Self-Supervised Learning with In-Domain and Out-of-Domain Image Classification

ECE 285 – Project Presentation

Team
Murali M K Dandu
Srinivas Rao Daru

Agenda

- Introduction
- Method
- Datasets
- Experiments
- Conclusion

Introduction

Background

- Lack of labelled data and time consuming annotations
- Supervised models don’t fully learn generalized visual representations for any downstream tasks
- Dealing with few labelled scenarios

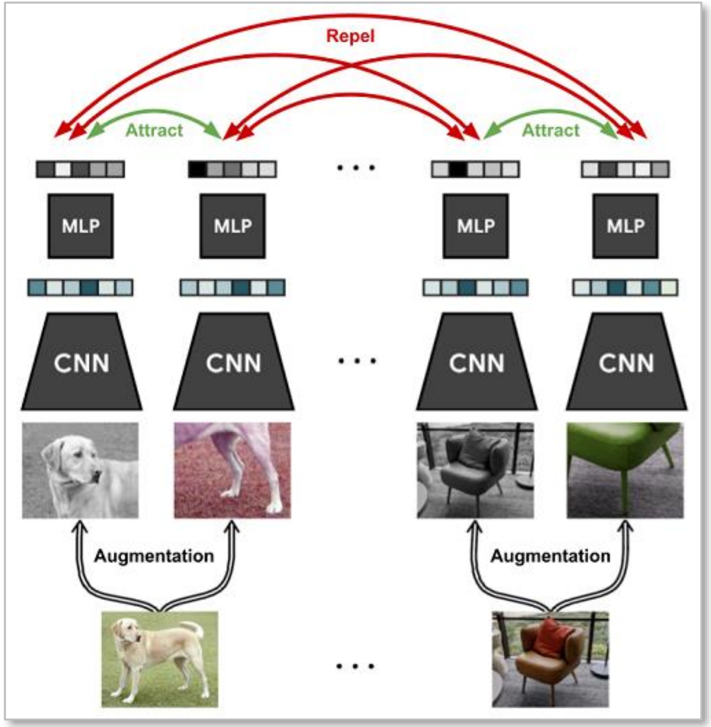
Solution and Plan

- Self-Supervised Learning
- Contrastive Learning framework using data augmentations - SimCLR
- Plan to test sampling variations and transfer learning robustness

Method

Method – SimCLR

Simple Framework of Contrastive Learning for Visual Representations



Batch	1	2	3	4	5	6	..	..	..	N
Augment 1	1a	2a	3a	4a	5a	6a	..	..	..	Na
Augment 2	1b	2b	3b	4b	5b	6b	..	..	..	Nb

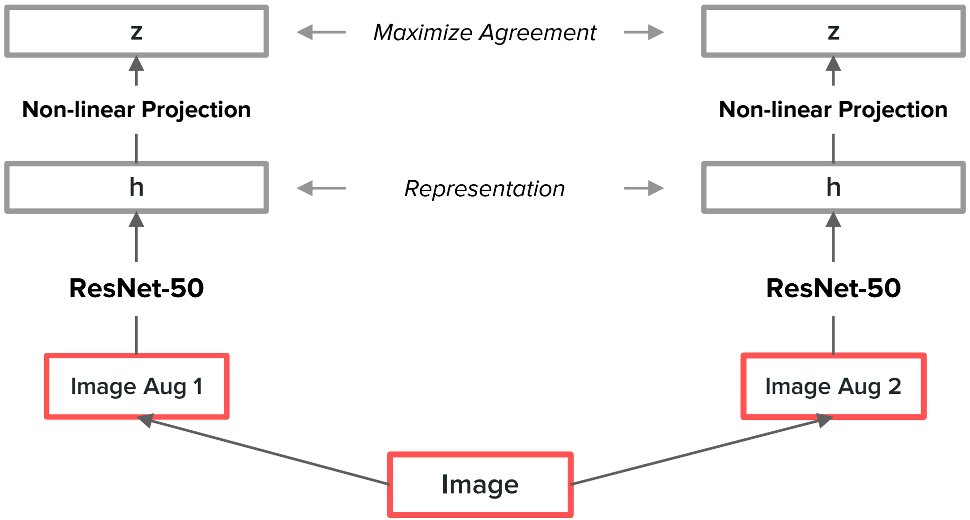
In a batch of size N:
For 1 positive pair, there are 2(N-1 negative pairs)

Loss for one positive pair (i, j):

