Learning with Delayed Feedback

Pranavan Theivendiram(pranavan@u.nus.edu), **Terence Sim**(tsim@comp.nus.edu.sg)
National University of Singapore

Problem Overview

- Human learns continuously, the distinction between training and testing is never realized.
- One key problem in continual learning is catastrophic forgetting whereas the other one is Delayed Feedback which is heavily overlooked in the literature.

Delayed Feedback

Somebody gives feedback with a car to the learner.

Car
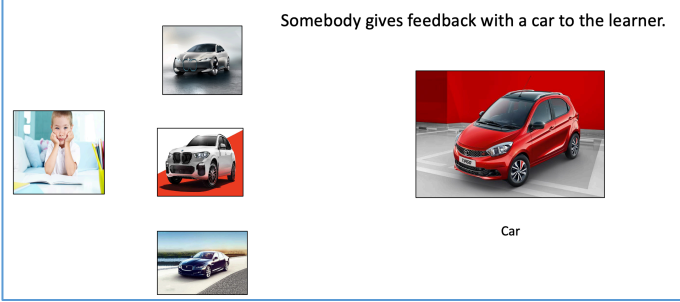
Contributions

- A Continual Learning strategy that enables an Agent to be always learning, without an explicit training/testing phase. This mimics human learning, and is beneficial because the Agent can learn over its lifetime.
- Our strategy also handles delayed feedback, in which the Agent receives a supervisory signal at a later time and is still able to learn from it. Moreover, the feedback can come in a different form. This allows the Agent to learn flexibly and efficiently.
- Our strategy is model-agnostic, and may be applied to any base learner.

Related Work

Machine Learning Paradigm	Characteristics	References
Continual Learning or Lifelong Learning	Learns a sequence of tasks over time, while not forgetting tasks learned previously. [41].	[13], [15], [14], [16], [42]
Curriculum Learning	Learns a sequence of tasks that progressively gets harder.	[28], [29], [43], [44]
Transfer Learning	Adapts a pre-trained model to another task in a different domain.	[33], [34], [45]
Meta-Learning	Monitors the base learner in order to guide it optimally (e.g., by choosing hyper-parameters or sampling the training set).	[35], [36]
Unsupervised Learning	Learns to cluster similar data points together, in a way that is semantically meaningful.	[37], [38]
Self-supervised Learning	Learns good representations from unlabelled data by generating pseudo-labels from parts of the input.	[39], [40]
Learning with delayed feedback (ours)	Updates the model using the given feedback on the current and previous inputs that match, or postpone the updating until suitable feedback arrives in the future.	-

Growing Phase
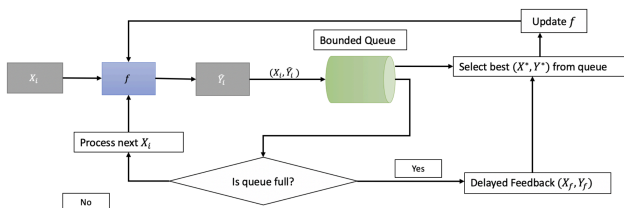
Model Update
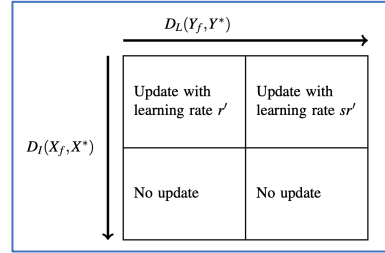
Results – Image Classification

- CIFAR10 dataset is used with the division of 10000 images for Seeding phase and 50000 images for Growing Phase.
- A modified version of AlexNet is used as the learner.
- Base line accuracy is 69.52%.

Queue Size	Output Similarity(%)	Input and Output Similarity(%)	Input and Output Similarity and Batch Update(%)
1000	69.92	69.92	68.37
2000	70.42	70.69	69.46
3000	71.39	71.50	70.23
5000	73.09	73.08	72.01
10000	77.31	77.23	75.89

Results – Image Captioning

- COCO dataset is used with the division of 113287 images for Seeding phase and 5000 images for Growing phase.
- A neural network with LSTM is used as the learner.
- Baseline BLUE4 score is 0.225.

Queue Size	Output Similarity	Input and Output Similarity	Bounding Boxes	Image Segmentation
100	0.233	0.237	0.228	0.230
200	0.232	0.236	0.227	0.228
300	0.224	0.235	0.220	0.222
400	0.226	0.228	0.223	0.221
500	0.226	0.231	0.221	0.223

Conclusion and Future Work

- Test the strategy with a completely different forms of feedback such as audio.
- Implement our strategy with other base learners such as SVM, and Decision Trees.