Counting Semantic Part Types of 3D Objects

Lucy Zhu *Panos Achlioptas** *Kaichun Mo** *Srinath Sridhar* *Leonidas J. Guibas*

Motivation

- Learn to **distinguish** the individual parts of 3D object point-clouds
- 3D models often have distinct parts **separable** in 3D modeling programs
- Same cannot be said for objects in the real world such as models obtained via 3D scanners
  - Manually model each piece
- Applications to computer vision and **3D Q&A**

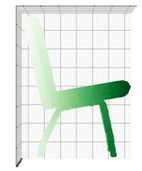
swivel chair

Approach

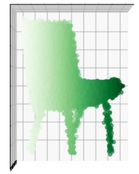
1. Prepare supervised dataset (whole/parts)
2. Deploy 2 distinct Deep Neural Networks
   i. Specializing in *single* part type
   ii. Generalizing across *multiple* parts
3. Train on different parts of 3D object point clouds

Dataset

- 3D Object Point Clouds (**ShapeNet**) with labeled parts (**PartNet**)
  - 2048 unique chair point clouds (split equally between testing and training dataset)



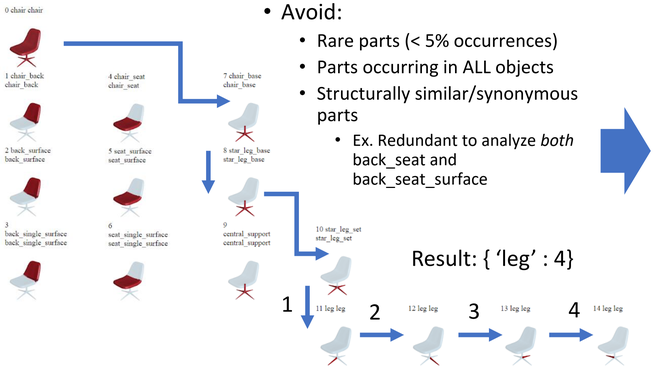
Pre-augmentation



Jitter and rotation applied.

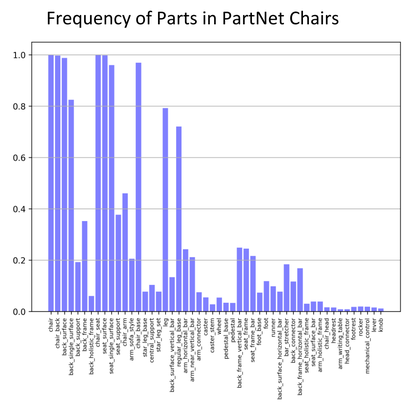
- Randomly select 2,500/10,000 points for input
- Data augmentation
  - Preserve invariance
  - Increase dataset size

Processing the Dataset



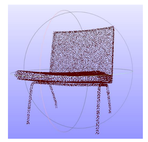
- Avoid:
  - Rare parts (< 5% occurrences)
  - Parts occurring in ALL objects
  - Structurally similar/synonymous parts
    - Ex. Redundant to analyze *both* back_seat and back_seat_surface

Result: { 'leg' : 4}



Model

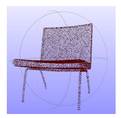
Specialized (specific part)



n x 3 → PointNet → output scores k_1 → log softmax k_1 → max → x

(x = number of occurrences of part)

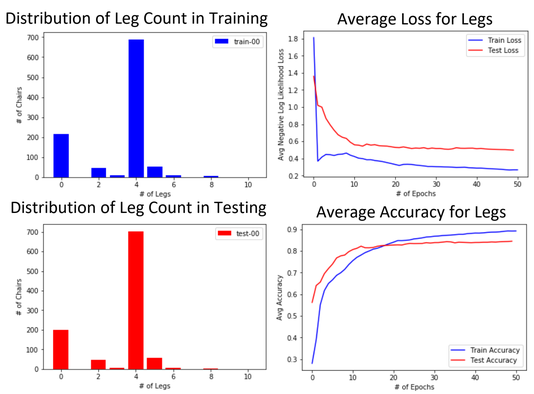
General (multiple parts)



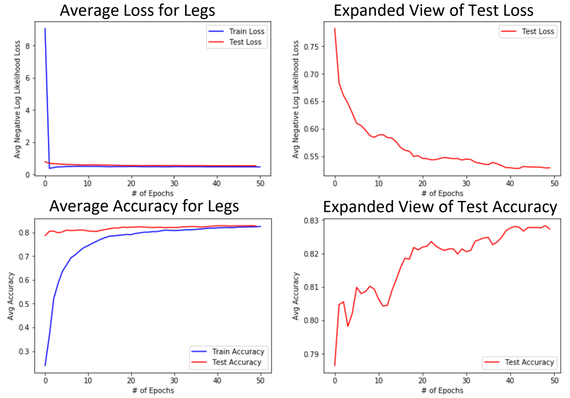
n x 3 + vector of parts counts (ground truth) → PointNet → 512 → 256 → linear layers → output scores (varying values for k): k_1, k_2, k_3, ..., k_x → log softmax → max → (n_x = part *x* occurs *n* times): n_1, n_2, n_3, ..., n_x

Experiment Results

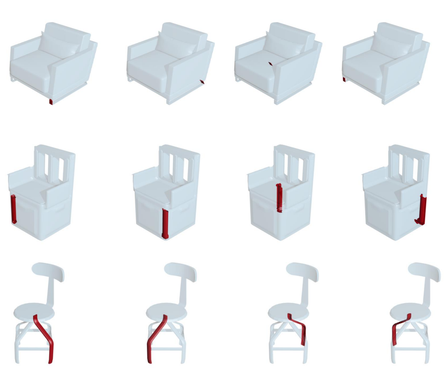
Specialized (specific part)



General (6 parts)



Failure Cases



Legs small in comparison to overall shape

(pred: 0 target: 4)

Legs are indistinct

(pred: 0 target: 4)

Complex leg model

(pred: 5 target: 4)

Parts Accuracy Results from each Deep Neural Network

Part Name	Specialized	General
back_frame_vertical_bars	0.776371875	0.8173828125
back_frame_horizontal_bars	0.718750	0.7734375
bar_stretcher	0.880859375	0.8115234375
chair_arm	0.95703125	0.8994140625
foot	0.8603515625	0.8984375
leg	0.859375	0.7900390625

Establishing correlations between different part structures potentially **enhances** the generalization capacity for *rarer* parts as evident in the general network.

Conclusions

- An approach that demonstrates the feasibility of solving the counting problem given proper part labels
- Specialized deep neural network not optimal when training dataset is sparse
- Possible future plans:
  - Point-cloud attention
  - Methods that reduce the amount of training data needed