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

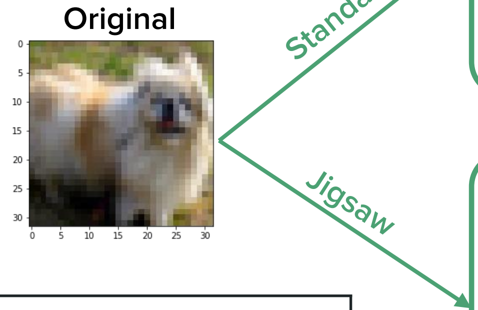
Loss is calculated across all positive pairs in a given batch

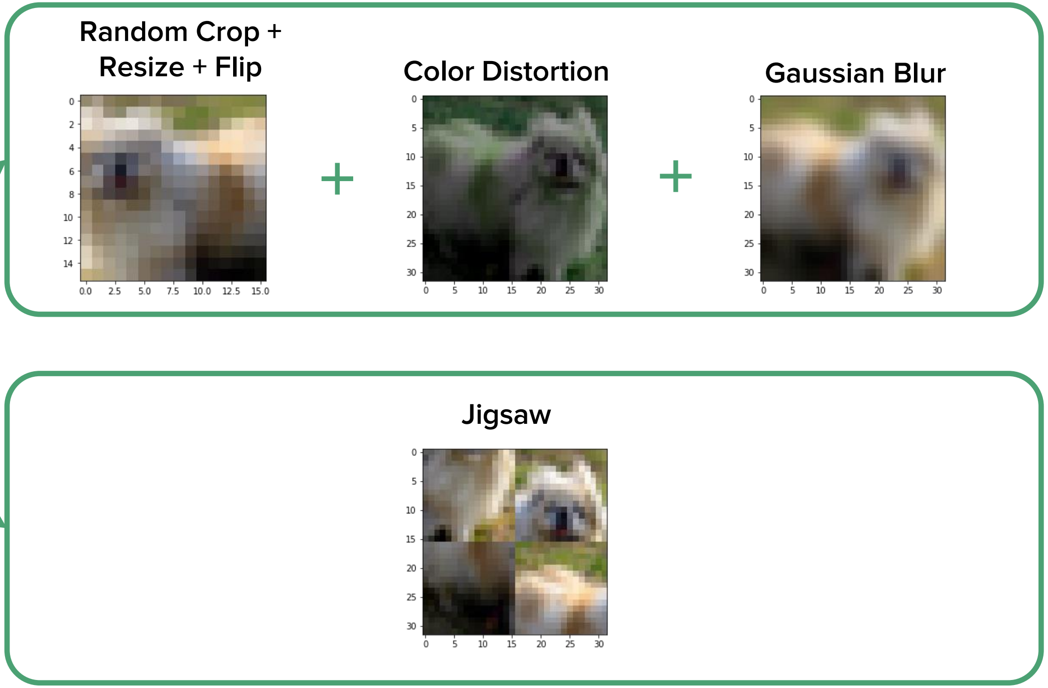
Method - Architecture



	Augmentation 1	Augmentation 2
V1	Standard	Standard
V2.1	**Jigsaw**	Standard
V2.2	Jigsaw	Standard + **Jigsaw**

Method – Data Augmention Strategies





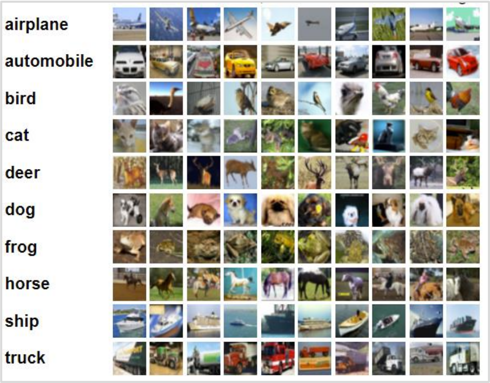
Model	Augment 1	Augment 2
V1	Standard	Standard
V2.1	**Jigsaw**	Standard
V2.2	Jigsaw	Standard + **Jigsaw**

Datasets

Datasets

- ## CIFAR-10

  60000 32x32 colour images in 10 classes, with 6000 images per class.



