Abstract:

Photos of everyday life are inherently multi-label in nature. Hence, multi-label classification is commonly used to analyze their content. In this talk, we discuss major challenges in multi-label classification and present an overview of three novel techniques that aim at tackling them. As a typical photo contains a few positive labels and many negative ones, a negative-positive imbalance may harm the optimization process. To this end, we introduce an asymmetric loss (ASL) which enables to dynamically down-weight easy negative samples. Second, large-scale multi-label classification datasets are commonly, and perhaps inevitably, partially annotated. That is, only a small subset of labels are annotated per sample. To handle partial annotation, we propose a selective approach that treats each class differently in the loss function. Finally, we deal with zero-shot learning for multi-label classification and propose an end-to-end model training that supports the semantic diversity of images and labels by using an embedding matrix with multiple principal embedding vectors.