Continual learning on Cifar-10 from Scratch

László Freund, Anna Székely, Aryan Prabhudesai

Introduction

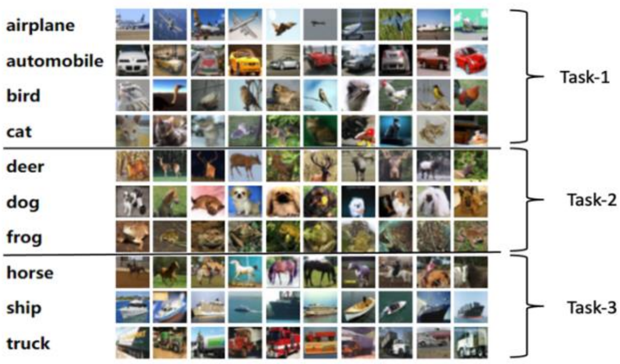
Continual learning is one of the main challenges that current machine learning agents face. Natural organisms learn continuously from the experiences that they are exposed to, but to train artificial agents, we need radically different training methods. AI algorithms require the training data to be i.i.d., stationary and that different training examples are distributed evenly across the training data. These assumptions seriously constrain the applicability of neural networks in multiple natural situations (Aljundi, 2019; Flesch et al., 2018; Hadsell et al., 2020; Parisi et al., 2019). Thus, we concentrate on a possible neural network architecture that could mitigate these constraints.

Previous solutions

Multiple solutions were proposed to overcome the issues of continual learning and prevent catastrophic forgetting. Two main directions are the experience replay and a generative replay of samples (Aljundi et al., 2019; Shin et al., 2017). Both directions aim to "update the model's memory" with samples from previous tasks. The difference is whether the samples shown to the model are actual previous samples or fake ones generated only for the sake of updating the model.

Dataset

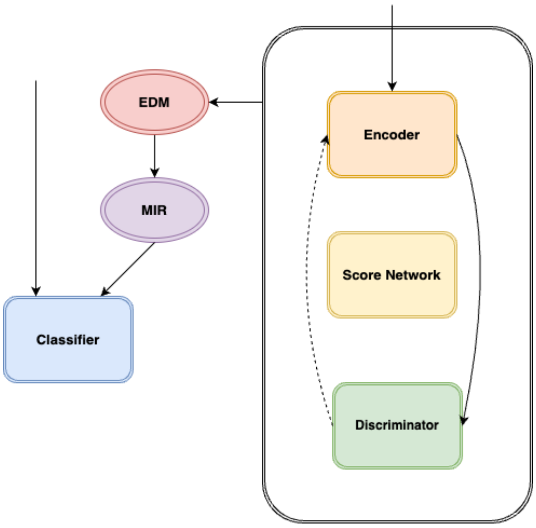
Our solution concentrates on an image recognition task, thus we use the widely used Cifar-10 dataset. We define different tasks to learn continually by splitting the dataset into subsets of pictures of three categories, thus within Cifar-10, we have three tasks to learn (see fig. 1.). The goal is to teach our model for the first task and then continue to learn on the second task without falling back to the performance level before training. In other words, we would like our model to transfer the knowledge acquired during previously learned tasks to solve new challenges. Furthermore, we expect our model not to forget previous tasks when learning a new one.



1. Figure: Cifar-10 split to 10 tasks

Proposed method

We propose a network architecture that consists of a classifier and a generator model (see fig XX.). The classifier's task is to classify the incoming input images into the appropriate categories. The task of the generator is to update the classifier model with generated pictures from other categories that are likely to be forgotten[1]. For the generator model, we used the state-of-the-art guided diffusion process, which consists of an encoder and a discriminator network (trained through the task), and a pre-trained score network that optimizes the data generation process. The generator models' output is then fed into EDM G++ before arriving in the classifier.



