Fine-Tuning DARTS for Image Classification

**Muhammad Suhaib Tanveer[1], Muhammad Umar Karim Khan[2], and Chong-Min Kyung[1]**

[1]School of Electrical Engineering, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, 34141 South Korea

[2]Centre of Integrated Smart Sensors (CISS) KAIST Daejeon, 34141 South Korea,

suhaibtanveer@kaist.ac.kr

1. Introduction

- Neural Architecture Search (NAS) methods have gained popularity due to superior classification performance.
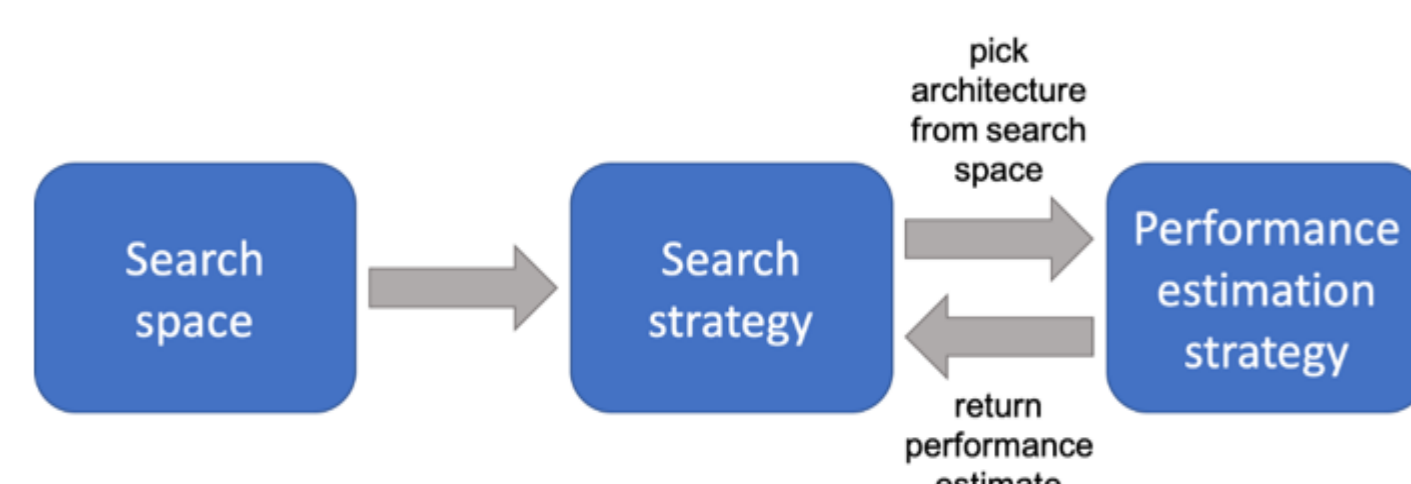

- Differentiable Architecture Search (DARTS) is an efficient NAS method, but it makes numerous approximations.

- We propose to add fixed operations in the form of attention modules to make DARTS independent of those approximations.

2. Motivation

- Neural Architecture Search methods do not make use of previous research done for designing manual architectures.

- Fine-tuning is a proven method for improving performance of a neural network.

- Attention modules have proven to be effective for improving classification accuracy.

- Performance of a neural network is extremely sensitive to the input layer.

3. Methodology

- We propose to introduce an attention module after each cell.

