Towards Sampling Technique in Points Cloud

Hanchen Wang

Ph.D Candidate in
Engineering, University of Cambridge

August 11, 2019

Overview

1. Week Summary
2. project I - points cloud sampler
   - progress: sampling from meshes, normal calculations
   - next step: other task pipelines, new sampling techniques
3. project II - Latent Graph Neural Nets with Reinforcement Learning
   - progress: literature review, pipeline construction
   - next step: pipeline construction
4. project III - Fairness Metric Learning
   - progress: writing manuscripts for AAAI
   - next step: more analysis needed
5. project IV - Distributed Medical Imaging Recognition
   - progress: finish evaluation/visualization of trained seg model
   - next step: construct with a federate learning framework
6. misc
   - some new project thoughts
   - i really really want to attend ICCV in this October

week summary

- notified that the supervisor transition procedure has been completed
- received the reply from Charles @ Stanford Graphics Lab and Paul @ UCL Geometry Processing Group
  - get familiar with their sampling techniques, calculating normals and curvatures from points cloud
- not many experiments
- tools dev/literature survey on existing/new projects
- I really really want to go to ICCV2019

Benchmark of Last Week's Training Model

- no significant difference, but there is still space for improvement

batch	# of points	Gaussian Noise	Rotation	Accuracy	Comment
1	1024	None	None	85.7%	default setting
2	512	None	None	87.8%	
3	2048	None	None	88.4%	
4	1024	0.01 * $\mathcal{N}(0,1)$	None	**88.7%**	best in record
5	1024	0.05 * $\mathcal{N}(0,1)$	None	85.7%	
6	1024	0.5 * $\mathcal{N}(0,1)$	None	65.6%	
7	1024	None	identical rotation	**88.7%**	best in record
8	1024	None	random rotation	82.7%	

Table: benchmark on ModelNet40 Classification Task (Test Set)

- batch 4, 7 are not identical, different accuracy for the same class
- reported best accuracy in the paper was 89.2%, similar with ours
- it is better to align all objects with orientation before training

How PointNet/PointNet++ Preprocessing - Sampling

- for each object, either airplane or keyboard,
  - first uniformly sample 10,000 points from the **meshes**,
  - then apply random/farthest point/Poisson-Disk point sampling 512/1024/2048 points from these **points sets**,
  - the algorithms are described as following:

Algorithm 1 uniformly sampling from the meshes

---

meshes_area_sum = 0
for $i = 1$ to N_{faces} **do**
    meshes_area_sum += triangle_area({3 vertices of i-th face})
end for
probability_i = area_face_i/meshes_area_sum
for $j = 1$ to $N_{samples}$ **do**
    randomly choose face_j w.r.t $\{\text{probability\_i}\}_{i=1}^{N_{faces}}$
    sample a from $\mathcal{U}[0, 1]$, sample b from $\mathcal{U}[0, 1 - a]$
    point_j = a * vertex_1 + b * vertex_2 + $(1 - a - b)$ * vertex_3
end for

How PointNet/PointNet++ Preprocessing - Sampling

- now consider sampling from the points set with sufficient number of points

Algorithm 2 random sampling

```
for i = 1 to N_samples do
    random select one point from the set
end for
```

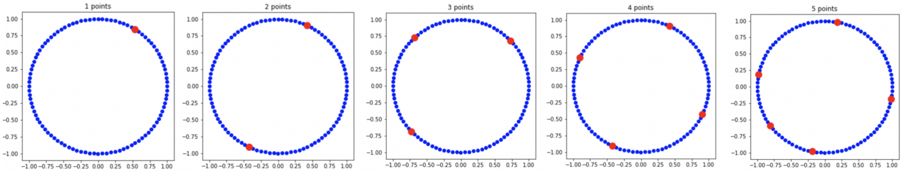
Algorithm 3 farthest point sampling

for $i = 1$ to $N_{samples}$ **do** in Algorithm 2 above and the steps below are typeset as:

- randomly choose one point p_i as a start, add p_i to the selection set $\mathbb{S}$
- **for** $i = 2$ to $N_{samples}$ **do**
    - **for** points $\notin \mathbb{S}$ **do**
        - choose $p_i = \arg\max_{p_j} \sum_{p_k \in \mathbb{S}} d(p_j, p_k)$
    - **end for**
    - add p_i to the set $\mathbb{S}$
- **end for**

