

Title of Project:

(English)	Discrimination-Enhanced Generative Learning for Data-Scarce Visual Recognition
(Chinese)	用於數據稀缺視覺識別的歧視性增強生成學習

Abstract:

Visual recognition applications face difficulties in acquiring sufficient training data, while **generative learning** offers a potential solution by synthesizing surrogate samples for training. However, existing methods **face challenges** in (1) domain shift between synthetic features and real features, (2) insufficient discriminability between samples, and (3) under exploration of new unlabeled data after model deployment. These issues hinder the applicability of generative learning in real-world settings, where available annotated data are scarce.

This project aims to overcome these challenges by developing novel methods for generative learning in data-scarce scenarios. (1) We will develop a vision-guided dynamic semantic modeling approach to reduce domain shift. Existing feature synthesis methods mainly rely on class-level semantic information, while overlooking intra-class diversity and the guidance available from external knowledge. As a result, the synthesized features exhibit distributional discrepancy from real features. Our approach will integrate instance-specific semantic distributions and global domain knowledge (e.g., domain-specific descriptions), enabling the generator to better align synthetic features with real-world data. (2) We will enhance the discriminability of the visual recognition model by improving both the feature encoder and classifier. Insufficient discriminability results in feature confusion and recognition bias. Current approaches mainly optimize the generator to produce discriminative features, neglecting the roles of the encoder and classifier. To address this issue, we will simultaneously augment the discriminability for the visual encoder and classifier. By exploiting discriminative information from key regions of the input images, the entanglement of features will be reduced. A scoring-and-filtering mechanism will help the classifier reject low-quality synthetic features and encourage the generator to produce higher-quality features. This joint approach will improve feature separability and classification performance in visual recognition tasks. (3) We will design a test-time adaptation system to utilize new, unlabeled data after model deployment. Once deployed, models will face large amounts of new, unlabeled data that can enhance performance if properly utilized. We will implement a test-time adaptation approach to continually enhance the model's performance after deployment by leveraging new data. Additionally, to prevent catastrophic forgetting brought by the above continual training, we will introduce a data replay mechanism to retain important knowledge from previous domains.

This project aims to advance generative data-scarce visual recognition approaches by addressing key challenges in domain shift, discriminability, and continuous adaptation. This research could contribute to fields such as healthcare, autonomous systems, and environmental monitoring, where data are often scarce and adaptive models are essential.