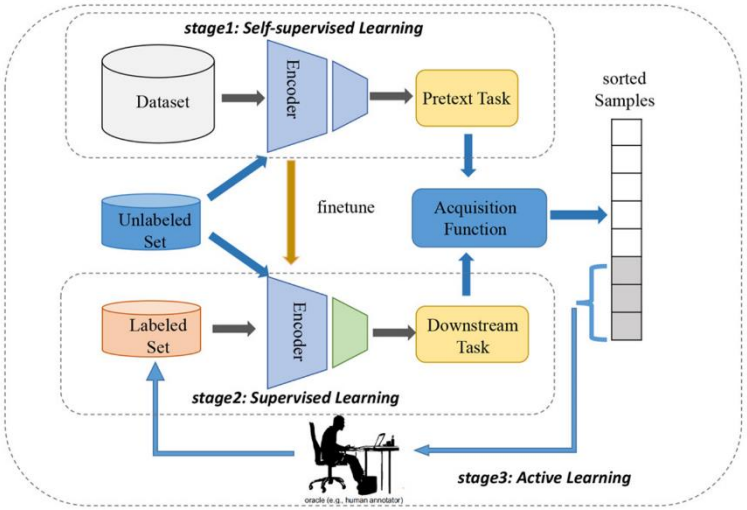
Deep Active Sampling with Self-Supervised Learning

Haochen SHI, Hui ZHOU

Frontiers of Computer Science, DOI: 10.1007/s11704-022-2189-z

Problems & Ideas

- Problems of combine self-supervised learning and active learning:
  - This method is difficult to effectively improve the model performance because it does not consider the feature representation performance of the examples on the pretext task.
- Ideas : A deep active sampling framework with self-supervised representation learning.



Algorithm 1: Deep active sampling algorithm based on Self-supervised representation learning

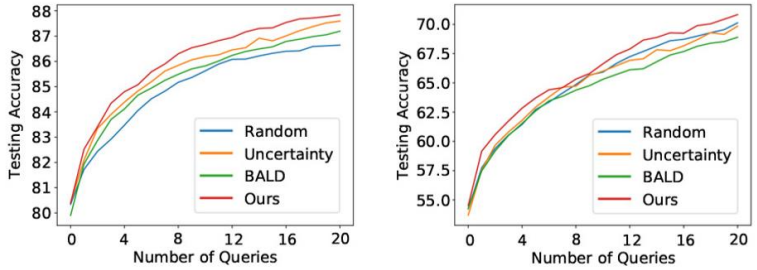
Input: Unlabeled data pool D_u, Labeled dataset D_l, Hyperparameter λ
Output: Target model f

Initialize the model f';
Train the model f' with pretext based on D_l and D_u;
Fine-tune the model f' with the downstream task based on D_l and get the draft target model f;
while *cost* < *budget* **do**
  Calculating scores for the each sample in D_u according to Eq.1;
  Select K samples with the largest value as Query set Q;
  Query for labeling of each samples in Q;
  Update $D_l = D_l \cup Q$;
  Update $D_u = \frac{D_u}{Q}$;
  Fine-tune f based on D_u;
end

Left: Deep active sampling framework with self-supervised learning; Right: the active sampling algorithm based on Self-supervised representation learning.

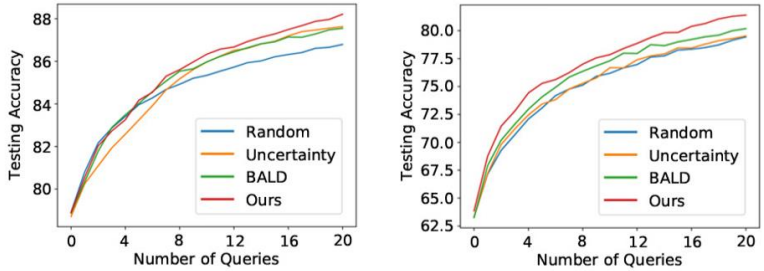
Main Contributions

- Contributions:
  - PROPOSE a deep active sampling framework based on self-supervised representation learning;
  - INITIATE a new active sampling strategy to fine-tune the self-supervised pre-training model on the downstream task;
  - EXPERIMENTS show that the proposed method is significantly superior to the state-of-the-art deep active sampling methods.



(a) fashion-MNIST (b) CIFAR10

Performance comparison of various methods when the self-supervised model is MoCo



(a) fashion-MNIST (b) CIFAR10

Performance comparison of various methods when the self-supervised method is to predict the image rotations