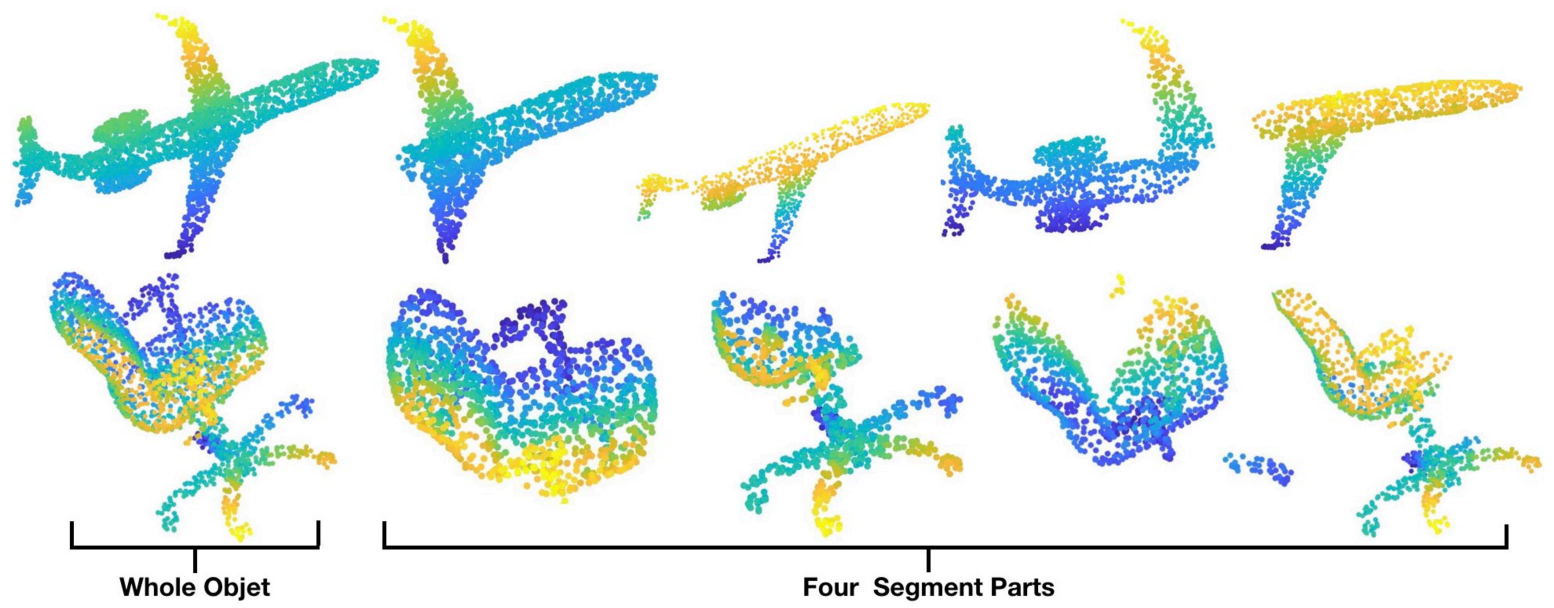
Unsupervised Feature Learning for Point Cloud by Contrasting and Clustering

Ling Zhang and Zhigang Zhu
The City College of The City University of New York

This research is supported by the National Science Foundation via awards #CNS-1737533 & #IIP-1827505, and by Bentley Systems, Inc. Additional travel support provided by the Defense Intelligence Agency via the Rutgers University Consortium for Critical Technology Studies.

3DV 2019
International Conference on 3D Vision
Québec City, Canada
September 16 - 19, 2019

Motivation



- Human can easily recognize objects and the locations of segments (parts) of the objects.
- We train GCNNs to learn features from unlabeled datasets by recognizing whether two segments are from the same object.

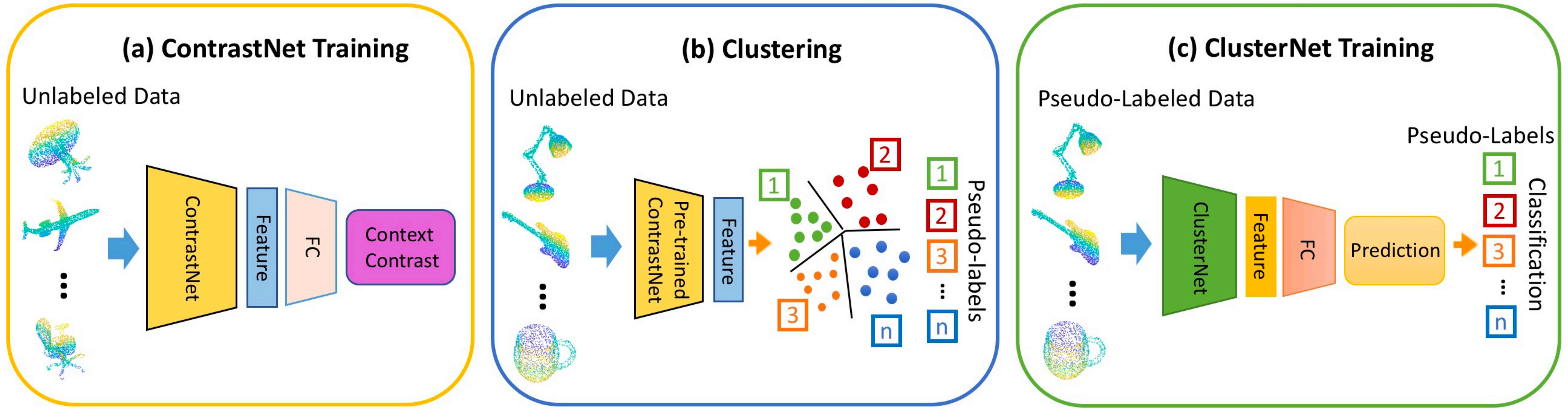
Achievements

- An unsupervised feature learning framework is proposed for point cloud data to learn semantic features.
- Our proposed approach outperforms most of the state-of-the-art unsupervised learning methods.