- ## Caltech-101

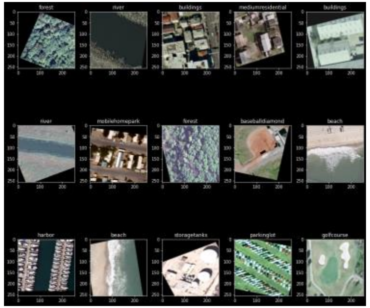
  101 classes, roughly 40-800 images totalling of 9k images (size 200-300)

- ## Oxford Pets

  37 category pet images with 200 images in each class

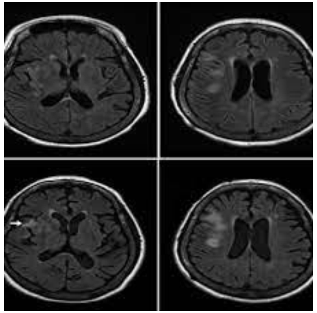
- ## Satellite Land Use

  21 classes scene classification with 100 images (256x256) per class



- ## Brain Tumor MRI

  Brain MRI scans of 7k images classified into 4 classes



Experiments

Experiments – Linear Evaluation

Evaluating linear-classifier freezing the pretrained representations

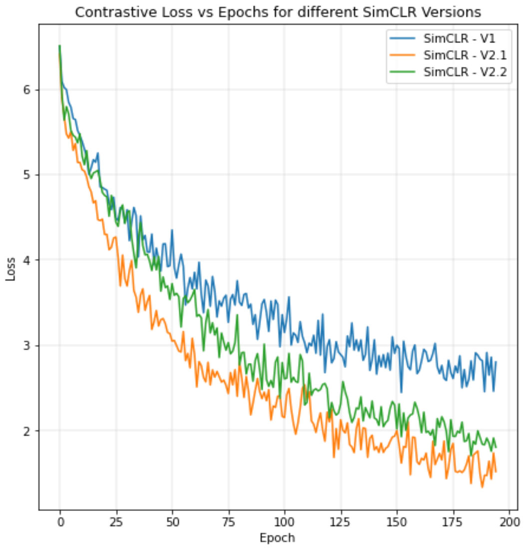
Method	Pretrained Data	Architecture	Top-1 Accuracy
Supervised	CIFAR-10	ResNet-50	82.1%
SimCLR v1 (Standard - Standard)	CIFAR-10	ResNet-50	58.3%
SimCLR v2.1 (Jigsaw - Standard)	CIFAR-10	ResNet-50	59.7%
SimCLR v2.2 (Jigsaw – Standard+Jigsaw)	CIFAR-10	ResNet-50	58.8%

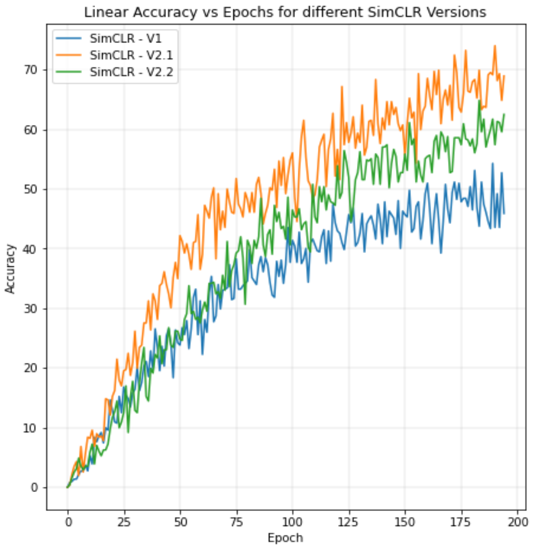
Given same architecture, SimCLR achieves ~60% accuracies on CIFAR-10

Experiments – Comparing different augmentations

Evaluating linear-classifier freezing the pretrained representations

The loss curves show better convergence upon adding jigsaw augmentation





Model	Augment 1	Augment 2
V1	Standard	Standard
V2.1	**Jigsaw**	Standard
V2.2	Jigsaw	Standard + **Jigsaw**

Experiments – Semi-Supervised Evaluation

Finetuning CIFAR-10 with few labels

		Fine Tuning Accuracy	
Method	**Pretrained Data**	**1% Labels**	**10% Labels**
Supervised Baseline	-	29.7%	43.8%
SimCLR v1	CIFAR-10	52.4%	70.5%
SimCLR v2.1	CIFAR-10	52.4%	69.9%
SimCLR v2.2	CIFAR-10	49.3%	67.2%

Note that dropout has been added in the models to prevent overfitting

Experiments – Transfer Learning

Method	**Caltech 101** (In domain)	**Oxford Pets** (natural images but out of domain for CIFAR10)	**Satellite Land Use** (out of domain)	**Brain Tumor MRI** (out of domain)
Linear Evaluation				
Supervised	40.6%	11.3%	35.2%	81.1%
SimCLR	49.0%	17.2%	52.1%	85.8%
Fine-Tuning				
Random Init	65.8%	18.3%	49.3%	96.5%
Supervised	76.3%	45.1%	63.1%	96.5%
SimCLR	73.0%	41.6%	58.8%	96.8%

