Particle Swarm Optimization for Evolving Deep Convolutional Neural Networks for Image Classiﬁcation

Bin Wang

Supervisors: Mengjie Zhang, Bing Xue, Harith Al-Sahaf, Yanan Sun

**Abstract**

Convolutional neural networks (CNNs) is one of the most effective deep learning method to solve image classification problems, but the best architecture of CNN to solve a specific problem can be extremely deep which is hardly designed by humans. This project focuses on utilising Particle Swarm Optimisation (PSO) to automatically search the optimal architecture of CNN without any manual work involved.

1. **Introduction**

Convolutional neural networks (CNNs) have demonstrated their exceptional superiority in numerous machine learning tasks such as speech recognition [1], sentence classification [2] and image classification [3]. However, designing CNNs for specific tasks can be extremely complex which can be seen from some existing efforts done by researchers such as LeNet [4][5], AlexNet[3], VGGNet[6] and GoogLeNet[7]. Since the architecture of CNN gets deeper, the hyperparameters and the weights become more complex which makes the further improvement of the CNN architecture harder. In addition, we cannot expect to get the optimal performance by applying the same architecture on various tasks and we need to adjust the CNN architecture for each specific task which will bring tremendous work as there are thousands of types of machine learning tasks in the industry.

In order to solve the complex problem of the CNN architecture design, evolutionary computation has recently been leveraged to automatically design the architecture without any human effort involved. Interested researchers have done some excellent work on the automatic design of the CNN architecture by using Genetic Programming [8] and Genetic Algorithm [9]. Since PSO is a very simple and effective swarm intelligence algorithm in the evolutionary computation family, in this project, I would like to develop a PSO method to achieve the goal of automatically seeking the optimal CNN architecture.

1. **Goals**

The overall goal of this project is to design and develop an effective and efficient PSO method to automatically discover good architectures and corresponding weight initialisation values of CNNs. The specific objectives of this project are to

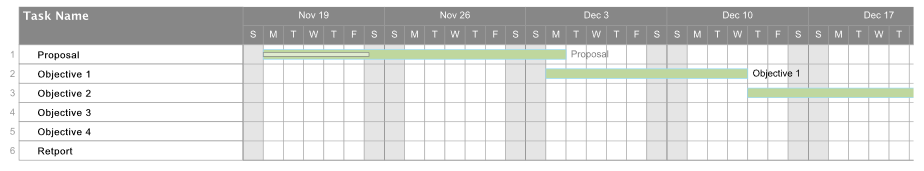
* Design a ﬂexible particle encoding scheme of the architecture, which does not constrain the maximal length of the building blocks in CNNs which are comprised of Convolution layer, Pooling layer and Fully-connected layer. With this PSO encoding scheme, the evolved architecture is expected to beneﬁt CNNs to achieve good performance in solving different tasks at hand.
* Investigate the connection weight encoding strategy, which is capable of representing tremendous numbers of the connection weights in an economical way. More specifically, the standard deviation and mean value of each layer will be encoded in PSO so that the connection weights can be sampled from the corresponding Gaussian distribution. With this encoding approach, the weight connection initialization problem in CNNs is expected to be effectively optimized by the proposed PSO.
* Develop the associated update equation that can cope with the designed particle encoding strategies of both architectures and connection weights.
* Propose an effective ﬁtness measure of the individuals representing different CNNs, which does not require intensive computational resources.

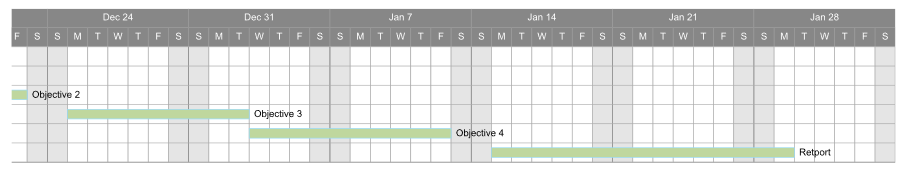
Given the time constraint of the project, I will accomplish the above four objectives first and leave the two following objectives in the future research study.

* Investigate two PSO topology models - global version and local version and develop the topology strategy that converges fast and is not easy to be trapped into local optima.
* Investigate whether the new approach signiﬁcantly outperforms the existing methods in both classiﬁcation accuracy and number of weights.

1. **Plans**

The plan for my study has been broken down into weekly blocks shown in the following chart.





The first two weeks starting from 20 NOV 2017 is allocated for the study proposal, the following 6 weeks will be used to research and develop each objective of the four objectives, and the final report is expected to be accomplished in the final two weeks ending at 29 JAN 2018. If the project goes well, I will convert the report to a conference paper to be submitted to AI 2018.

1. **Evaluation Solution**

In these experiments, nine widely used image classiﬁcation benchmark datasets are used to examine the performance of the proposed PSO method. They are the Fashion [10], the Rectangle [11], the Rectangle Images (RI) [11], the Convex Sets (CS) [11], the MNIST Basic (MB) [11], the MNIST

with Background Images (MBI) [11], Random Background (MRB) [11], Rotated Digits (MRD) [11], and with RD plus Background Images (MRDBI) [11] benchmarks.

The classification accuracy will be compared with some human-designed CNNs such as AlexNet and GoogLenet.

1. **Ethics and Resourcing**

**5.1. Budget**

There is no additional budget for this project.

**5.2 Ethical**

There are no ethical considerations for this project as all data used will be public and freely available.

**5.3 Resources**

Since the computational cost may be high, ECS grid access available to ECS students will be needed to save time.

**Datasets**

MNIST and Fashion MNIST will be used as the datasets, and the URLs are:

<http://yann.lecun.com/exdb/mnist/>

<https://github.com/zalandoresearch/fashion-mnist>

**Bibliography**

[1] Ossama Abdel-Hamid, Li Deng and Dong Yu, “Exploring Convolutional Neural Network Structures and Optimization Techniques for Speech Recognition”, INTERSPEECH 2013, 5 - 29 August 2013, Lyon, France

[2] Yoon Kim, “Convolutional Neural Networks for Sentence Classification”, arXiv:1408.5882v2 [cs.CL] 3 Sep 2014

[3] Alex Krizhevsky, Ilya Sutskever and Geoffrey E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks”, Advances in Neural Information Processing Systems 25 (NIPS 2012), 2012

[4] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, “Backpropagation applied to handwritten zip code recognition,” Neural Computation, vol. 1, no. 4, pp. 541–551, 1989.

[5] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” Proceedings of the IEEE, vol. 86, no. 11, pp. 2278–2324, 1998.

[6] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” arXiv preprint arXiv:1409.1556, 2014.

[7] 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.

[8] Masanori Suganuma, Shinichi Shirakawa, and Tomoharu Nagao, “AGeneticProgrammingApproachtoDesigningConvolutional NeuralNetworkArchitectures∗

”, arXiv:1704.00764v2 [cs.NE] 11 Aug 2017

[9] K. O. Stanley and R. Miikkulainen, “Evolving neural networks through augmenting topologies,” Evolutionary computation, vol. 10, no. 2, pp. 99–127, 2002.

[10] H. Xiao, K. Rasul, and R. Vollgraf, “Fashion-mnist: a novel image dataset for benchmarking machine learning algorithms,” arXiv preprint arXiv:1708.07747, 2017.

[11] H. Larochelle, D. Erhan, A. Courville, J. Bergstra, and Y. Bengio, “An empirical evaluation of deep architectures on problems with many factors of variation,” in Proceedings of the 24th International Conference on Machine Learning. ACM, 2007, pp. 473–480.