DAAE: A Double Accelerating Contrastive Learning Framework

Zixuan Pan, Zihao Wei
04/26/2022

MICHIGAN ENGINEERING
UNIVERSITY OF MICHIGAN

Background & Motivation

Background & Motivation

- The inherent augmentation effect of resizing has been neglected.
- Directly use resize can have out of distribution problem / incoherence representation size problem.
- Contrastive Learning can make full advantage of augmentations.
- Inspired by Masked Autoencoder:
  - Masking be treated as a kind of strong blurring augmentation
  - Reconstruction target can be combined with contrastive models
- Downsampling is similar to masking, which can lower the images' dimension and speed up the training process.
- Vision Transformers can acquiring coherent-size representation from image of different shapes.

Methodology

Architecture - DAAE

- Asymmetric Contrastive Learning Framework
- Downsampling & Masking as augmentation
- Contrastive Learning Loss between two tracks
- Reconstruction Loss of the MAE track
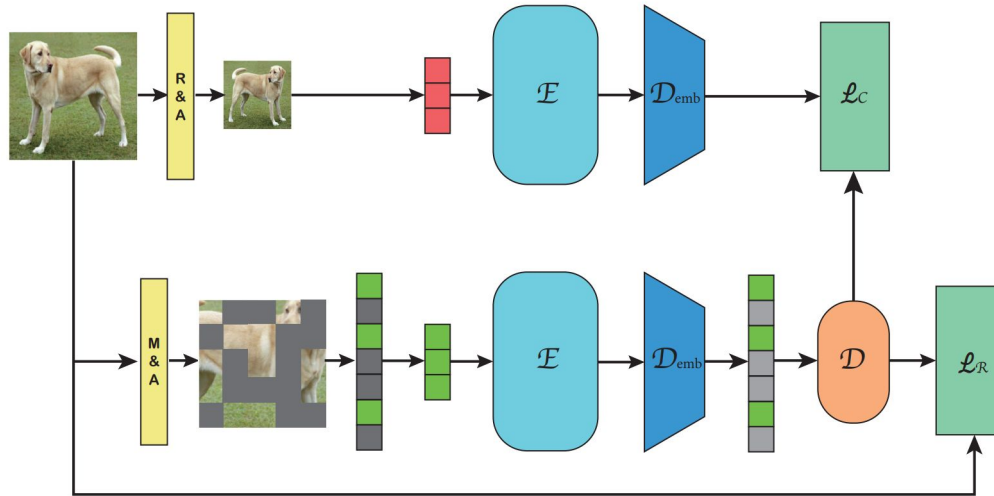
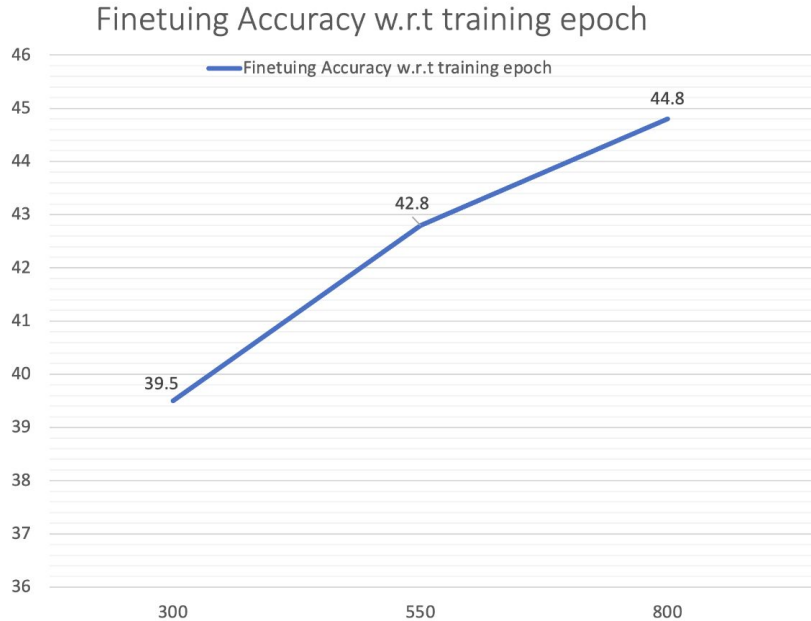# Experiment

ImageNet Classification

Faster & Better

Model	Top1 Acc	Epoch	Wall time
MAE	83.3	1600	2.05×
DAAE(ours)	**83.5**	600	**1×**

ImageNet-1K

Model	Top1 Acc	Epoch	Wall time
MAE	42.5	800	1.08×
DAAE(ours)	42.8	550	**1×**
DAAE(ours)	**44.8**	800	1.45×

ImageNet-mini (about 2% 1K)

ImageNet Classification

Finetuing Accuracy w.r.t training epoch

Model	Top1 Acc	Epoch	Wall time
MAE	42.5	800	1.08×
DAAE(ours)	42.8	550	**1**×
DAAE(ours)	**44.8**	800	1.45×

ImageNet-mini (about 2% 1K)

Ablation

Data Augmentation

Augmentation	Top1 Acc
resize only	38.9
resize + crop (0.2 - 1)	39.2
resize + crop + color jitter + grey scale	**39.5**
resize (0.08 - 1)	38.7

Feature Position

Augmentation	Top1 Acc
symmetric feature	39.4
asymmetric feature	**39.5**

Q & A

Thank you !