Visual Cognition Inspired Network for Few-Shot Fine-Grained Image Classification

Yongqi Liu*
Chongqing University, Chongqing, China

Yao Chen*
Chongqing University, Chongqing, China

Zhongpu Chen
Southwestern University of Finance and Economics, Chengdu, Sichuan, China

Zhi-Jie Wang†
Chongqing University, Chongqing, China

* denotes equal contribution
† denotes corresponding author (Zhi-Jie Wang, cszjwang@cqu.edu.cn)

Abstract

Few-shot fine-grained image classification (FSFGIC) is the task of developing computational models to recognize highly similar sub-categories from minimal training data. This task poses a core dilemma: over-emphasizing local details makes models vulnerable to noise and overfitting, while relying on global features often lacks the sensitivity required for fine-grained discrimination. Inspired by the human visual cognition strategy of dynamically balancing global and local information, we propose an Adaptive Global–Local Balance Network (AGLB-Net), a framework that computationally implements adaptive global– local integration for FSFGIC. AGLB-Net introduces two key modules: a Hierarchical Discriminative Feature Refinement (HDFR) module that progressively integrates representations from global semantics to fine-grained details, and an Adaptive Regional Re-Attention Module (ARRM) that automatically localizes and emphasizes discriminative regions without additional supervision. Extensive experiments on widely-used benchmarks demonstrate that AGLB-Net consistently achieves state-of-the-art performance across various few-shot settings, validating the effectiveness of human visual cognition-inspired adaptive global–local balancing in FSFGIC.

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*