Convex Optimization for Enhancing CNN-Based Unsupervised Object Discovery

Huairu Chen
Electrical & Computer Engineering
University of Toronto
Toronto, Canada
huairu.chen@mail.utoronto.ca

Shiming Zhang
Electrical & Computer Engineering
University of Toronto
Toronto, Canada
shim.zhang@mail.utoronto.ca

brief description of the problem

Unsupervised object discovery and tracking in video collections rely heavily on Convolutional Neural Networks (CNNs) for feature extraction. However, raw CNN feature maps often contain noise, redundancies, and irrelevant background details, which can degrade object discovery performance. To improve feature quality, convex optimization techniques, such as sparse coding and total variation (TV) regularization, can be used. Sparse coding helps in learning compact and discriminative representations, while TV regularization smooths feature maps, reducing noise while preserving object boundaries. The problem is how to efficiently integrate these convex optimization techniques with CNN-based object discovery to enhance accuracy and robustness.

brief survey of existing approaches

CNNs extract features from video frames, but without additional constraints, these features can be noisy and contain unnecessary details. Traditional methods rely on unsupervised clustering or heuristic filtering for feature refinement.

And Sparse coding has been used to learn compact feature representations in image processing tasks, removing irrelevant features and improving discriminative power (Mairal et al., 2009). Low-rank factorization has been used to compress CNN representations, but its application in unsupervised object discovery remains limited.

TV regularization has been widely used in image denoising to suppress noise while preserving sharp edges (Rudin, Osher, and Fatemi, 1992). It has been applied in CNN feature smoothing but not extensively explored for object discovery in videos.

This project aims to integrate these convex optimization techniques into CNN-based object discovery to enhance feature extraction, improve robustness, and reduce false positives.

a list of references

[1] Mairal, J., Bach, F., Ponce, J., & Sapiro, G. (2009). Online Learning for Matrix Factorization and Sparse Coding. Journal of Machine Learning Research (JMLR).

[2] Rudin, L. I., Osher, S., & Fatemi, E. (1992). Nonlinear Total Variation Based Noise Removal Algorithms. Physica D: Nonlinear Phenomena.

[3] Chambolle, A., & Pock, T. (2011). A First-Order Primal-Dual Algorithm for Convex Problems with Applications to Imaging. Journal of Mathematical Imaging and Vision.

[4] Beck, A., & Teboulle, M. (2009). A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems. SIAM Journal on Imaging Sciences.

[5] Vo, H. V., Pérez, P., & Ponce, J. (2021). Large-Scale Unsupervised Object Discovery. NeurIPS 2021.