Review-4

Data-Efficient Image Recognition With Contrastive Predictive Coding

Growing advancements in unsupervised learning have given rise to data-efficient algorithms which commonly fall in the regime of self-supervised learning. These methods provide improved performance in comparison to supervised learning with their representations readily transferrable to downstream tasks. The work presents an improved version of the Contrastive Predictive Coding method (CPCv2) which makes use of latent representations for image recognition in a data-efficient manner. In comparison to CPCv1, CPCv2 demonstrates improved performance as a result of 5 improvements carried out to the training setting. These improvements aid in data-efficient learning when trained on 1% of ImageNet dataset. Furthermore, representations learnt by the model are finetuned to for efficient object-detection on the PASCAL-VOC 2007 dataset.

CPCv2 is a significant improvement in comparison to the previous counterpart. While the model makes use of pre-existing InfoNCE Loss, the training setup introduces 7 improvements which lead to data-efficient learning of images. (1) CPCv2 consists of increased model capacity with 46 residual blocks resulting in ResNet-161. (2) Batch normalization is replaced by Layer normalization as it does not harm downstream performance by eliminzating dependency between image patches. (3) Instead of carrying out bottom-up predictions of patches, CPCv2 increases the prediction directions to bottom, up left and right. (4) CPCv2 consists of patch-based augmentations such as color-dropping and color-transforms in order augment low-level variability. (5) Lastly, CPCv2 makes use of larger patch sizes in conjunction with other randomized patch-based augmentations.