IROS 2019 Lifelong Object Recognition competition : Team *NTU_LL*

Duvindu Piyasena, Sathursan Kanagarajah, Siew-Kei Lam and Meiqing Wu
Nanyang Technological University, Singapore
piya0003@e.ntu.edu.sg, ksathursan1408@gmail.com, ASSKLam@ntu.edu.sg, meiqingwu@ntu.edu.sg

***Abstract*—The challenge *Lifelong Object Recognition*, requires exploring how knowledge acquired on previous tasks can be leveraged when learning future tasks, while also memorizing the past tasks, efficiently. This capability is explored under a *difficulty-incremental* scenario in IROS 2019, lifelong object recognition challenge, where a model is required to perform continual learning under different environmental conditions such as illumination, occlusion, clutter, resolution and camera pose. We show that by using a combination of regularization based lifelong learning and data augmentation, model learning can be made generalized under varying environmental conditions.**

***Index Terms*—lifelong-learning, data-augmentation, computer-vision, robotics**

I. INTRODUCTION

The objectives of the challenge are

- Efficiently retain previous knowledge acquired
- Leverage past knowledge to generalize to new tasks

when continuously learning under difficulty-incremental environmental conditions such as illumination, occlusion, clutter, resolution and camera pose.

Conventionally, Deep Neural Networks have shown to perform poorly when learning continuously from changing data distributions, leading to *catastrophic forgetting* or *catastrophic interference*, leading to a complete or partial loss of previously acquired knowledge [3]. Lifelong Learning, is a branch of deep learning that aims to tackle this challenge and build models that can progressively acquire and retain knowledge from changing data distributions.

As popular literature suggests, current approaches in Lifelong learning can be identified under 3 broad categories : ***1) Regularization 2) Architectural 3) Replay***, based approaches [4], [5].

Regularization approaches protects previous knowledge by modifying the normal loss to retain previous knowledge [6], [7], [8]. Architectural approaches, modifies network architecture such that new knowledge can be learnt without interfering with previous knowledge. Replay methods, are based on using a subset of samples from previous tasks, or generating past data and leverage them to while training on new task. (See [4], [5]) for detailed review of these approaches).

We use regularization based approach along with data augmentation to solve the challenge [1], [2].

II. PROPOSED METHODOLOGY

A. Regularization

Regularization is a popular lifelong learning approach used to prevent catastrophic interference in lifelong learning. Broadly, regularization approaches fall under 2 categories [5].

1) Knowledge Distillation Methods
2) Preventing changes to parameters important to old tasks

Learning Without Forgetting, [6], uses knowledge distillation to protect old task performance, and falls under the first category. *Elastic Weights Consolidation* [7] and *Synaptic Intelligence* [8], are popular work that falls to the second category. Both these methods, measure importance of each parameter to previous knowledge and use that information to augment the loss term during training. This helps prevent drift of the important weights to the previous task and only change ones which are not critical. We used *Synaptic Intelligence*, based regularization to solve this task.

B. Data Augmentation

The dataset used in the challenge [1] is imbalanced in terms of the different environmental conditions. Hence to prevent overfitting of model and for better generalization we use data augmentation. This allows the model learning be performed with some level of invariance to illumination, resolution, occlusion and clutter. Table I, summarises applied data augmentations.

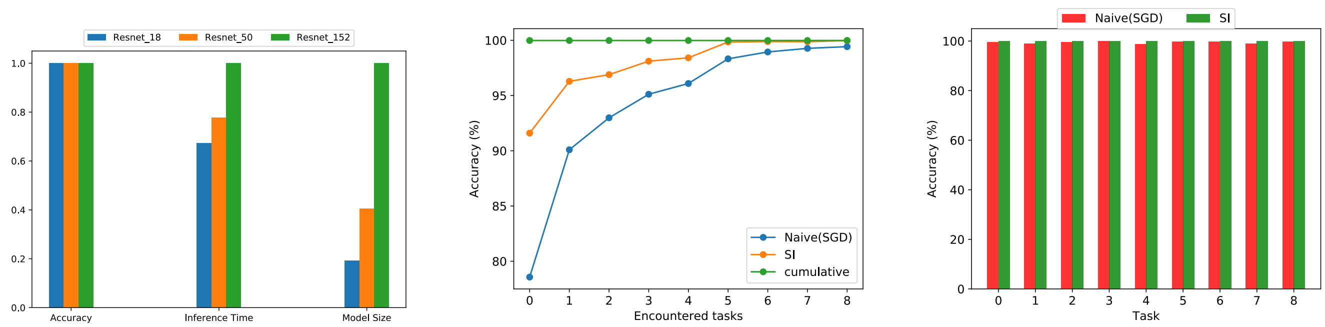
III. EXPERIMENTAL RESULTS

The training was performed using the following methods

- ***Naive*** : Incremental batch learning using SGD
- ***Cumulative*** : Combined training of all tasks (to be used as upperbound accuracy)
- ***SI*** : Incremental batch learning using SGD with *Synaptic Intelligence* based regularization
- ***SI + Aug*** : Incremental batch learning using SGD with *Synaptic Intelligence* based regularization and data augmentation

Image Augmentation	Configuration	Purpose
Color Jitter	*Random brightness & contrast ((b : 0.5,1), (c : 0.5, 2))*	Illumination invariance
Gaussian Blur	*mean = 0, std. dev = 0.3, p=0.1*	Resolution invariance
Random Affine	*degrees = +/- 10*	Camera Pose invariance
Horizontal Flip	*p = 0.2*	Camera Pose invariance

TABLE I: Data Augmentations Used



(a) *Model Selection* **(b)** *Accuracy with tasks encountered* **(c)** *Final accuracy by task*

Fig. 1: *First round results*

	Parameter	Round 1	Final Round
SGD	Model	**Resnet-18**	**Resnet-18**
	Batch size	128	128
	Epochs	4	5
	Optimizer	SGD	SGD
	Learning rate	0.001	0.001
Synaptic Intelligence	Regularization factor(SI)	0.2	4
	Epochs	8	8

TABLE II: Training Parameters of two rounds

Method	Accuracy
Naive(SGD)	99.42
Cumulative(SGD)	99.98
SGD + SI	99.98

TABLE III: First round results summary

The training parameters used in two rounds are summarised in Table II.

