Semantic Image Clustering

Abstract:

Can we automatically group images into semantically meaningful clusters when ground-truth annotations are absent? The task of unsupervised image classification remains an important, and open challenge in computer vision. Several recent approaches have tried to tackle this problem in an end-to-end fashion.

In this project, we deviate from recent works, and advocate a two-step approach where feature learning and clustering are decoupled. First, a self-supervised task from representation learning is employed to obtain semantically meaningful features. Second, we use the obtained features as a prior in a learnable clustering approach. In doing so, we remove the ability for cluster learning to depend on low-level features, which is present in current end-to-end learning approaches.

Existing System:

Neural net classifiers trained on data with annotated class labels can also capture apparent visual similarity among categories without being directed to do so. We study whether this observation can be extended beyond the conventional domain of supervised learning: Can we learn a good feature representation that captures apparent similarity among instances, instead of classes, by merely asking the feature to be discriminative of individual instances?

We formulate this intuition as a non-parametric classification problem at the instance-level, and use noise contrastive estimation to tackle the computational challenges imposed by the large number of instance classes. Our experimental results demonstrate that, under unsupervised learning settings, our method surpasses the state-of-the-art on ImageNet classification by a large margin.

Proposed System:

Image classification is the task of assigning a semantic label from a predefined set of classes to an image. For example, an image depicts a cat, a dog, a car, an airplane, etc., or abstracting further an animal, a machine, etc.

Nowadays, this task is typically tackled by training convolutional neural networks on large-scale datasets that contain annotated images, i.e., images with their corresponding semantic label. Under this supervised setup, the networks excel at learning discriminative feature representations that can subsequently be clustered into the predetermined classes.

What happens, however, when there is no access to ground-truth semantic labels at training time? Or going further, the?

The desired goal in this case is to group the images into clusters, such that images within the same cluster belong to the same or similar semantic classes, while images in different clusters are semantically dissimilar. Under this setup, unsupervised or self-supervised learning techniques have recently emerged in the literature as an alternative to supervised feature learning.

Software Tools:

1. VS Code
2. Jupyter Notebook
3. Anaconda
4. TensorFlow
5. Keras
6. Matplotlib
7. Scikit-learn
8. NumPy

Hardware Tools:

1. Laptop
2. Operating System: Windows 11
3. RAM: 16GB RAM
4. ROM: 5GB
5. Fast Internet Connectivity
6. GPU

Applications:

1. Image Clustering can be applied to all Facebook/Instagram Photos to classify the feed
2. Image Clustering can also be applied for medical datasets.