



THE UNIVERSITY OF
CHICAGO

DEPARTMENT OF STATISTICS

MASTER'S THESIS PRESENTATION

Xinyi Gu

Department of Statistics
The University of Chicago

Non-Contrastive Self-Supervised Learning with Video Frames

THURSDAY, February 9, 2022, at 2:00 PM
Zoom Meeting

ABSTRACT

This paper is an exploration of using non-contrastive self-supervised learning methods in the image classification field. We present the experiments using YouTube videos as training data. Many experiments have been done using ImageNet, CIFAR10, and CIFAR100 datasets as input, where the images are finely selected for each category, and downstream evaluation is done based on images of same type. Typical self-supervised learning methods are using artificial augmentation to create two versions of the same image that should be embedded into similar vectors. In contrast, this paper uses 16 videos (9 hours total length, a total of 33k frames) to test if treating nearby frames from videos as pairs could reduce the use of augmentation and still serve the purpose of learning representations. This idea was motivated by studying how infants discover the world since, when at home crawling, they could observe objects from varied angles and in various lighting conditions. The main result is that our method reaches 60.5% accuracy in downstream KNN classification with CIFAR10.