Report: 3/7

Ling Haoyang 519021910043

Learning the SimCLR

- proxy task: (not quite understand)
  - image colorization, jigsaw puzzles, Image in-painting, rotation prediction, instance discrimination
  - summary: context based, temporal based, and contrastive based (negative/positive pairs)
- the loss function uses the InfoNCE/NT-Xent/NT-Logistics
- cropping, color jittering
- larger batch sizes and longer training

Reading the paper GraphCL

- semi-supervised, unsupervised, and transfer learning as well as adversarial attacks.
- GNNs often have shallow architectures to avoid over-smoothing or information loss: res_gcn in semi-supervised_TU

GNN	CNN
node cropping	cropping
edge perturbation	cropping/color jittering
Attribute masking	color jittering
Subgraph (random walk/motif)	convolution?

- projection heads: similar to transformers? (gsimclr.py: self.proj_head)
- results:
  - Edge perturbation benefits social networks but hurts some biochemical molecules
    - solution: random walk to count a certain group of motifs?
  - Applying attribute masking achieves better performance in denser graph
  - fed into a down-stream **SVM** classifier
  - **Semi-supervised (10%) < Unsupervised** from the results?

Read the GraphCL Code

- read unsupervised_TU and run some tests on AIDS, NCI1, PROTEINS
- look part of the code in semi-supervised_TU: ResGCN and GAE remain to be understood