Point Cloud Transformer

CS 323 Project 4

Overview

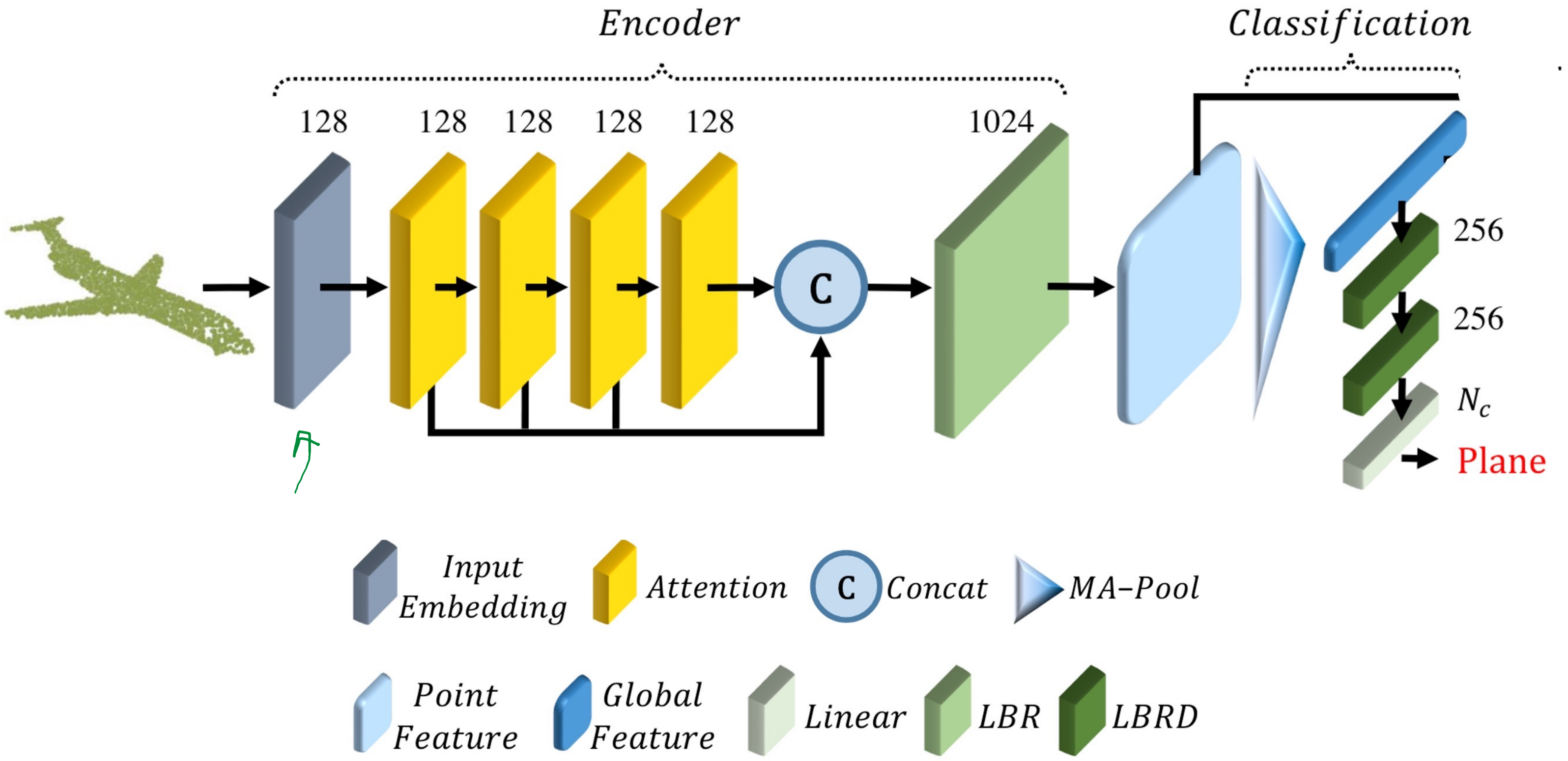
- In this project, you will implement PCT, an attention-based object classifier for 3D point clouds

- The requirements will be as follows:
  - Implement model 1: Naive PCT
  - Implement model 2: PCT with Offset-Attention
  - Implement model 3: PCT with Offset-Attention and Local Neighbor Embedding layer

