MEDICAL UNIVERSITY OF VIENNA

CIR

Dynamic memory to alleviate catastrophic forgetting in continuous learning settings

Matthias Perkonigg[1*], Johannes Hofmanninger[1*], James A. Brink[2], Oleg Pianykh[2], Helmut Prosch[1], Christian Herold[1], Georg Langs[1]

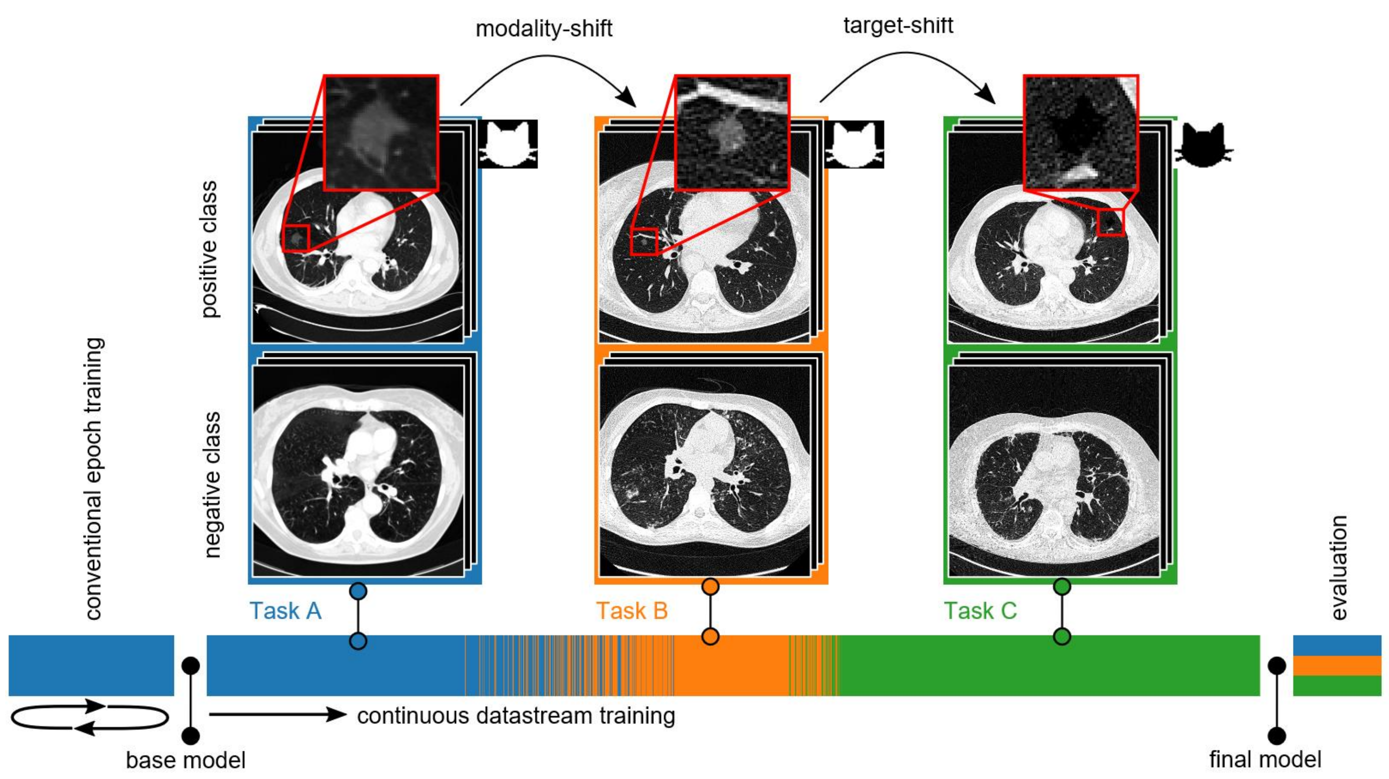
[1] Department of Biomedical Imaging and Image-guided Therapy, Medical University of Vienna
[2] Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Boston, USA

Objective

In medical imaging, technical progress or changes in diagnostic procedures lead to a continuous change in image appearance. While radiologists can easily adapt to changing appearance, such domain and task shifts limit the applicability of machine learning algorithms in the clinical routine by rendering models obsolete over time. Continuously training the model on data with new appearance leads to poor performance on old domains, a phenomena called **catastrophic forgetting.**

We address the problem of data shifts in a continuous learning scenario by adapting a deep learning model to unseen variations in a data stream while counteracting catastrophic forgetting effects. Our method uses a dynamic memory to facilitate rehearsal of a small, but diverse training data subset to mitigate forgetting. The decision on which images from the continuous data stream to keep in memory is based on the distance of gram matrices which encodes the style of an image.

