Open-world continual learning via knowledge transfer

Li, Y.

English Summary

This thesis explores and advances Open-World Continual Learning (OWCL), a new learning paradigm that addresses the core limitations of traditional deep learning in dynamic, uncertain, and non-stationary environments. Unlike conventional models that assume closed-world training with fixed task boundaries and complete label space, OWCL seeks to endow learning systems with the ability to continuously adapt to new tasks, detect and accommodate unknown classes, and retain prior knowledge without retraining from scratch. This paradigm fundamentally blends elements of continual learning, open-set recognition, and out-of-distribution generalization, and introduces new conceptual and algorithmic challenges at their intersection.

The theoretical importance of OWCL lies in its disruption of the classical i.i.d. learning assumption, motivating new mathematical formulations of decision spaces, transferability, and uncertainty. From a research perspective, OWCL redefines the structure of lifelong learning and opens new questions on how knowledge should be represented, stored, and adapted over time. From a practical perspective, OWCL is directly relevant to numerous real-world applications such as fraud detection, autonomous systems, clinical diagnostics, and recommendation systems—scenarios where the data evolves, classes appear or disappear, and full supervision is rarely guaranteed.

To systematically address the multifaceted challenges of OWCL, this thesis presents a coherent line of contributions centered around four proposed models, each designed to target a core bottleneck in the field:

- **Pro-KT (Prompt-enhanced Knowledge Transfer)** tackles the fundamental issue of balancing knowledge retention (stability) and plasticity in OWCL. It introduces a principled use of prompt learning to modularize and organize past knowledge into a reusable prompt space, enabling dynamic task adaptation. By leveraging a prompt repository that encodes both task-specific and task-agnostic information, Pro-KT enables flexible transfer across evolving tasks and achieves robust performance without catastrophic forgetting.
- **OFCL (Open-World Few-Shot Continual Learning)** expands OWCL into the low-resource regime, where labeled data is sparse or expensive. OFCL

identifies and mitigates three core challenges: representation collapse, ambiguity in open-set boundaries, and degraded transferability under supervision scarcity. Through a modular design incorporating contrastive learning, prototype replay, and boundary-aware training, our proposed model in this chapter, PEAK, demonstrates the feasibility and importance of scalable OWCL even under extreme label constraints, opening the door for real-world deployments where the annotation is costly or delayed.

- **HoliTrans (Holistic Knowns-Unknowns Knowledge Transfer)** formalizes OWCL at both a theoretical and algorithmic level. It begins by identifying four practically motivated OWCL scenarios that have not been rigorously captured in prior literature, and proposes a unified decision space formulation. Algorithmically, HoliTrans bridges continual learning with open-set detection using Nearest Class Mean (NCM) dynamics and transformer-based architectures. The model offers strong empirical performance while aligning with a theoretically grounded optimization formulation, serving as a comprehensive OWCL solution with provable generalization and boundary consistency.
- **CCL (Cross-regional Continual Learning for Fraud Detection)** demonstrates the practical significance of OWCL in real-world decision-making, specifically through the lens of financial fraud detection. By modeling distribution shifts across geographic regions and temporal phases, CCL introduces a graph-based continual learner that adapts to evolving fraud patterns without discarding prior experience. Its ability to generalize across regional boundaries highlights the robustness and deployability of OWCL principles in dynamic, mission-critical applications.

Collectively, these contributions not only advance the algorithmic frontier of OWCL but also serve to redefine the learning protocol under open-world conditions. The thesis advocates for a shift from monolithic, static models toward modular, adaptive, and transparent learning systems, capable of learning continuously, transferring knowledge purposefully, and responding intelligently to novel conditions.

The work concludes with several forward-looking research directions, including MoE-based modular architectures, PEFT-integrated federated OWCL systems, cross-modal learning under task heterogeneity, and explainable continual reasoning. These proposed directions reflect a broader aspiration: to build intelligent systems that support lifelong learning, that understand their limitations, and that serve real-world goals with robustness, transparency, and adaptability.