- As the architecture of attention module is fixed so it is not affected by the approximations of the DARTS method.
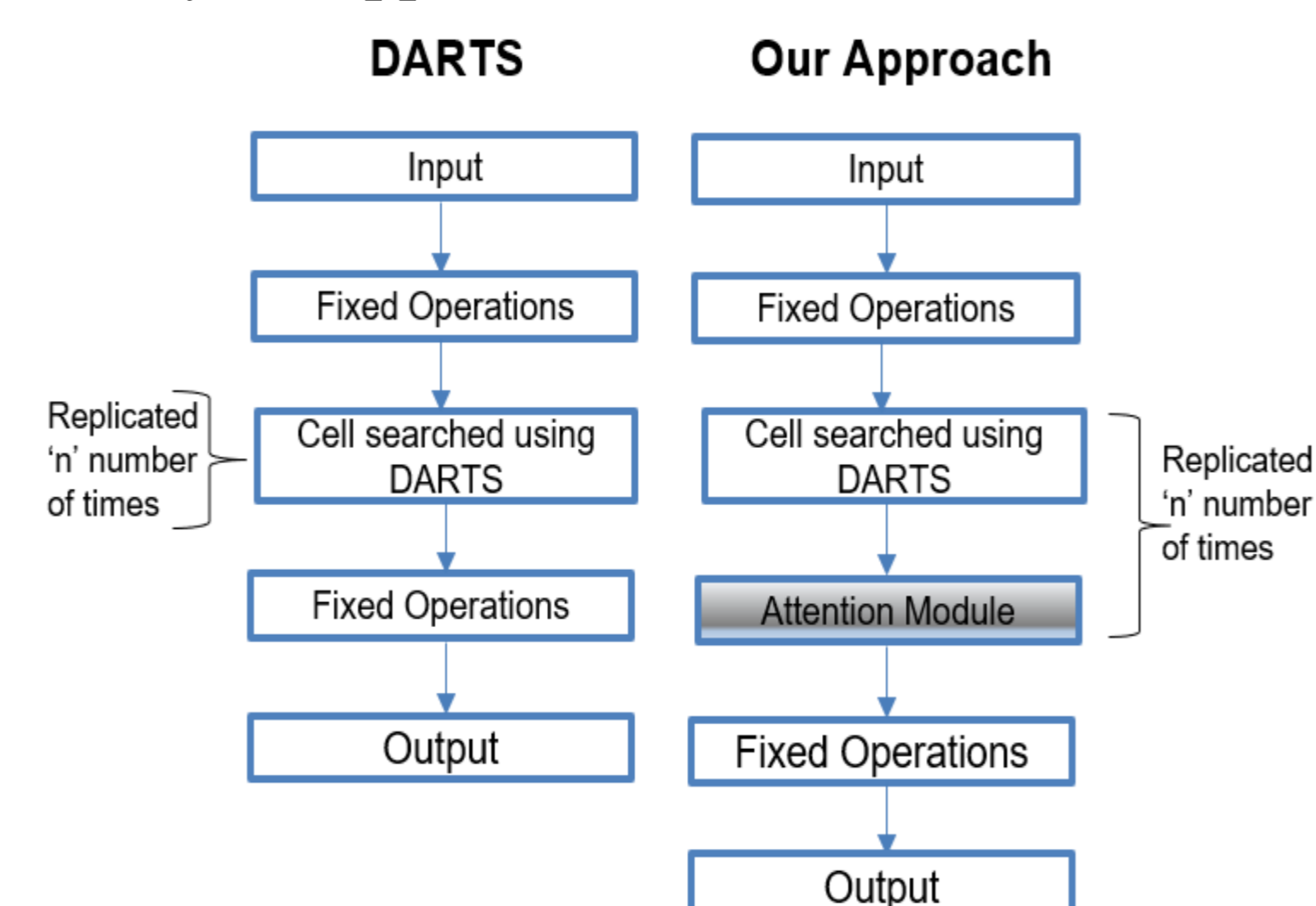
- Contrary to DARTS, we propose to pass different information to the first cell of our network.

