EVOLVE: Enhancing Unsupervised Continual Learning with Multiple Experts

Xiaofan Yu¹, Tajana Rosing¹, Yunhui Guo²

¹ University of California San Diego
² University of Texas at Dallas

https://github.com/Orienfish/EVOLVE

WACV 2024
Continual Learning

- Continual Learning
  - To continually learn over time by acquiring new knowledge as well as consolidating past experiences
  - Key assumption: continuously changing environments
  - Key challenge: catastrophic forgetting

- While great progress has been made in continual learning, there is still a gap between existing continual learning algorithms and real-world deployments, due to
  - Unpredictable streaming input
  - Lack of supervision and prior knowledge

We consider the unsupervised continual learning (UCL) problem
Related Work

- Most existing works in UCL rely on various prior knowledge to produce good results
  - **Cons**: these prior knowledge may not be available in real-world applications

<table>
<thead>
<tr>
<th>Papers</th>
<th>Single-pass</th>
<th>Non-iid</th>
<th>No task labels</th>
<th>No class labels</th>
</tr>
</thead>
<tbody>
<tr>
<td>VASE [2], CURL [61], L-VAEGAN [77]</td>
<td>✗</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>He et al. [31], CCSL [46], CaSSL [24], LUMP [49]</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td>Tiezzi et al. [70], KIERA [57]</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
</tr>
<tr>
<td>STAM [68]</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

- Self-supervised learning methods with a replay buffer can alleviate catastrophic forgetting on non-iid data streams [ECCV’22]
  - **Cons**: it is unclear how to effectively learn new patterns in an unpredictable world

- In this work, EVOLVE aims at closing the gap between continual learning algorithms and real-world applications by eliminating the prior assumptions.
An Empirical Study of Existing Self-Supervised Learning (SSL) Baselines

- To study whether SSL w/ replay buffer is sufficiently practical, we conduct empirical study
  - Dataset: CIFAR-10 (image), Stream-51 (video) [CVPRW’20]
  - Data streams: iid vs. temporally correlated

SSL baselines experience a significant accuracy drop when applied on video datasets with temporally correlated data streams.
Our Method: EVOLVE

- We propose EVOLVE, a hybrid UCL framework with (1) SSL training on the local client and (2) expert-guided training on the cloud, transmitting a small set of data and intermediate features to the cloud.
Two Key Designs in EVOLVE

- The design of EVOLVE has two key parts
- We use the Hilbert-Schmidt independence criterion (HSIC) for assessing a model’s ability in representation learning

\[
\text{HSIC}(X, Y) = \text{HSIC}(K, L) = \frac{1}{(n-1)^2} \text{Tr}(KHLH),
\]

where \( H = I - \frac{1}{N} 11^T \)

1. **Expert Aggregation Loss based on HSIC**

\[
\mathcal{L}_E = - \sum_{e=1}^{E} p^t_e \cdot \text{HSIC}(K^t_e, L^t),
\]

2. **Dynamic weight adjustment based on a confidence metric**

\[
q^t_e = \sum_i \frac{\exp(h^A_{e,i} \cdot h^B_{e,i} / \tau)}{\sum_k \exp(h^A_{e,i} \cdot h^A_{e,k} / \tau) + \sum_k \exp(h^A_{e,i} \cdot h^B_{e,k} / \tau)},
\]

\[
w^t_{e+1} = \alpha w^t_e + (1 - \alpha) q^t_e.
\]

\[
p^t_e = w^t_e / \sum_l w^t_l
\]

Moving average-based update
Experimental Setup

- **Datasets**: CIFAR-10, TinyImageNet, CORe50 [PMLR’17], Stream-51 [CVPRW’20]
- **Data streams**:

  - **Baselines**:
    - Synaptic Intelligence (SI) [ICML’17]
    - Progressive Neural Network (PNN) [arXiv’16]
    - Dark Experience Replay (DER) [NeurIPS’20]
    - CaSSLe [CVPR’22]
    - Lifelong Unsupervised Mixup (LUMP) [ICLR’22]
  - **Target model**: ResNet-18
  - **Experts**: pretrained ResNet-50, Swin Transformer downloaded from torchvision
  - **Metrics**: kNN accuracy, linear evaluation accuracy
Main Accuracy Results

- EVOLVE outperforms the top baseline using the same SSL by **3.6-20.0% in kNN accuracy** and **6.1-53.7% in top-1 linear evaluation accuracy** across diverse data streams.
Conclusion

- Existing UCL algorithms cannot generalize to real-world scenarios due to
  - Unpredictable streaming input
  - Lack of supervision and prior knowledge
- We propose a general expert-guided continual learning framework, called EVOLVE, with local SSL training on clients and expert-guided training on the cloud
- EVOLVE has two key designs
  - Expert aggregation loss based on HSIC to distill guidance
  - Dynamic weight update to adjust the “power” of diverse experts based on latest data input
- EVOLVE outperforms existing UCL baselines on both image- and video-based data streams using the same SSL backbone
References

- Lomonaco, Vincenzo, and Davide Maltoni. "Core50: a new dataset and benchmark for continuous object recognition." PMLR’17
- Zenke, Friedemann, Ben Poole, and Surya Ganguli. "Continual learning through synaptic intelligence." ICML’17
- Fini, Enrico, et al. "Self-supervised models are continual learners." CVPR’22