Setting



CT scans were performed on a Siemens scanner with either B3 reconstruction kernel and 3mm slice-thickness (B3/3) or B6 reconstruction kernel and 1mm slice-thickness (B6/1) and imprinted with a synthetic target structure in the form of a cat on random locations, rotations and varying scale at 50% of the cases. The scans are split into base training, continuous training, validation and test set.

	Task A	Task B	Task C	Total
Protocol	B3/3	B6/1	B6/1	
Target	high	high	low	
Base	1513	0	0	1513
Continuous	1513	1000	2398	4911
Validation	377	424	424	1225
Test	381	427	426	1234

Method

We continuously update the parameters of an already trained model with new training data, stored in a dynamic memory $\mathcal{M}$. We compose $\mathcal{M}$ from a continuous data stream to capture novel data characteristics while sustaining the diversity of the overall training corpus.
For each new image of the data stream, $\mathcal{M}$ is updated by three rules:

(1) every novel case will be stored in the memory
(2) a novel case can only replace a case in memory of the same class
(3) the case in memory that will be replaced is close according to a high level metric capturing the style of an image.

We evaluate the style using a high level metric, the gram matrix [2]:

$$G_{ij}^{l}(\mathbf{x}) = \frac{1}{N_l M_l} \mathbf{f}_{il}(\mathbf{x})^{\top} \mathbf{f}_{jl}(\mathbf{x})$$

where $f_{il}(x)$ and $f_{jl}(x)$ are the activations of two feature maps i and j in a layer l given a sample image x.

For a set of convolutional layers C we compare the style of two images by calculating the gram distance:

$$\delta(\mathbf{x}, \mathbf{y}) = \sum_{l \in \mathcal{L}} \frac{1}{N_l^2} \sum_{i=1}^{N_l} \sum_{j=1}^{N_l} (G_{ij}^{l}(\mathbf{x}) - G_{ij}^{l}(\mathbf{y}))^2$$

A new image $\mathbf{x}$ from the data stream replaces the image $\mathbf{y} \in \mathcal{M}$ that minimizes the distance δ.