- As performance of a neural network is sensitive to its input layer so our approach gives better accuracy.
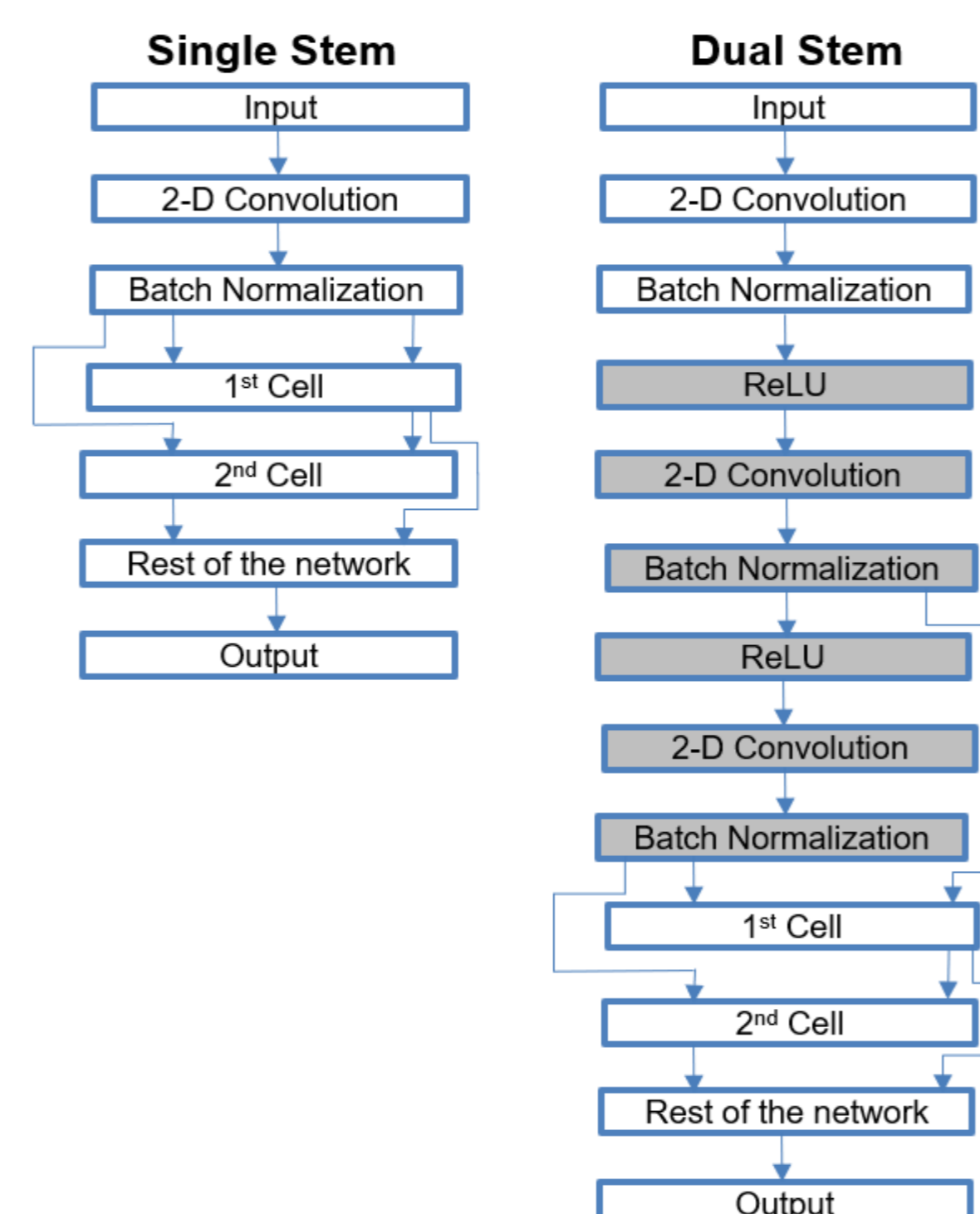
4. Results