Several perspective views of 3D point cloud of an object. The views from different perspectives might be totally different.

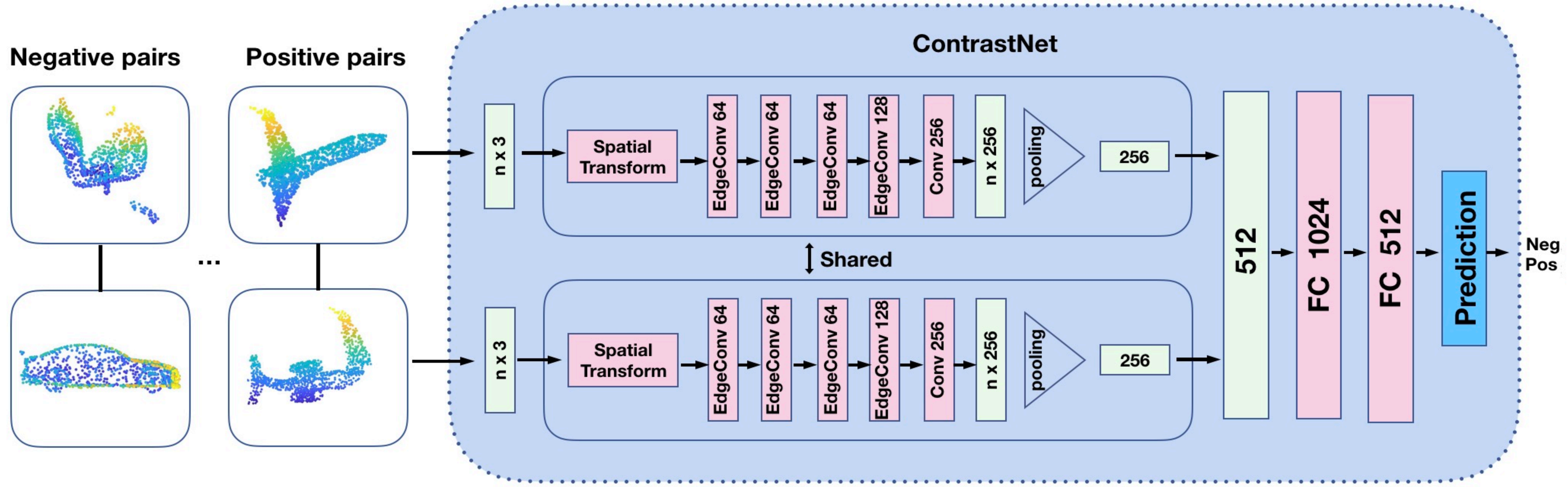
- Experiments show that our proposed approaches generate results that are practically useful for occluded objects and perspective views.

Unsupervised Feature Learning Pipeline



a) ContrastNet to verify whether two point cloud cuts belong to the same object;
b) Kmeans++ to cluster samples of 3D objects and assign cluster IDs, using the features learned from ContrastNet;
c) ClusterNet for object clustering learning, by training the network with pseudo-labels assigned by clustering.

Architecture of ContrastNet



- The positive pairs are generated by randomly sampling two segments from the same point cloud sample, while the negative pairs are generated by randomly sampling two segments from two different samples.
- A dynamic graph convolutional neural network (DGCNN) is used as the backbone network.
- The part contrast learning does not require any data annotations by humans.

Results for ContrastNet and ClusterNet

Training	Testing	ContrastNet(%)	ClusterNet(%)
ShapeNet	ModelNet40	84.1	86.8 (**+2.7**)
ShapeNet	ModelNet10	91.0	93.8 (**+2.8**)
ModelNet40	ModelNet40	85.7	88.6 (**+2.9**)

Clusters	ShapeNet ModelNet40	ShapeNet ModelNet10	ModelNet40 ModelNet40
100	86.2%	93.5%	87.7%
200	86.4%	93.6%	88.2%
300	**86.8%**	**93.8%**	**88.6%**
400	86.5%	93.2%	87.8%

Pre-training	Clusters	Clustering Acc.	Testing Acc.
ShapeNet	16	83.4%	86.1%
ModelNet40	40	64.2%	87.4%

Results for Occlusions and Perspective Views

Model	Acc. Full	Acc. Part	Acc. Perspective
ContrastNet	85.7%	79.4%	72.0%
ClusterNet	88.6%	82.4%	75.8%

Comparison with the State of the Art

Models	ModelNet40 (%)	ModelNet10 (%)
SPH [15]	68.2	79.8
LFD [7]	75.5	79.9
T-L Network [9]	74.4	-
VConv-DAE [30]	75.5	80.5
3D-GAN [36]	83.3	91.0
Latent-GAN [1]	85.7	**95.3**
FoldingNet [38]	**88.4**	94.4
ClusterNet (Ours)	***86.8***	***93.8***

Models	Acc.(%)	Models	Acc.(%)
PointNet [23]	89.2	DGCNN [35]	**92.2**
PointNet++ [25]	90.7		
PointCNN [17]	**92.2**	***ClusterNet(Ours)***	***88.6***