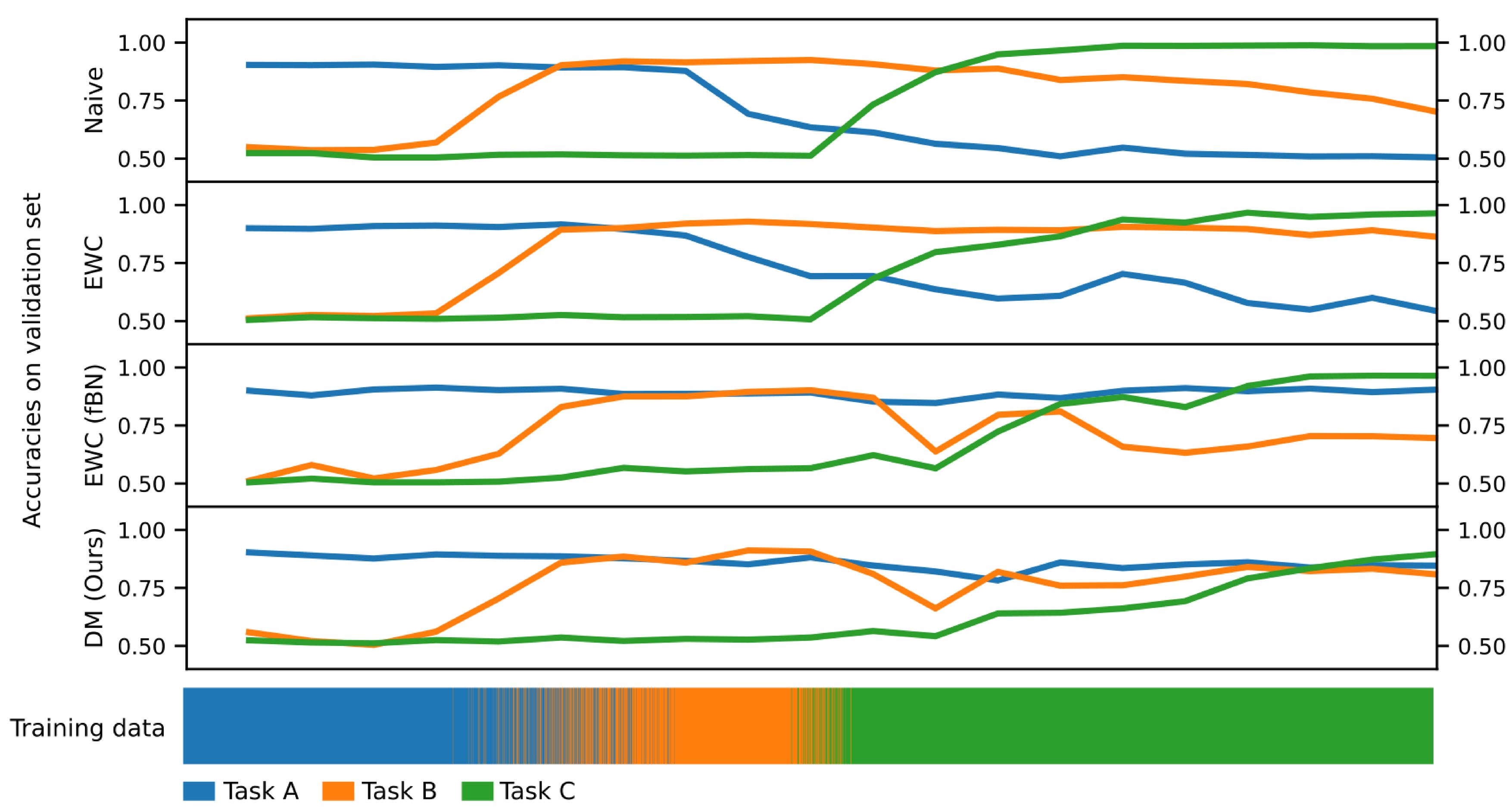
After each update to $\mathcal{M}$ the model is trained with on a mini-batch sampled from the memory.

Results

Different learning strategies are compared:

- Naive: Baseline, without strategy to counter catastrophic forgetting
- Elastic Weight Consolidation (EWC): Reference method presented in [3]. Weight regularization method.
- EWC-fBN: EWC with fixed batch norm layers.
- Dynamic Memory (DM): Our method developed in this work. For a memory size of 32 elements.

	ACC Task A	ACC Task B	ACC Task C
Naive	$\otimes$ 0.51 ± 0.00	$\otimes$ 0.71 ± 0.02	0.98 ± 0.01
EWC	$\otimes$ 0.57 ± 0.00	0.83 ± 0.00	0.91 ± 0.00
EWC-fBN	0.89 ± 0.02	$\otimes$ 0.74 ± 0.05	0.94 ± 0.04
DM (Ours)	0.81 ± 0.02	0.85 ± 0.02	0.92 ± 0.04
Full training	0.92 ± 0.01	0.91 ± 0.02	0.97 ± 0.00



[1] J. Hofmanninger, M. Perkonigg, J. A. Brink, O. Pianykh, C. Herold, and G. Langs, "Dynamic Memory to Alleviate Catastrophic Forgetting in Continuous Learning Settings", MICCAI, 2020.
[2] L. Gatys, A. Ecker, and M. Bethge, "A Neural Algorithm of Artistic Style," Journal of Vision, vol. 16, no. 12, p. 326, 2016.
[3] J. Kirkpatrick et al., "Overcoming catastrophic forgetting in neural networks," Proceedings of the National Academy of Sciences of the United States of America, vol. 114, no. 13, pp. 3521–3526, 2017.

matthias.perkonigg@meduniwien.ac.at