A Transfer Learning Based Evolutionary Deep Learning Framework to Evolve Convolutional Neural Networks

Bin Wang
School of Engineering and Computer Science
Victoria University of Wellington
Wellington
bin.wang@ecs.vuw.ac.nz

Bing Xue
School of Engineering and Computer Science
Victoria University of Wellington
Wellington
bing.xue@ecs.vuw.ac.nz

Mengjie Zhang
School of Engineering and Computer Science
Victoria University of Wellington
Wellington
mengjie.zhang@ecs.vuw.ac.nz

ABSTRACT

The manual design of CNNs has become exceptionally complex due to the more sophisticated CNN architectures. Thankfully, more and more researchers endeavour to mitigate the difficulty of manual design by designing automated process, but the computational cost of the automatic methods is extremely high due to the huge search space. In this paper, an evolutionary deep learning framework based on transfer learning is proposed to reduce the computational cost, while maintaining the classification at a competitive level. The main idea is to evolve a CNN block from smaller datasets, and then increasing the capacities of the evolved block to handle larger datasets. The proposed method obtains good CNNs with less than 40 GPU-hours. It also achieves a promising error rate of 3.46% on the CIFAR-10 dataset.

CCS CONCEPTS

• **Computing methodologies** → **Object recognition**; **Neural networks**; **Genetic algorithms**;

KEYWORDS

Evolutionary Deep Learning, Convolutional Neural Networks, Image Classification, Neural Architecture Search.

ACM Reference Format:
Bin Wang, Bing Xue, and Mengjie Zhang. 2021. A Transfer Learning Based Evolutionary Deep Learning Framework to Evolve Convolutional Neural Networks. In *2021 Genetic and Evolutionary Computation Conference Companion (GECCO '21 Companion), July 10–14, 2021, Lille, France.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3449726.3459455

1 INTRODUCTION

The CNN architectures have evolved significantly in recent years by increasing the depth and introducing more complex topology. From VGGNet [4] to DenseNet [1], the depth has grown from tens of layers to hundreds of layers. Also, the topology is from the feed-forward fashion to including complex shortcut connections. Therefore, it has become more complex to manually design CNNs. Besides, the more complex CNNs require more time to train, so the trials during the manual design process are taking a much longer time. To reduce human efforts for designing CNNs, both reinforcement learning (RL) and evolutionary computation (EC) methods are leveraged to automatically design CNN architectures. An outstanding RL method is NASNet [6], which has achieved the state-of-the-art classification accuracy, but it takes 2,000 GPU-days to obtain the CNN architecture. On the other hand, AmobaNet [2] has set the state-of-the-art classification accuracy for the first time by using EC methods. However, the computation cost is still high, which is 3150 GPU-days. In this paper, a transfer learning based evolutionary deep learning (EDL) framework is proposed to mitigate the issue of the huge computational cost. The proposed EDL framework learns CNN blocks from smaller datasets, which are then transferred to larger datasets.

GECCO '21 Companion, July 10–14, 2021, Lille, France

ACM ISBN 978-1-4503-8351-6/21/07…$15.00
https://doi.org/10.1145/3449726.3459455

2 THE PROPOSED EEDL FRAMEWORK

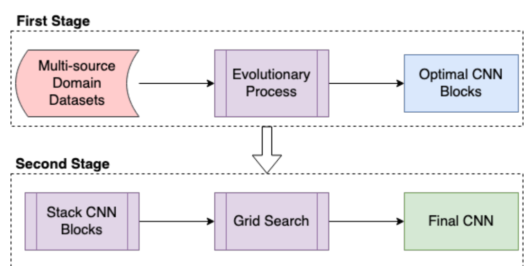

Figure 1: The proposed efficient EDL framework.

The proposed overall efficient EDL (EEDL) framework is composed of two parts/stages as shown in Fig. 1, which are the source domain learning and the target domain learning based on transfer learning. A CNN block is evolved from the first stage of the source domain learning, and a more complex CNN is learned by stacking the CNN block from the second stage of the target domain learning as the final CNN architecture. Another benefit of introducing the two-stage learning with multi-source domain datasets is that a generalised CNN block is expected to be evolved only once. For any target domain dataset, the evolved CNN block can be used for the second stage of the target domain learning, so the first stage does not have to be repeated, which significantly saves the computational cost. To further accelerate the evolutionary process, the surrogate model proposed in [5] is adopted.

