Learning Image Representations with SimCLR

David Bro Ludvigsen (s204102), Rasmus Steen Mikkelsen (s204135), Victor Tolsager Olesen (204141)

Introduction

Useful representations of images can be learned unsupervised, enabling models to solve tasks with much less labelled data than traditional supervised methods.

We used the SimCLR pretraining algorithm for this purpose on the CIFAR10 dataset

Learning Objectives

- Explore unsupervised representation learning
- Image classification with limited labels
- Semi-supervised-learning
- Train on different model sizes

Method

Pretraining method: SimCLR
Goal: Attract positive- and repel negative-pairs
Loss: Contrastive Loss

$$\ell_{i,j} = -\log \frac{\exp(\text{sim}(\mathbf{z}_i, \mathbf{z}_j)/\tau)}{\sum_{k=1}^{2N} \mathbb{1}_{[k \neq i]} \exp(\text{sim}(\mathbf{z}_i, \mathbf{z}_k)/\tau)}$$

Dataset : CIFAR10
60.000 images of resolution 3x32x32
50.000 for training and 10.000 for evaluation

Augmentations: random cropping, horizontal flip, color jitter (brightness, contrast, saturation and hue) and random gray scaling.

Models :
RexNet150, ResNet34, and ResNet50

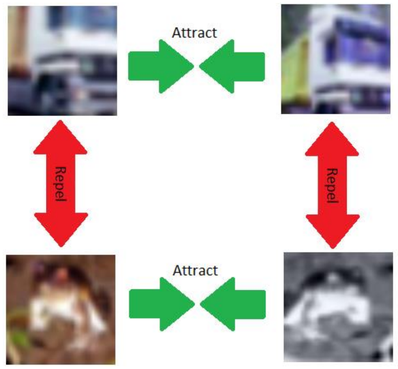
Each model architecture was trained in three ways:
- **Supervised:** Trained solely on the labelled data
- **Transfer:** Trained on the labelled data but the weights were initialized with transfer learning from a model trained on ImageNet
- **Semi-supervised:** Pretrained with SimCLR finetuned on labelled data

Pretraining for 500 epochs on the complete training set and evaluating with kNN (k=200) on the test set.

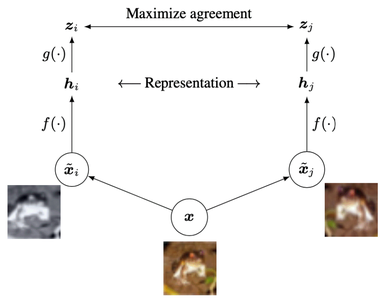
To fine-tune, we replaced the last linear layer with another linear layer where the output dimension match the number of classes (10).
Each model was trained supervised on (1%, 10%, 100%) of the training data.

Training Overview

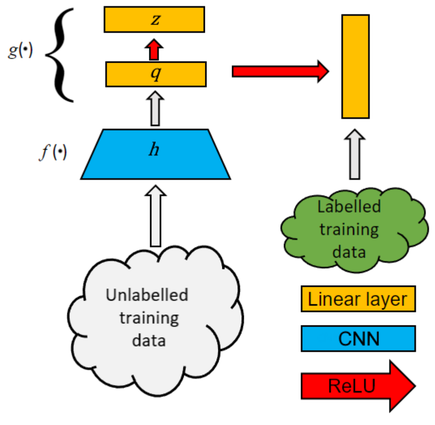
Pretraining Objective



Pretraining Implementation

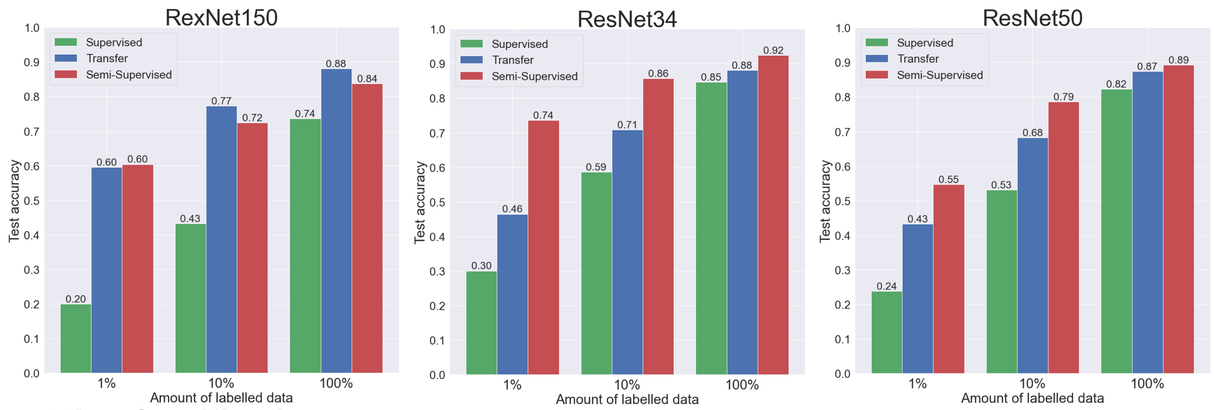


Model Workflow



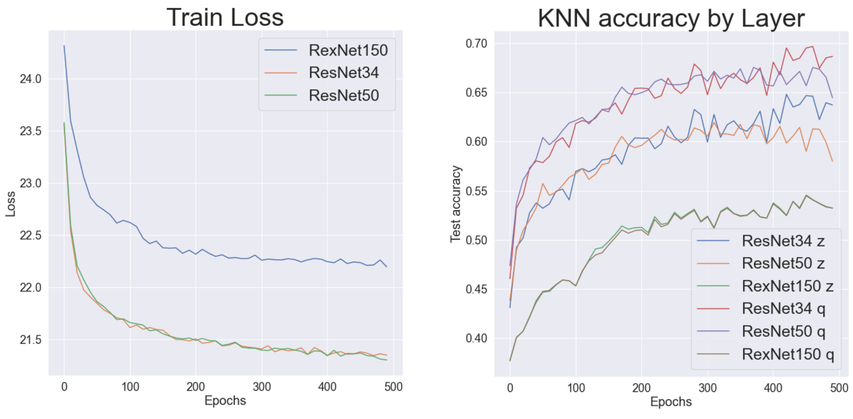
Results

Accuracies for different amount of labelled data



Missclassifications ResNet34

Pretraining Metrics



Model	RexNet150	ResNet34	ResNet50
#Parameters (M)	9.7	21.8	25.6

References

[1] T. Chen, S. Kornblith, M. Norouzi, G. Hinton, A simple framework for contrastive learning of visual representations, 2020
[2] T. Chen, S. Kornblith, K. Swersky, M. Norouzi, G. Hinton, Big Self-Supervised Models are Strong Semi-Supervised Learners, 2020
[3] K. He, X. Zhang, S. Ren, J. Sun, Deep Residual Learning for Image Recognition, 2015
[4] leftthomas, SimCLR, https://github.com/leftthomas/SimCLR, 2020