2. Figure: proposed network architecture

Evaluation method

Our models' performance could be evaluated by the following measures:

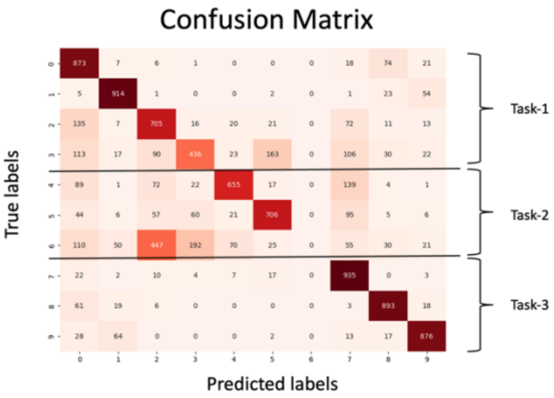
1) forward transfer: the learned tasks should contribute to the better acquisition of the new task.
2) backward transfer (or the lack of forgetfulness): training on new tasks should improve, but at least keep classification performance on previous tasks.

Results

Our model is able to perform reliably on different tasks even when trained sequentially. Confusion matrix shows the achieved accuracy when tested on samples coming from different tasks (see fig. 3).

Accuracy of the trained task and the previous tasks is shown in table 1. Increasing accuracy values from task-1 to task-3 on the diagonal shows that our model gets better throughout the whole training procedure, while values left from the diagonal show that our model does not forget catastrophically the previous tasks. These results show that our model is able to transfer previously acquired knowledge to new tasks (forward transfer), while our minimum expectation regarding the backward transfer (not being catastrophically forgetful) is also satisfied.

[1] Optimally, to find the "most likely to be forgotten" category, we would rely on the method of *maximally interfered retrieval* (Aljundi et al., 2019), however, memory constraints prevented us from training the model with MIR. Nevertheless, we regard the model, augmented with MIR a promising further improvement of our current model.



3. Figure: classification performance across tasks

Accuracy on currently trained and previous tasks

		Test		
		Task-1	Task-2	Task-3
Training	Task-1	0.853	-	-
	Task-2	0.620	0.885	-
	Task-3	0.732	0.453	0.901

Table 1: accuracy performances at each stage of the sequential training

Discussion

We proposed a new model architecture to overcome one major challenge of continual learning: the issue of catastrophic forgetting. Previous solutions for the same challenge offered different types of “memory updates” for the model (e.g., generative or replay based). Our solution exceeds the previous attempts by applying a more elaborate generator that is able to keep the model up to date with more realistic samples of the data. Current results are comparable to state-of-the-art solutions, nevertheless with greater computational capacity, the proposed model could be further improved (e.g., by introducing the Maximally Interfered Retrieval, which was not implementable due to RAM shortage). The proposed architecture for further improvement is shown in the appendix, on figure 4.

References

Aljundi, R. (2019). *Continual Learning in Neural Networks*. http://arxiv.org/abs/1910.02718

Aljundi, R., Caccia, L., Belilovsky, E., Caccia, M., Lin, M., Charlin, L., & Tuytelaars, T. (2019). *Online Continual Learning with Maximally Interfered Retrieval*. http://arxiv.org/abs/1908.04742

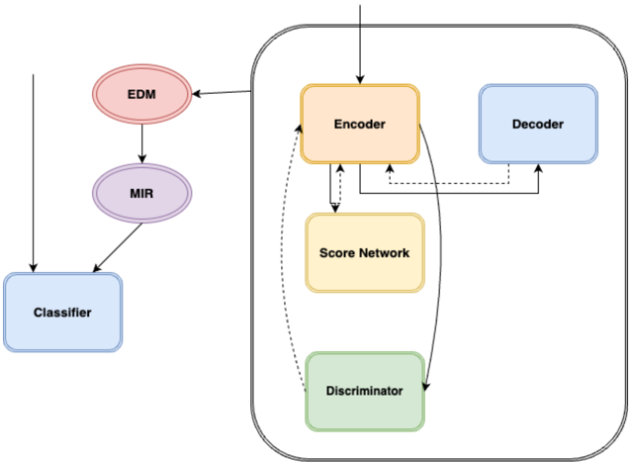
Flesch, T., Balaguer, J., Dekker, R., Nili, H., & Summerfield, C. (2018). Comparing continual task learning in minds and machines. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(44), E10313–E10322. https://doi.org/10.1073/pnas.1800755115

Hadsell, R., Rao, D., Rusu, A. A., & Pascanu, R. (2020). Embracing Change: Continual Learning in Deep Neural Networks. In *Trends in Cognitive Sciences* (Vol. 24, Issue 12, pp. 1028–1040). Elsevier Ltd. https://doi.org/10.1016/j.tics.2020.09.004

