

Postdoc proposal

Advancing Exemplar-Free Class-Incremental Learning: Towards More Efficient, Stable, and Adaptive Models

Grégoire Petit, PhD

Introduction

Exemplar-Free Class-Incremental Learning (EFCIL) is the specialized subfield within the broader domain of Machine Learning and Artificial Intelligence, where a Machine Learning model needs to continuously adapt and expand its knowledge to accommodate new batches of classes, and named states, while retaining the ability to recognize previously learned classes.

The *exemplar-free* aspect of EFCIL distinguishes it from other incremental learning methods that rely on retaining specific examples or exemplars from previous classes to aid in learning new classes. In EFCIL, the model does not have access to stored exemplars, making the learning process more challenging, though easily embeddable.

EFCIL is a complex and challenging domain that requires innovative solutions to enable continual learning of new classes while preserving previously acquired knowledge. Several key challenges contribute to the complexity of EFCIL, each of which requires careful consideration and specialized approaches.

1 Summary of My PhD Work and Its Significance

1.1 Overview of My Research in EFCIL

In my PhD research, I focused on Exemplar-Free Class-Incremental Learning (EFCIL), a challenging yet crucial subfield of machine learning. My work was driven by the need to develop models that can continually adapt and learn new classes while retaining knowledge of previous classes, without relying on storing exemplars from those classes.

1.2 Key Challenges Addressed

1. **Memory Management:** I tackled the challenge of efficiently learning a large number of classes in a memory-constrained environment.
2. **Computational Requirements:** My research balanced the need for efficient learning of new classes with limited computational resources.
3. **Catastrophic Forgetting:** I addressed the issue of retaining information about past classes in the absence of exemplar revisits.
4. **Scenario Variability:** I optimized models for different scenarios, mostly focusing on initial state composition.

1.3 My Contributions

- **PlaStIL (Plastic and Stable Memory-Free Class-Incremental Learning):**¹ I developed PlaStIL, which balances plasticity and stability in CIL by freezing the feature extractor post-initial state and incorporating multiple model tops. This approach demonstrated exceptional performance across diverse configurations on large datasets.
- **FeTrIL (Feature Translation for Exemplar-Free Class-Incremental Learning):**² I introduced FeTrIL, a novel pseudo-feature generator using centroid representations of past classes. This method is not only computationally efficient but also effective in challenging data scenarios.
- **An Analysis of Initial Training Strategies for Exemplar-Free Class-Incremental Learning:**³ I established a framework for analyzing initial training strategies in EFCIL, providing insights into the impact of using the first batch of the target dataset versus pre-trained weights from an auxiliary dataset.
- **AdvisIL (A Class-Incremental Learning Advisor):**⁴ In collaboration with my colleagues, I co-developed AdvisIL, a method that recommends the most suitable combination of learning algorithm and neural architecture for class-incremental learning based on the characteristics of the incremental process. This approach simplifies the design of class-incremental learning systems by eliminating the need for extensive preliminary experiments.

1.4 Significance of My Work

My PhD research significantly advances the field of EFCIL by:

- Defining EFCIL, outlining its importance, and setting up its boundaries for the whole community.
- Proposing innovative solutions to the plasticity-stability dilemma, leading to models that are both adaptable and knowledgeable of past classes.
- Introducing PlaStIL and FeTrIL, 2 state-of-the-art methods, which address memory and computational efficiency challenges in EFCIL.
- Providing a new statistical analysis framework that enhances the understanding and optimization of initial training strategies for incremental learning models.

These contributions are pivotal in developing more efficient, stable, and adaptive machine learning models, essential for applications requiring continual learning capabilities.

¹**Grégoire Petit**, Adrian Popescu, Eden Belouadah, David Picard, and Bertrand Delezoide. 'PlaStIL: Plastic and Stable Memory-Free Class-Incremental Learning'. In: *Proceedings of The 2nd Conference on Lifelong Learning Agents*. Aug. 2023

²**Grégoire Petit**, Adrian Popescu, Hugo Schindler, David Picard, and Bertrand Delezoide. 'FeTrIL: Feature Translation for Exemplar-Free Class-Incremental Learning'. In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*. Jan. 2023.

³**Grégoire Petit**, Michael Soumm, Feillet Eva, Adrian Popescu, David Picard, and Bertrand Delezoide. 'An Analysis of Initial Training Strategies for Exemplar-Free Class-Incremental Learning'. In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*. Jan. 2024.

⁴Eva Feillet, **Grégoire Petit**, Adrian Popescu, Marina Reyboz, and Céline Hudelot. 'AdvisIL - A Class-Incremental Learning Advisor'. In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*. Jan. 2023.

2 Objectives of My PostDoc Research

Building upon the foundation laid during my PhD in Exemplar-Free Class-Incremental Learning (EFCIL), my PostDoc research aims to advance this field by addressing several critical challenges and exploring new frontiers. The overarching goal is to enhance the efficiency, adaptability, and stability of EFCIL models, making them more suitable for real-world applications.