3 EXPERIMENTAL DESIGN

3.1 Benchmark Datasets

The widely-used benchmark dataset for image classification — CIFAR-10 is used to evaluate the proposed method, which is the target domain. Based on the availability and similarity of smaller datasets to the target domains, the MNIST and Fashion-MNIST are chosen as the source domain datasets for learning the knowledge to solve image classification problems.

3.2 Parameter Settings

Table 1: Parameter settings

Parameter	Value
inertia weight w	0.7298
acceleration coefficient $c1$	1.49618
acceleration coefficient $c2$	1.49618
velocity range	[-12.5, 12.5]
population size	30
number of generations	50

The parameters of the experiments are depicted in Table 1. The PSO parameters are set based on the community convention [3].10 runs are performed to achieve the experimental results for statistical tests.

4 RESULT ANALYSIS

4.1 Performance Comparisons on CIFAR-10

Table 2 lists the performance on the CIFAR-10 dataset including the error rate, the number of parameters and the computational cost of searching the network architectures (including both the source and target domain learning). The bold values mean the corresponding competitors outperform the proposed method. While the others indicate the proposed method excels. For the proposed method, the results from 10 runs are presented at the bottom of the table with *the best value* and *the mean value ± the standard deviation*. In regard to the error rate, i.e. the classification accuracy, two out of the seven peer competitors achieve better performance than the proposed method. However, the CNN models from three competitors with smaller error rates substantially outsize the models obtained by the proposed method. In terms of the number of parameters, the proposed method ranks the 3rd among the seven peer competitors and itself. Again, the two competitors with smaller model sizes significantly sacrifice the classification comparing to the proposed method. Looking at the computational cost, the proposed method takes only 40 GPU-hours to accomplish the neural architecture search task, which is faster than all of the 7 peer competitors. Overall, the proposed has demonstrated superior performance against the 7 peer competitors on the CIFAR-10 dataset by considering the classification accuracy, the model size and the computational cost.

Table 2: Performance comparison with peer competitors on CIFAR-10

Method	CIFAR-10 (Error rate%)	Number of Parameters	Computational Cost
ResNet-110	6.43	**1.7M**	–
NASNet-A (7 @ 2304)	**2.97**	27.6M	2,000 GPU-days
NASH (ensemble across runs)	4.40	88M	4 GPU-days
AmoebaNet-B (6,128)	**2.98**	34.9M	3150 GPU-days
Hier. repr-n, evolution (7000 samples)	3.75	–	300 GPU-days
HGAPSO	4.37	–	7 GPU-days
CGP-CNN(ResSet)	5.98	**1.68M**	29.8 GPU-days
EEDL (Best classification accuracy)	3.46	2.29M	< 40 GPU-hours
EEDL (10 runs)	3.53±0.0092	2.41M±0.04M	< 40 GPU-hours

5 CONCLUSIONS

The overall goal of proposing a new evolutionary deep learning framework has been successfully achieved. From the experimental results, the proposed method has achieved very competitive performance in terms of the classification accuracy, the model size of the evolved CNNs, and the computational cost.

REFERENCES

[1] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger. 2017. Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 4700–4708.

[2] Esteban Real, Alok Aggarwal, Yanping Huang, and Quoc V. Le. 2019. Regularized Evolution for Image Classifier Architecture Search. In *The Thirty-Third AAAI Conference on Artificial Intelligence (AAAI)*. 4780–4789.

[3] Yuhui Shi and Russell C Eberhart. 1998. Parameter selection in particle swarm optimization. In *International conference on evolutionary programming*. Springer, 591–600.

[4] Karen Simonyan and Andrew Zisserman. 2015. Very Deep Convolutional Networks for Large-Scale Image Recognition. In *3rd International Conference on Learning Representations (ICLR)*. http://arxiv.org/abs/1409.1556

[5] B. Wang, B. Xue, and M. Zhang. 2021. Surrogate-Assisted Particle Swarm Optimization for Evolving Variable-Length Transferable Blocks for Image Classification. *IEEE Transactions on Neural Networks and Learning Systems* (2021), 1–14. https://doi.org/10.1109/TNNLS.2021.3054400

[6] Barret Zoph, Vijay Vasudevan, Jonathon Shlens, and Quoc V Le. 2018. Learning transferable architectures for scalable image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 8697–8710.