A. First Round

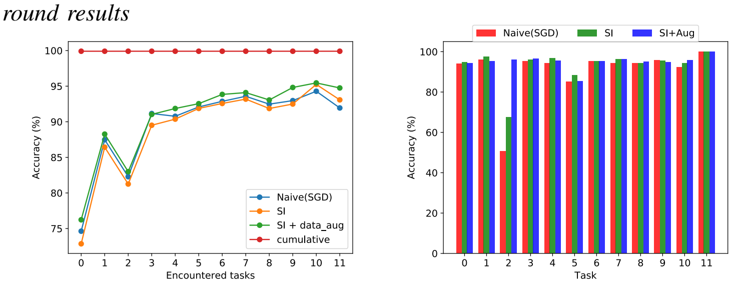
The task was evaluated using three different variants of Resnet, namely Resnet-18/50/152. Cumulative training on the three models, gave the same accuracy. Hence Resnet-18 was selected, due to the smaller model size and faster inference speed. The model selection results are shown in Fig. 1a.

During the first phase, *Naive*, *Cumulative* and *SI* methods were adopted. The incremental batch training results are shown in Fig. 1b. All 3 training methods achieve same accuracy eventually. The difference in the change of accuracy with tasks encountered, could be attributed to the lesser number of epochs used in *Naive* training (4) than in *SI* training (8). The final accuracies are summarised in table III. Fig. 1c, shows the accuracy for each task in the final model.

B. Final Round

In addition to the previous training methods, data augmentation was used in combination with regularization, during final round. The change of accuracy with the encountered tasks is indicated in Fig. 2a. Generally, it can be seen under different difficulty levels accuracy levels fluctuating. Noticeably, for task 3, under low illumination and task 12, where clutter is high, accuracy can be seen dropping.

However, by analyzing Fig. 2b, it can be seen using *SI* and *SI+Aug*, the model generalizes well across different tasks better than *naive* approach. Particularly, for task 3, it was observed *naive* approach accuracy is 50.79%, while using *SI* improves accuracy to 67.52%. Using *SI+Aug*, improves this



(a) *Accuracy with tasks encountered* **(b)** *Final accuracy by task*

Fig. 2: *Final round results*

	Method	
Final Accuracy(%)	Naive(SGD)	91.94
	Cumulative(SGD)	99.91
	SGD + SI	93.08
	SGD + SI + Aug (Color Jitter)	94.11
	SGD + SI + Aug (Color Jitter + Blur)	**95.04**
	SGD + SI + Aug (Color Jitter + Blur + Affine + Hor. Flip)	91.11
Train Time(min)	SGD + SI	215 min
	SGD + SI + Aug (Color Jitter + Blur)	269 min

TABLE IV: Final round results summary

further to 95.99%. Overall, *SI+Aug*, generalizes better in terms of all the environmental conditions.

IV. Conclusion

Synaptic Intelligence based Regularization and data augmentation increases generalization of model and helps to reduce overfitting of model to specific environmental conditions.

V. Acknowledgement

This research project is partially funded by the National Research Foundation Singapore under its Campus for Research Excellence and Technological Enterprise(CREATE) programme with the Technical University of Munich at TUM-CREATE.

References

[1] Q. She, F. Feng, X. Hao, Q. Yang, C. Lan, V. Lomonaco, X. Shi, Z. Wang, Y. Guo, Y. Zhang, F. Qiao, and R. H. M. Chan, “Openloris-object: A robotic vision dataset and benchmark for lifelong deep learning,” 2019.

[2] F. Feng, R. H. M. Chan, X. Shi, Y. Zhang, and Q. She, “Challenges in task incremental learning for assistive robotics,” *IEEE Access*, vol. 8, pp. 3434–3441, 2020.

[3] M. McCloskey and N. J. Cohen, “Catastrophic interference in connectionist networks: The sequential learning problem,” ser. Psychology of Learning and Motivation, G. H. Bower, Ed. Academic Press, 1989, vol. 24, pp. 109 – 165. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0079742108605368

[4] G. I. Parisi, R. Kemker, J. L. Part, C. Kanan, and S. Wermter, "Continual lifelong learning with neural networks: A review," *Neural Networks*, vol. 113, pp. 54 – 71, 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0893608019300231

[5] T. Lesort, V. Lomonaco, A. Stoian, D. Maltoni, D. Filliat, and N. D. Rodríguez, "Continual learning for robotics," *ArXiv*, vol. abs/1907.00182, 2019.

[6] Z. Li and D. Hoiem, "Learning without forgetting," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 12, pp. 2935–2947, Dec 2018.

[7] J. Kirkpatrick, R. Pascanu, N. Rabinowitz, J. Veness, G. Desjardins, A. A. Rusu, K. Milan, J. Quan, T. Ramalho, A. Grabska-Barwinska, D. Hassabis, C. Clopath, D. Kumaran, and R. Hadsell, "Overcoming catastrophic forgetting in neural networks," *Proceedings of the National Academy of Sciences*, vol. 114, no. 13, pp. 3521–3526, 2017. [Online]. Available: https://www.pnas.org/content/114/13/3521

[8] F. Zenke, B. Poole, and S. Ganguli, "Continual learning through synaptic intelligence," in *ICML*, 2017.