ToddlerNet: Data Diversity vs View Diversity

Presenters: Kshama Nitin Shah, Kemmannu Vineet Rao
Motivation

▪ What is more important for representation learning data diversity vs view diversity?

▪ Toddlers learn with multiple views of a small number of objects

▪ What kind of learning objective do we need for such view-diverse category limited data?
Introduction

- Toddlers have wide-ranging knowledge of the world before learning to speak

- Question to address - better learning objective or better statistics in the form of inductive biases present in the learning data around them?

- Data determined by viewpoint, infant / toddler position, partial scene, curriculum based
Related Work

Infant curriculum learning

- Infant training set changes as the sensorimotor abilities of the infants develop
- Changing environments forms a curriculum

Related Work

Supervised Contrastive Learning

- Goal of contrastive learning: positive and negative pairs

Methodology: Datasets Used

ToyBox
Methodology : Datasets Used

ToyBox - Multiple Views
Methodology: Datasets Used
ToyBox vs CIFAR
Methodology: Supervised Learning

Pred_class: horse
Methodology: Supervised Contrastive Learning

(a) Supervised Cross Entropy

(b) Self Supervised Contrastive

(c) Supervised Contrastive

### Results: Quantitative

Supervised Learning
Classification Accuracies

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Model</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIFAR10</td>
<td>ResNet-18</td>
<td>92.68%</td>
</tr>
<tr>
<td>ToyBox</td>
<td>ResNet-18</td>
<td>97.56%</td>
</tr>
</tbody>
</table>
## Results: Quantitative

Supervised Contrastive Learning
Classification Accuracies

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Model</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIFAR10</td>
<td>ResNet-18+Linear Classifier</td>
<td>94.64%</td>
</tr>
<tr>
<td>ToyBox</td>
<td>ResNet-18+Linear Classifier</td>
<td>96.39%</td>
</tr>
</tbody>
</table>
Results: Qualitative

Supervised Learning

t-SNE

ToyBox Dataset

CIFAR-10 Dataset
Results: Qualitative

Supervised Contrastive Learning

t-SNE

ToyBox Dataset

CIFAR-10 Dataset
Conclusion

- ToyBox Dataset outperforms CIFAR10 for both learning objectives

- Supervised Contrastive learning objective outperforms Supervised Learning using Cross Entropy Loss

- Not just the data, but learning objective matters for learning that promotes generalizable performance
Thank You!