Minimax Active Learning

Sayna Ebrahimi, William Gan, Kamyar Salahi, Trevor Darrell
UC Berkeley
{sayna, wjgan, kam.salahi, trevordarrell}@berkeley.edu
website link (TBD)

Abstract

Active learning aims to develop label-efficient algorithms by querying the most representative samples to be labeled by a human annotator. Current active learning techniques either rely on model uncertainty to select the most uncertain samples or use clustering or reconstruction to choose the most diverse set of unlabeled examples. While uncertainty-based strategies are susceptible to outliers, solely relying on sample diversity does not capture the information available on the main task. In this work, we develop a semi-supervised minimax entropy-based active learning algorithm that leverages both uncertainty and diversity in an adversarial manner. Our model consists of an entropy minimizing feature encoding network followed by an entropy maximizing classification layer. This minimax formulation reduces the distribution gap between the labeled/unlabeled data, while a discriminator is simultaneously trained to distinguish the labeled/unlabeled data. The highest entropy samples from the classifier that the discriminator predicts as unlabeled are selected for labeling. We extensively evaluate our method on various image classification and semantic segmentation benchmark datasets and show superior performance over the state-of-the-art methods.

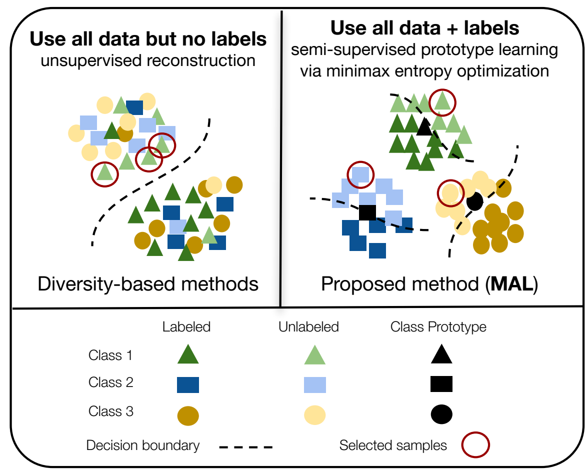

Figure 1: **Left:** Diversity-based task-agnostic active learning approaches perform unsupervised learning on all data and adversarially train a discriminator to predict whether to label a sample. The red-circled samples show a failure case for such methods, where the unlabeled samples closest to the discriminator's decision boundary are all from the same class and do not provide a diverse set of data to label. **Right:** Our proposed approach (MAL) performs semi-supervised learning on all data and uses the label information to learn per-class prototypes and then train a binary classifier to predict which samples to label. In contrast to the diversity-based approach, the selected red-circled samples are chosen because they are farther from the class prototypes, and as a result, come from different classes.

1. Introduction

The outstanding performance of modern computer vision systems on a variety of challenging problems [14] is empowered by several factors: recent advances in deep neural networks [11], increased computing power [10], and an abundant stream of data [33], which is often expensive to obtain. Active learning focuses on reducing the amount of human-annotated data needed to obtain a given performance by iteratively selecting the most *informative* samples for annotation.

Notably, a large research effort in active learning, implicitly or explicitly, leverages the uncertainty associated with the outcomes on the main task for which we are trying to annotate the data [66, 32, 18, 23, 68, 4]. These approaches are vulnerable to outliers and are prone to annotate redundant samples in the beginning because the task model is equally uncertain on almost all the unlabeled data, regardless of how representative the data are for each class [56]. As an alternative direction, [56] introduced a *task-agnostic* approach in which they applied reconstruction as an unsupervised task on the labeled and unlabeled data, which allowed for train-