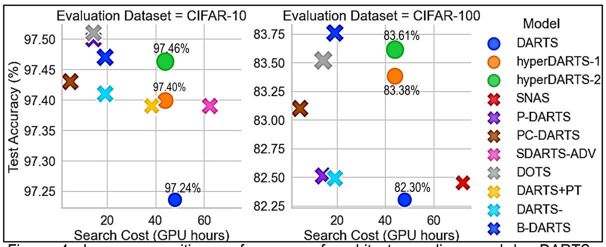


Figure 4. Image recognition performance of architectures discovered by DARTS, hyperDARTS-1, hyperDARTS-2, and other state-of-the-art (SOTA) DARTS variants on CIFAR-10 (left) and CIFAR-100 (right).

As shown in Figure 4, architectures discovered by both **hyperDARTS-1 and -2 outperform DARTS**-discovered architecture in two aspects of image recognition task performance:

1. Image recognition capability, with higher test accuracies on CIFAR-10 and -100.
2. Transferability (generalization capability), with higher test accuracy on CIFAR-100.

The improvement is not trivial as it is **comparable to other** SOTA **DARTS variants.**

HyperDARTS-2 outperforms hyperDARTS-1 with its non-linear hypernetwork capacity to model $\boldsymbol{w}$ more accurately, discovering architectures with higher task performance.

Future Work to further improve hyperDARTS performance in finding optimal architectures:

1. Combine hyperDARTS with other DARTS variants since most of them still employ the approximation-based search strategy proposed by DARTS.
2. Improve hypernetwork capacity in modelling $\boldsymbol{w}$ to lessen candidate architecture interference issue. This issue can be reduced in hyperDARTS which uses hypernetworks, but it is an unsolvable issue in most other DARTS-based models.

CONCLUSION

HyperDARTS can find architectures with higher image recognition task performance than DARTS. The improvement is not trivial as it is comparable to other DARTS variants. The aim and all 3 objectives have been achieved.

REFERENCES

[1] H. Liu, K. Simonyan, and Y. Yang, "DARTS: Differentiable architecture search," in International Conference on Learning Representations, 2019. [Online]. Available: https://openreview.net/forum?id=S1eYHoC5FX