- We achieve state-of-the-art results on Fashion-MNIST, COMP-CARS and MIO-TCD datasets.
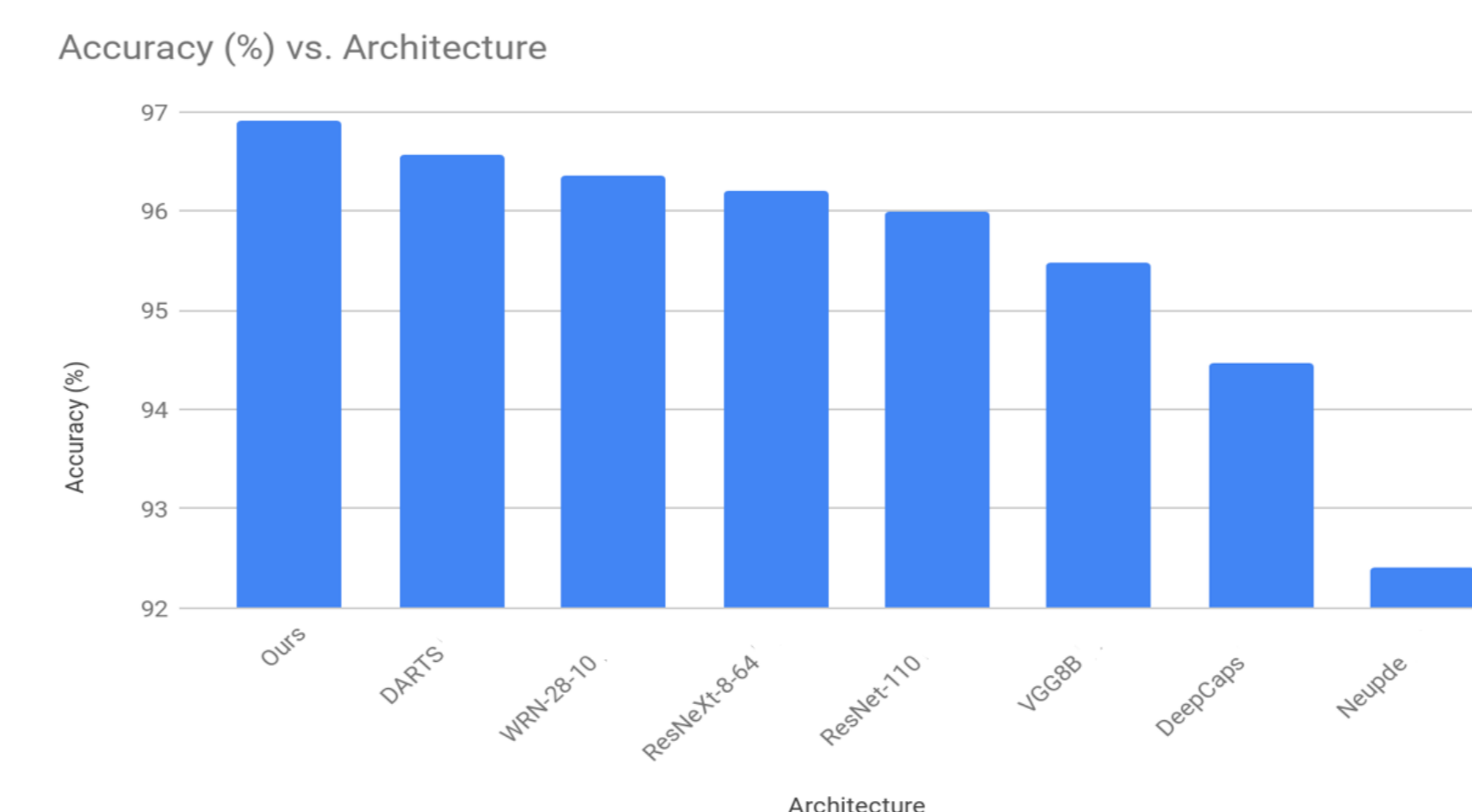
Fashion-MNIST

Accuracy (%) vs. Architecture

Image	Ours	DARTS	SNAS	Ground Truth
	Audi RS7	Audi RS5	Audi RS3	**Audi RS7**
	Porsche Panamera	Porsche Panamera Hybrid	Porsche Panamera Hybrid	**Porsche Panamera**

Image	Ours	DARTS	SNAS	Ground Truth
	Cadillac ATS	Cadillac ATS-L	Cadillac ATS-L	**Cadillac ATS**
	Zhonghua H330	Zhonghua H320	Zhonghua H320	**Zhonghua H330**

5. Conclusion

- Based on the proven success of fine-tuning in manually designed architectures, we propose to fine-tune DARTS by adding fixed operations (in the form of attention module).

- These operations are independent of the approximations used in DARTS, so it results in superior performance.

- Contrary to DARTS, we use a dual-stem approach at the inputs of the first cell.

- We conducted experiments on CIFAR-10, CIFAR-100, Fashion-MNIST, COMP-CARS and MIO-TCD, and our results demonstrate the validity of our claim.

- We were able to obtain state-of-the-art results on Fashion-MNIST, COMP-CARS and MIO-TCD datasets while our results on other datasets were also competitive.

6. Future Work

- Applying our proposed method for NLP tasks (e.g., text classification, sentiment analysis, NER, and Relationship Extraction).

- Applying our proposed method for GAN training(using adversarial loss for training).

- Designing a NAS centered attention module for better adaptability to the said task.

- Exploring other approaches for fine-tuning Neural Architecture Search method.