How PointNet/PointNet++ Preprocessing - Sampling

- the visualization procedure of *Farthest Point Sampling*
  - take 2d circle as illustration, the distance metric $d(\cdot, \cdot)$ is Euclidean



- extended thought #1: **sample faster**,
  - $O(N^2)$ for sampling single point, for 10k points, sampling 1k point requires 10^{11} floating point operations,
  - for single CPU, its performance is around 100G ($\approx 10^{11}$)flops(floating point operations per second), which is compatible
  - maybe it becomes a bottleneck when it comes to application scenario which has requirement for fast sampling(such as fast scene detection in autonomous driving etc) or the sampling size is a bit larger

How PointNet/PointNet++ Preprocessing - Sampling

- extended thought #2: **sample more efficiently**,
    - use smaller number of points to achieve similar performance on classification/segmentation/detection tasks
    - novel methods for sampling, it is possible to apply MCMC/Gibbs sampling on the points set???
        - basically such sampling techniques are used to estimate the posterior, so it is possible to deploy them to get a representation of objects
    - maybe with a more complicated architecture, such as latent graph neural nets[1]
- what I have done:
    - incorporate these sampling methods into the Python open source module: *pyntcloud*(500+ stars, 100+ forks)
- what to do next:
    - investigate on more sampling methods

[1] LatentGNN: Learning Efficient Non-local Relations for Visual Recognition, ICML 2019
[2] https://github.com/daavoo/pyntcloud

How PointNet/PointNet++ Preprocessing - Normals

- for the normal calculations,
  - basically for points sampled within faces, their assigned normals is identical with that triangular faces
  - for points on the edges/vertices, the normals are the area-weighted sum of the faces containing that point
- what to do next:
  - integrate these calculations into *pyntcloud*
  - use analytical geometry models(with ground truth of surface normals) to testify the effectiveness of this normal calculation procedure(corporation with Hugo), as discussed last week

[1] https://github.com/daavoo/pyntcloud

Points Cloud Sampler - Future Work

- more sampling techniques
- normal calculation testify
- go through pipelines for other tasks with Points Cloud as well:
  - 3D Object Classification(completed)
  - 3D Object Part Segmentation(uncompleted)
  - Semantic Segmentation in Scenes(uncompleted)
  - 3D object detection in scenes(uncompleted)

[1] LatentGNN: Learning Efficient Non-local Relations for Visual Recognition, ICML 2019

[2] https://github.com/daavoo/pyntcloud

LatentGNN: progress

- literature review on the previous work
  - **Attention Mechanism**: GCNet(arXiv 2019); An Empirical Study of Spatial Attention...(arXiv 2019); Attention Augmented Convolutional Networks(arXiv 2019); Self-Attention with Relative Position Representations(NAACL 2018); Attention Is All You Need(NIPS 2017)
  - **Object Detection**: ResNet(CVPR 2016); Mask R-CNN(ICCV 2017); Feature Pyramid Network(CVPR 2017)
  - **autoML, RL and GNN**: Neural Architecture Search with Reinforcement Learning(ICLR 2017); latentGNN(ICML 2019); SkipNet(ECCV 2018)
- basically finish constructing the model architecture:
  - ResNet(backbone)
  - latent GNN(feature learner)
  - skip mechanism with RL(computation efficient)
- me as co-principal contributor, in collaboration with
  - Yutong Bai, Prof. Alan Yuille @ JHU,
  - Prof. Xuming He @ ShanghaiTech

LatentGNN: future work

- understanding ResNet, Mask-RCNN, LatentGNN by running their pipelines on the object detection tasks respectively
- finish developing the architecture of the model, hopefully get early stage results(performance in object detection tasks)

Fairness Metric Learning

- **progress**: writing manuscripts for AAAI 2020(ddl: 5th Sep)
- **next step**: more analysis needed
- me as principal contributor, with Dr. Adrian Weller et al.

