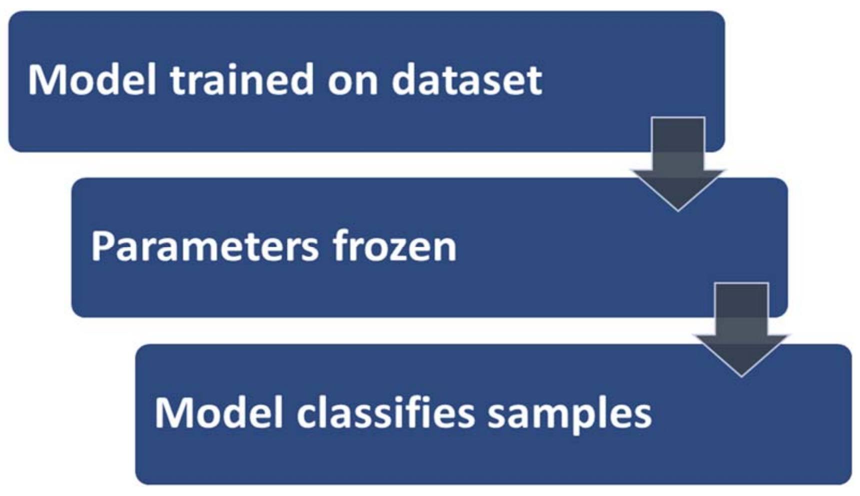
Incremental Learning for Image Classification

Finlay Boyle supervised by Dr Donald Sturgeon

Catastrophic Forgetting | Computer Vision | Machine Learning | Optical Character Recognition | Performance Evaluation

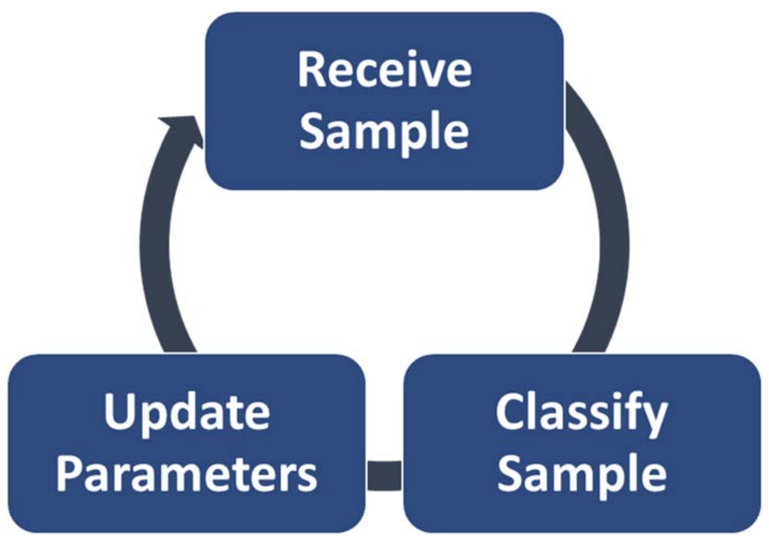
Incremental Learning

Typically when training a neural network, the whole dataset is available and split into batches. These overlapping batches are repeatedly used to train the model—this is called **offline training.**



Offline Training Flow

Incremental Learning Flow



For **incremental/continual learning**, data arrives sequentially then the network is trained on each sample as they appear with the purpose of **adapting to changes in the dataset** and **overcoming catastrophic forgetting.**

Abstract and Aims

Catastrophic forgetting continues to be a challenge for neural networks trained using the traditional batched approach. Incremental learning aims to overcome the phenomenon by providing an **alternative philosophy to train models**. There exist multiple different solution paradigms within incremental learning. This project **aims to explore different approaches** and provide a **unified evaluation of the effectiveness of the state of the art** techniques for incremental learning when applied to image classification and optical character recognition problems.

Catastrophic Forgetting: The complete and sudden loss of previously acquired knowledge in a neural network [1].

Related Work

Initial designs for incremental learning techniques focused on **regularisation**. To limit forgetting, weights that are deemed important are protected from changes by **penalising weight changes with respect to the importance of the connection using loss functions**. Elastic Weight Consolidation [2] estimated the importance from observations of weights and the results were pioneering for incremental learning.

Memory-based approaches are at the forefront of the literature. The sequential data is too much store, instead **exemplars are sampled from incoming data** [3]. These are **used periodically to remind the network** of acquired knowledge. An emerging extension of this is **using generative modelling to implicitly store exemplars** and is an active area of research in the literature [4].

State-of-the-art approaches include:

- **Mnemonics**—bilevel optimisation to find exemplars [5]
- **FearNet**—multiple networks imitating short and long term memory [6]
- **ACAE-REMIND**—compressing exemplars to network features [7]

Methodology

Research Question: What are the most effective strategies to improve incremental learning performance for image classification in neural networks?

The major issues in the literature are the **lack of comparability between proposed techniques and existing state-of-the-art** and a **lack of standardisation in the formulation of the problem** [8]. I intend to implement different approaches in **PyTorch** in order to compare and **evaluate their effectiveness on realistic datasets.**

Promising techniques are often evaluated on MNIST and CIFAR-10, **limiting the applicability of their results**. Realistic datasets have **greater complexity** such as a CIFAR-100, ImageNet and **Kuzushiji** recognition.

ImageNet is far more complex than MNIST

Techniques will be evaluated using **average forgetting, top-5 accuracy, wall-clock time, memory consumption and overall accuracy** [9].

Validity

It is important that any reputable research accounts for validity. The selected metrics are clearly designed to measure the performance of incremental learning techniques ensuring **face** validity. My planned experiments are derived from the literature and are reflective of the experimental design of other researchers enhancing the **concurrent** and **construct** validity. I will also use standard datasets for **predictive** validity to ensure that future work is comparable.

References

[1]: M. McCloskey and N. J. Cohen, "Catastrophic interference in connectionist networks: The sequential learning problem," in Psychology of learning and motivation. Elsevier, 1989, vol. 24, pp. 109– 165.
[2]: J. Kirkpatrick et al., "Overcoming catastrophic forgetting in neural networks," Proceedings of the national academy of sciences, vol. 114, no. 13, pp. 3521–3526, 2017
[3]: A. Prabhu, P. H. Torr, and P. K. Dokania, "GDumb: A simple approach that questions our progress in continual learning," in European conference on computer vision. Springer, 2020, pp. 524– 540.
[4]: H. Shin et al., "Continual learning with deep generative replay," Advances in neural information processing systems, vol. 30, 2017.
[5]: Y. Liu et al., "Mnemonics training: Multi-class incremental learning without forgetting," in Proceedings of the IEEE/CVF conference on Computer Vision and Pattern Recognition, 2020, pp. 12 245–12 254
[6]: R. Kemker and C. Kanan, "Fearnet: Brain-inspired model for incremental learning," arXiv preprint arXiv:1711.10563, 2017.
[7]: K. Wang, J. van de Weijer, and L. Herranz, "ACAE-REMIND for online continual learning with compressed feature replay," Pattern Recognition Letters, vol. 150, pp. 122–129, 2021
[8]: G. I. Parisi et al., "Continual lifelong learning with neural networks: A review," Neural Networks, vol. 113, pp. 54–71, 2019
[9]: A. Chaudhry et al., "Riemannian walk for incremental learning: Understanding forgetting and intransigence," Proceedings of the European Conference on Computer Vision (ECCV), 2018, pp. 532–547.