Task-Free Continual Learning

Klaas Kelchtermans*, Rahaf Aljundi* and Tinne Tuytelaars

Introduction

Our goal is to learn **continually** from an infinite stream of data
without the constraint of **storing** the previous data
without catastrophic **forgetting** due to gradual data distribution shifts
without relying on **task boundaries**

James Speakman - Mercury Press

Method

Memory Aware Synapses - method for task-incremental learning

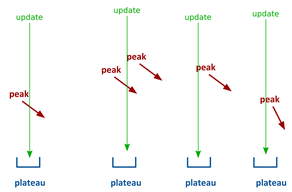
While training a new task, changes on important parameters are penalized:

$$F(x_k;\theta+\delta)-F(x_k;\theta)\approx\sum_i g_i(x_k)\delta_i \qquad \Omega_i=\frac{1}{N}\sum_{k=1}^{N}\|\,g_i(x_k)\,\| \qquad g_i(x_k)=\frac{\partial F(x_k)}{\partial\theta_i}$$

When training the new task, deviations are punished according to their importance:

$$L(\theta)=L_n(\theta)+\frac{\lambda}{2}\sum_i\Omega_i(\theta_i-\theta_i^*)^2$$

1) When to update importance weights?



Loss plateaus are stable moments to update the importance weights, once per plateau.

2) Which data to update them?

Small buffer with hard samples, B, (size 100) is used to stabilize the online setting. Hardest samples are selected.

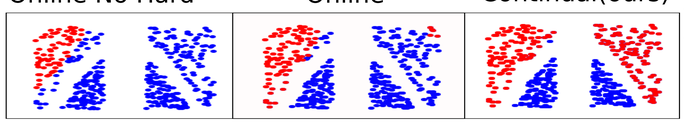
3) How to accumulate them?

Instead of adding the importance weights, each updates computes a **moving average.**

Algorithm 1 Online Continual Learning

Input: $\delta_\mu, \delta_\sigma, \mathcal{N}$
Initialize: $\mathcal{B}=\{\}, \mathcal{W}=\{\}$,
$\Omega=\vec{0}, \mu_L^{old}=0, \sigma_L^{old}=0, \mathcal{P}=0$
repeat
  Receive K recent samples X, Y
  for n in $\mathcal{N}$ **do**
    $\mathcal{L}^T=\mathcal{L}_\theta(X,Y)+\mathcal{L}_\theta(X_\mathcal{B},Y_\mathcal{B})+\frac{\lambda}{2}\sum_i\Omega_i(\theta_i-\theta_i^*)^2$
    $\theta \leftarrow \mathrm{SGD}(\theta,\mathcal{L}^T)$
    if $n=1$ **then**
      $\mathcal{W}\leftarrow\mathrm{update}(\mathcal{W},\mathcal{L}(X,Y),\mathcal{L}(X_\mathcal{B},Y_\mathcal{B}))$
    end if
  end for
  if $\neg\mathcal{P}\wedge\mu(\mathcal{W})<\delta_\mu\wedge\sigma(\mathcal{W})<\delta_\sigma$ **then**
    $\Omega\leftarrow\mathrm{update}(\Omega,\theta,(X_\mathcal{B},Y_\mathcal{B}))$
    $\mu_L^{old}=\mu(\mathcal{W}), \sigma_L^{old}=\sigma(\mathcal{W})$
    $\mathcal{W}=\{\}, \mathcal{P}=1$
  end if
  if $\mu(\mathcal{W})>\mu_L^{old}+\sigma_L^{old}$ **then**
    $\mathcal{P}=0$
  end if
  $(X_\mathcal{B},Y_\mathcal{B})\leftarrow\mathrm{update}((X_\mathcal{B},Y_\mathcal{B}),(X,Y),\mathcal{L}_\theta(X,Y))$

Synthetic experiment



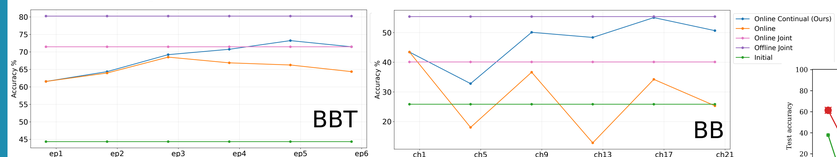
- Classify points in or out of the unit circle.
- Data sampled from 2 quadrants representing task 1 and 2.
- Accuracy on first quadrant (and total performance).

Face recognition in soap series

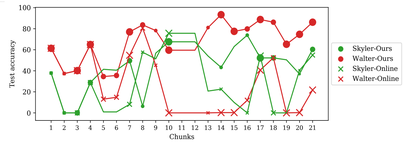
Goal: recognize faces of different actors.
How:
using an online face detector module and a multi-object face tracker.
Actors are classified according to the distance to their closest template within an embedding space.

The Big Bang Theory (BBT), Breaking Bad (BB), Mad Men (MM)

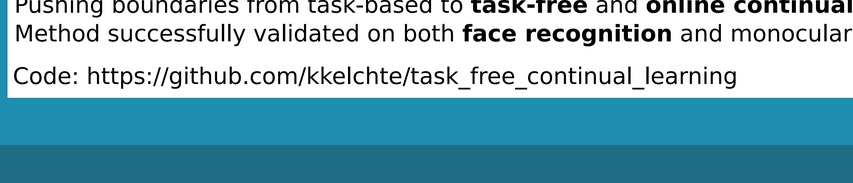
Weak Supervision: an annotator tells whether two consecutive tracks have the same identity.



After chunks with few actor samples performance drops with no continual learning



Self Supervision: assume tracks of two faces from same frame must belong to different actors.

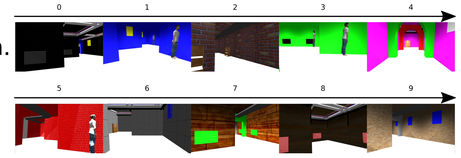


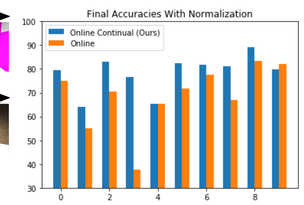
Monocular collision avoidance

Drone in ROS-Gazebo simulation
using off-policy simulation-supervision.

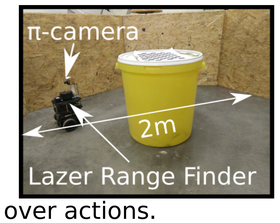
Depth based heuristic demonstrates navigation in sequence of 10 corridor.

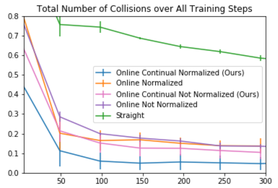
Continual learning avoids forgetting flight in earlier corridors.





Turtlebot in the real world:
- Supervision from heuristic with range finder
- On-policy: agent controls while training (harder)
- Gradual changes in environment.





Output normalization:
sample experiences to have an equal distribution over actions.

Conclusion

Online importance weight-based **continual learning** by defining
when, **on which data** and **how** to update importance weights.
Pushing boundaries from task-based to **task-free** and **online continual learning**.
Method successfully validated on both **face recognition** and monocular **collision avoidance**.

Code: https://github.com/kkelchte/task_free_continual_learning

*Rahaf Aljundi and Klaas Kelchtermans contributed equally to this work.

Acknowledgments
Rahaf Aljundi's PhD is funded by an FWO scholarship. This work was further supported by the CAMETRON research project (GOA) of the KU Leuven and the FWO SBO project Omnidrone.