*SimCLR performs **better** indicating better representation learning*

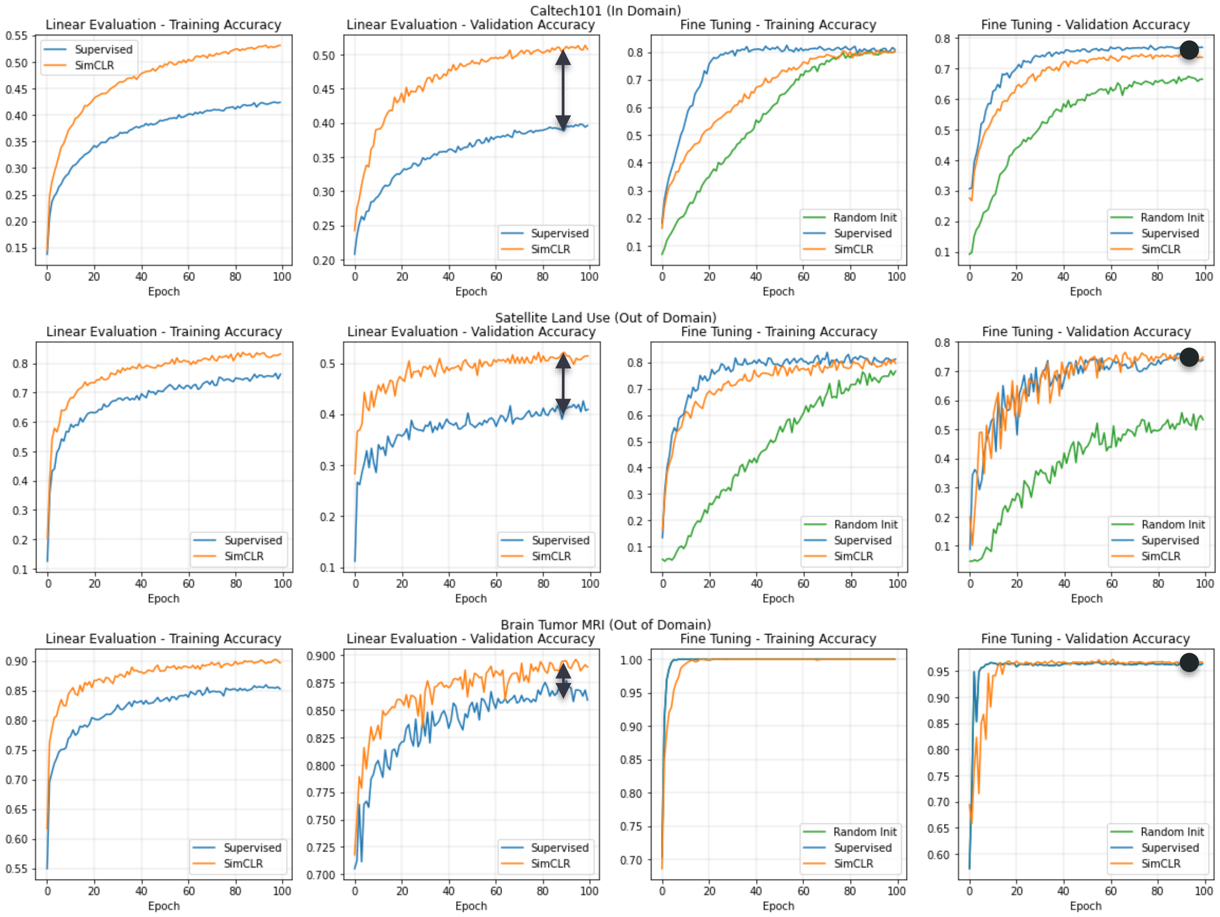
*SimCLR **almost matches** performance*

Experiments – Transfer Learning Deepdive

Caltech-101

Satellite Land Use

Brain Tumor MRI



*SimCLR has **better** validation performance in linear evaluation*

*SimCLR has **almost equal** performance in fine-tuning*

Key Takeaways

- *SimCLR helps learn good visual representations* ***without any explicits labels***
- ***Jigsaw augmentations*** *while training helps for* ***better convergence*** *and feature representations*
- *SimCLR can be used as an* ***effective semi-supervised*** *learning technique*
- *SimCLR representations generalize well in both in-domain and out-of-domain datasets, and can be used as a* ***common feature extraction model***
- *SimCLR* ***works inpar with supervised models*** *in standard transfer learning with fine-tuning approach*

Thank You!