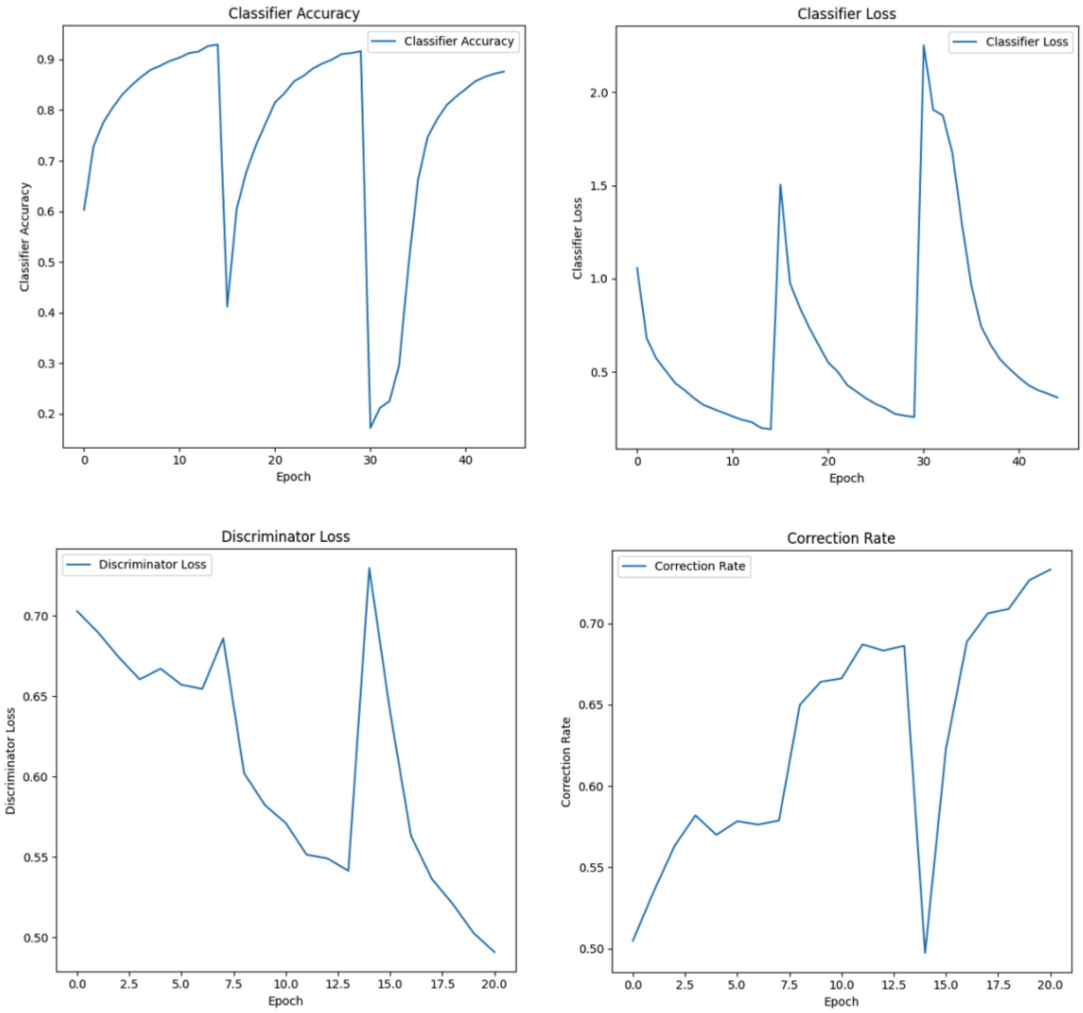
Parisi, G. I., Kemker, R., Part, J. L., Kanan, C., & Wermter, S. (2019). Continual lifelong learning with neural networks: A review. In *Neural Networks* (Vol. 113, pp. 54–71). Elsevier Ltd. https://doi.org/10.1016/j.neunet.2019.01.012

Shin, H., Lee, J. K., Kim, J., & Kim, J. (2017). *Continual Learning with Deep Generative Replay*. http://arxiv.org/abs/1705.08690

Appendix



4. Figure: further improvment of the implemented model architecture



5. Figure: Loss and accuracy metrics through the training procedure