2.1 Primary Objectives

1. **Improving the Feature Generalization:** Recent trends in the CIL community involve the embedding of a pretrained network, and demonstrate high performance, as confirmed by my PhD research. Building on this, multiple strategies can be inspired, *using this time ONLY the legitimate data* (i.e. data from the first state), to design more transferable EFCIL methods. This point will be the background task for every methods developed during my postdoc.
2. **Advancing Model Efficiency:** To develop methods that further reduce the computational and memory requirements of EFCIL models without compromising their learning effectiveness.
3. **Enhancing Model Stability:** To innovate techniques that further mitigate the issue of catastrophic forgetting, thus improving the model’s ability to retain knowledge of previously learned classes over extended periods.
4. **Improving Adaptability to New Classes:** To explore adaptive architectures and training methodologies that allow the model to seamlessly integrate new classes, particularly in diverse and dynamic environments.

2.2 Specific Research Directions

- **Novel Continual Learning Strategies:** Exploring new strategies in continual learning that enable the model to learn from a continuous stream of data without the need for retaining exemplars.
- **Efficient Transfer Learning Approaches:** Assessing how transfer learning can be more effectively leveraged in EFCIL, particularly focusing on the pretraining aspect to provide a robust foundation for subsequent incremental learning.
- **Application in Real-World Scenarios:** Applying EFCIL methodologies to practical applications, such as robotics or natural language processing, to validate and refine the models under new conditions.
- **Community Contribution and Benchmarking:** Aiming to contribute to the EFCIL community by developing open-source tools, libraries, and benchmark datasets, which are crucial for standardizing and accelerating research in this area.

2.3 Anticipated Challenges and Solutions

1. **Balancing Plasticity and Stability:** Continuously innovating techniques that address the plasticity-stability dilemma, ensuring that the model remains effective as it learns new classes while retaining old ones.
2. **Scalability and Resource Management:** Focusing on scalability issues and efficient resource management to enable the practical implementation of EFCIL models on long-tailed scenarios.

The objectives outlined for my PostDoc research are designed to build upon my PhD work, pushing the boundaries of what is possible in Exemplar-Free Class-Incremental Learning. By addressing these key areas, I aim to significantly contribute to the evolution of machine learning models that are not only more capable and efficient but also versatile enough to meet the demands of real-world applications.

3 Research Plan for My PostDoc Research

The research plan for my PostDoc is designed to build upon my PhD work in EFCIL, aiming to achieve the objectives laid out previously. This plan outlines a series of research activities, each with defined milestones, spread throughout the PostDoc period.

3.1 Timeline and Milestones

Note: The timeline is indicative and may be adjusted based on research progress and unforeseen challenges.

3.1.1 Months 1-3: Preliminary Research and Setup

- Conduct an extensive literature review to identify current trends and gaps in EFCIL.
- Set up computational infrastructure and finalize data sources.
- Begin networking with potential collaborators.
- *Milestone:* Completion of literature review and setup of the research environment.

3.1.2 Months 4-6: Advanced Continual Learning Strategies

- Develop novel strategies for continual learning without relying on exemplars.
- Conduct initial experiments to assess efficacy and efficiency.
- *Milestone:* Development and preliminary validation of new continual learning methods.

3.1.3 Months 7-9: Enhancing Transfer Learning Techniques

- Investigate improved transfer learning methods specifically tailored for EFCIL.
- Experiment with various pretraining and finetuning approaches.
- *Milestone:* Establishment of enhanced transfer learning techniques.

3.1.4 Months 10-12: Optimization and Resource Efficiency

- Focus on optimizing the computational efficiency of EFCIL models.
- Explore algorithmic improvements for reducing resource consumption.
- *Milestone:* Development of more resource-efficient EFCIL models.

3.1.5 Months 13-18: Application to Real-World Scenarios

- Apply developed models to practical use cases in domains such as NLP or computer vision.
- Gather and analyze feedback to refine models.
- *Milestone:* Successful application and refinement of models in real-world scenarios.

3.1.6 Months 19-24: Dissemination and Community Contribution

- Prepare research findings for publication in top-tier journals and conferences.
- Develop open-source tools or libraries based on the research.
- Engage in knowledge exchange through workshops or seminars.
- *Milestone:* Significant contribution to the EFCIL community and successful dissemination of research outcomes.

3.2 Anticipated Challenges and Adaptive Strategies

Throughout the research, I anticipate facing challenges related to model scalability, computational efficiency, and real-world applicability. To address these, I plan to:

- Regularly review and adapt research strategies based on ongoing results and feedback.
- Engage in continuous learning and collaboration with other researchers to overcome technical challenges.
- Consolidate the interactions I built in the Continual Learning community⁵

4 Expected Outcomes

- **Publications in Top-Tier ML/AI Conferences:**
 - Contribution to conferences such as CVPR, ECCV, ICCV, NeurIPS, etc.
 - Collaborations with lab colleagues, resulting in significant publications in the field.
- **Community Contributions:**
 - Development of open-source tools and libraries tailored to EFCIL.
 - Creation and maintenance of benchmark datasets to standardize research in EFCIL.
 - Aiming to accelerate research and collaboration within the EFCIL community.

5 Budget and Resources

High-Performance Computing is crucial for the success of my research for the following reasons:

- **Handling Large Datasets:** Long-tail EFCIL involves working with extensive datasets, requiring significant computational power to process efficiently.
- **Complex Model Training:** The development and training of advanced machine learning models, especially those in the transfer part of my proposal, demand substantial computational resources.
- **Experimentation and Analysis:** Running multiple experiments simultaneously and performing complex data analyses are computationally intensive tasks.

⁵ContinualAI