

Convolutional Neural Networks (CNNs): Improving Performance

Further Explore the Potential of Invariant Information
Clustering Using Video Datasets

Cai Shengqu

k1763920

Progress since the last meeting

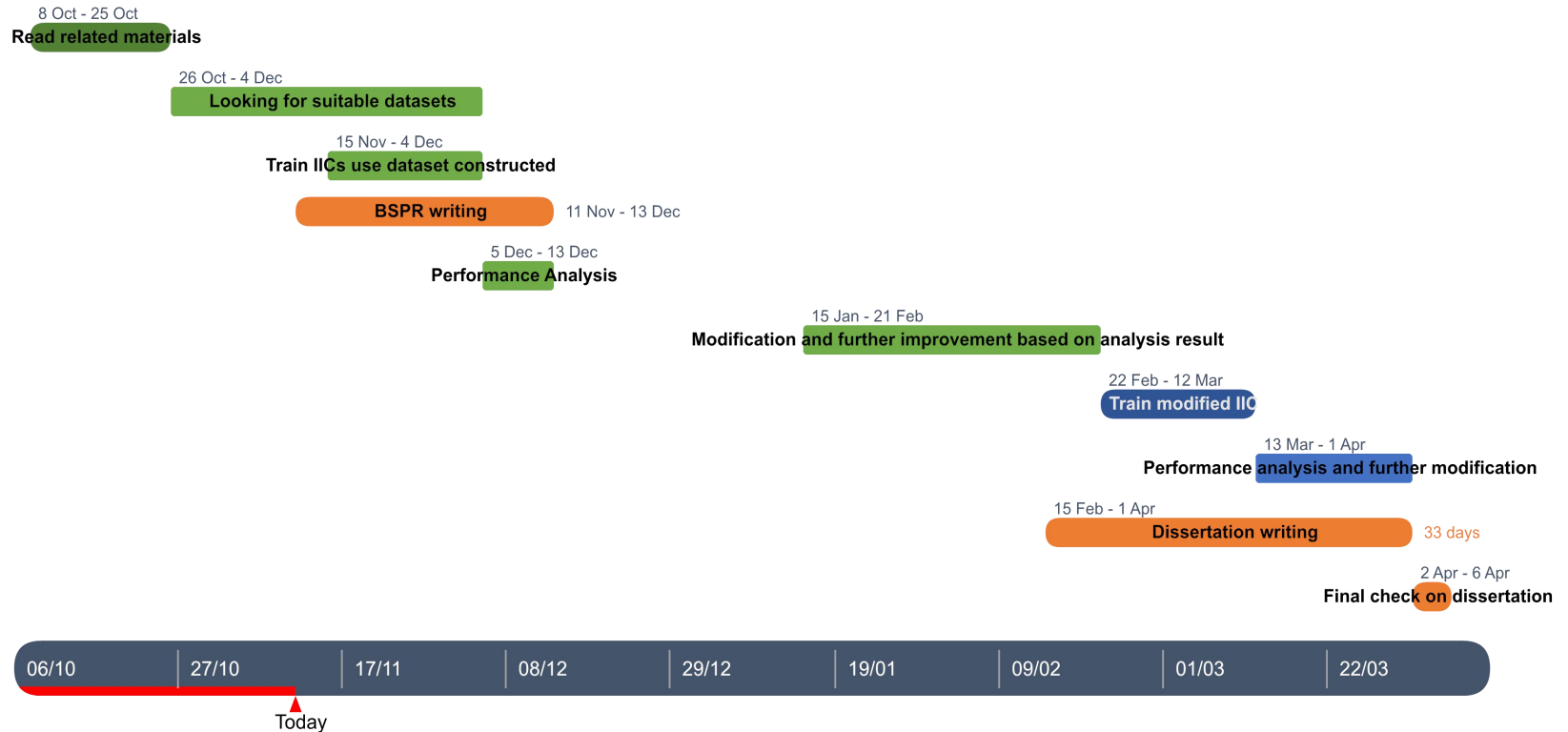
- Implement Youtube-BoundingBox in PyTorch.
 - Done.
 - csv annotations->local storage.
 - reason: efficiency, have to load the video while building dataset, which also takes much more memory.
 - videos->image of frames only.
 - reason: storage size. 4TB videos need to be fitted in maximum 40gb storage.
 - subset cutting?
 - 23 classes->less?
 - Implement related scripts.
 - Done(?).
 - Modified training scripts->yielding pairs instead of transform.
 - Need to be tested.
 - Perform the first training on TY-BB, gather performance data and evaluate further improvement space.
 - Scripts ready, hardware problems remains unsolved.

Current challenges and blocks

- Hardware requirement
 - disk space
 - 40gb home+20gb scratch for taught students on HPC.
 - memory
 - at least 22.9gb needed, 64gb available on HPC node with v100 gpu(?).
 - gpu
 - two nodes with v100 gpu available on HPC, others use k10 which is too old to compile with current Pytorch.
- Code reading
 - Un-explained parts.

My next steps

- Updated plan:



- Next steps:

- BSPR writing.
- Hardware confirmation.
- Perform initial training.

Reference list

Xu Ji, Joao F Henriques, and Andrea Vedaldi. Invariant information distillation for unsupervised image segmentation and clustering. *arXiv preprint arXiv:1807.06653*, 2018.

E. Real, J. Shlens, S. Mazzocchi, X. Pan, and V. Vanhoucke. Youtube-boundingboxes: A large high-precision humanannotated data set for object detection in video. *arXiv preprint arXiv:1702.00824*, 2017.