Federated Medical Imaging Detection

- **progress**: finish evaluation/visualization of trained segmentation model(handcrafted input + deeplab v3+)
- **next step**:
  - try some other developed models, hopefully result in better performance(higher mIOU, current: 0.3 - 0.4)
  - construct federated learning framework
  - literature research:
    - Link IJCAI-19 Workshop on Federated Learning(just ended)
    - Link NeurIPS 2019 Workshop on Federated Learning(ddl: Sep 9)
    - Link IEEE Big Data 2019 Special Track on Federated Learning
- me as principal contributor, with panakeia.ai

new thoughts

- 3D segmentation
  - there are some SOTA segmentation techniques in 2D, such as dilated Conv(ICLR 2016), Dilated ResNet(CVPR 2017), Learning localized generative models for 3D points cloud via graph convolution(ICLR 2019)
- Learning from the graph
  - one of KDD 2019's hottest topics, there should be lots of fresh papers
  - How Powerful are Graph Neural Networks?(ICLR 2019 oral)

I really really want to attend ICCV2019

my argument:

- though it might be expensive
  - registration fee £170, air tickets £500, accommodation £450, visa £31
  - discounted since I am an IEEE Young Professional and Student Member
- it can be covered from my scholarships
  - Xuxin fellowship from my alma mater
- held once every two years, a great opportunity to meet with people, seek for collaborators and discussion, I have already connected with:
  - Yutong Bai and Prof. Yuille(Cambridge Alumni) @ JHU
  - Dr. Charles Ruizhongtai Qi @ Facebook AI, Stanford Graphics Lab
  - Yongcheng Liu @ Chinese Academy of Science
  - ...
- I am confident to have some results on the points cloud project by then, it will be beneficial to discuss this with them as well

I really really want to attend ICCV2019

- there are lots of events matching my research interests
- interested workshops:
  - Link Statistical Deep Learning in Computer Vision
  - Link Visual Recognition for Medical Images
  - Link 3D Reconstruction in the Wild
  - Link Autonomous Driving(Alex Kendall will be there as well)
  - Link Recovering 6D Object Pose
  - Link Geometry Meets Deep Learning
- interested tutorials Link:
  - Global Optimization for Geometric Understanding with Provable Guarantees
  - Holistic 3D Reconstruction: Learning to Reconstruct Holistic 3D Structures from Sensorial Data
  - Visual Recognition for Images, Video, and 3D
  - 3D Deep Learning and Applications in Autonomous Driving
  - Second- and Higher-order Representations in Computer Vision
  - Visual Learning with Limited Labeled Data

I really really want to attend ICCV2019

- interested long papers:
- Link Deep Non-Rigid Structure from Motion(Oral)
- Link Scalable Place Recognition Under Appearance Change for Autonomous Driving(Oral)
- Link Molding Humans: Non-parametric 3D Human Shape Estimation from Single Images
- Link Permutation-invariant Feature Restructuring for Correlation-aware Image Set-based Recognition
- Link Pixel2Mesh++: Multi-View 3D Mesh Generation via Deformation
- Link GP2C: Geometric Projection Parameter Consensus for Joint 3D Pose and Focal Length Estimation in the Wild
- Link Neural 3D Morphable Models: Spiral Convolutional Networks for 3D Shape Representation Learning and Generation
- Link Metric Learning with HORDE: High-Order Regularizer for Deep Embeddings
- Link AutoGAN: Neural Architecture Search for Generative Adversarial Networks

I really really want to attend ICCV2019

- interested long papers:
- Link Multi-Angle Point Cloud-VAE: Unsupervised Feature Learning for 3D Point Clouds from Multiple Angles by Joint Self-Reconstruction and Half-to-Half Prediction
- Link FrameNet: Learning Local Canonical Frames of 3D Surfaces from a Single RGB Image(Oral)
- Link Learning to Reconstruct 3D Manhattan Wireframes from a Single Image
- Link PointFlow : 3D Point Cloud Generation with Continuous Normalizing Flows(Oral)
- Link 3D-RelNet: Joint Object and Relational Network for 3D Prediction
- Link 3D Point Cloud Learning for Large-scale Environment Analysis and Place Recognition
- Link Joint Monocular 3D Detection and Tracking
- Link Deep Hough Voting for 3D Object Detection in Point Clouds(Oral)
- Link Variational Adversarial Active Learning(Oral)
- ...

The End