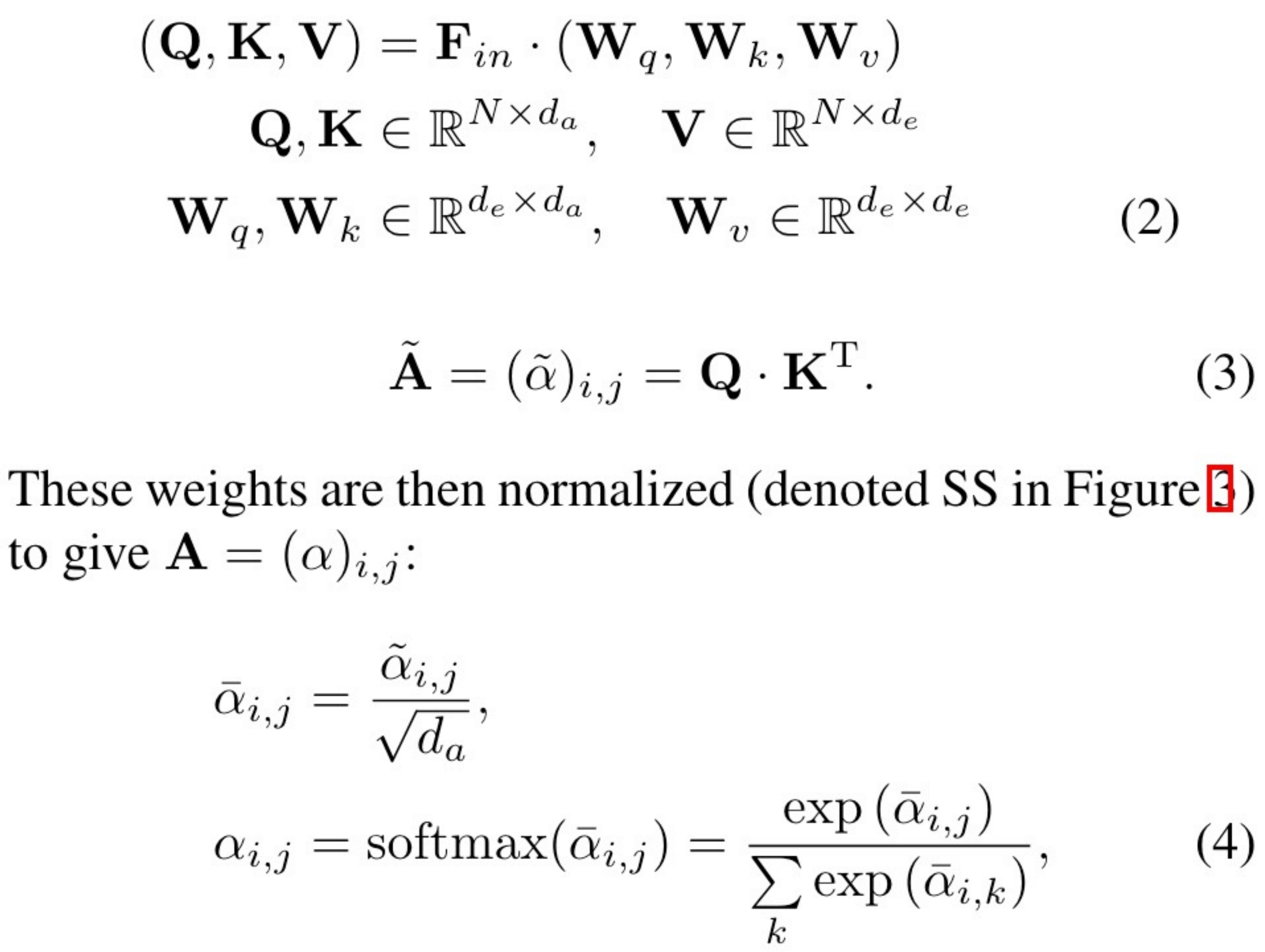
Dataset: ModelNet40

- 40 categories (such as airplane, car, plant, lamp)
- 9,843 meshes will be used for training
- 2,468 meshes will be used for validation
- The point cloud data points are uniformly sampled from the mesh surfaces
- You will be provided the point cloud data

PCT


Encoder
Classification
128
128
128
128
128
1024
C
256
256
N_c
Plane
Input Embedding
Attention
C
Concat
MA–Pool
Point Feature
Global Feature
Linear
LBR
LBRD

Self-Attention Layer



$$\begin{aligned}(\mathbf{Q},\mathbf{K},\mathbf{V}) &= \mathbf{F}_{in}\cdot(\mathbf{W}_q,\mathbf{W}_k,\mathbf{W}_v)\\ \mathbf{Q},\mathbf{K}\in\mathbb{R}^{N\times d_a}, &\quad \mathbf{V}\in\mathbb{R}^{N\times d_e}\\ \mathbf{W}_q,\mathbf{W}_k\in\mathbb{R}^{d_e\times d_a}, &\quad \mathbf{W}_v\in\mathbb{R}^{d_e\times d_e}\end{aligned} \tag{2}$$

$$\tilde{\mathbf{A}} = (\tilde{\alpha})_{i,j} = \mathbf{Q}\cdot\mathbf{K}^{\mathrm{T}}. \tag{3}$$

These weights are then normalized (denoted SS in Figure 3) to give $\mathbf{A} = (\alpha)_{i,j}$:

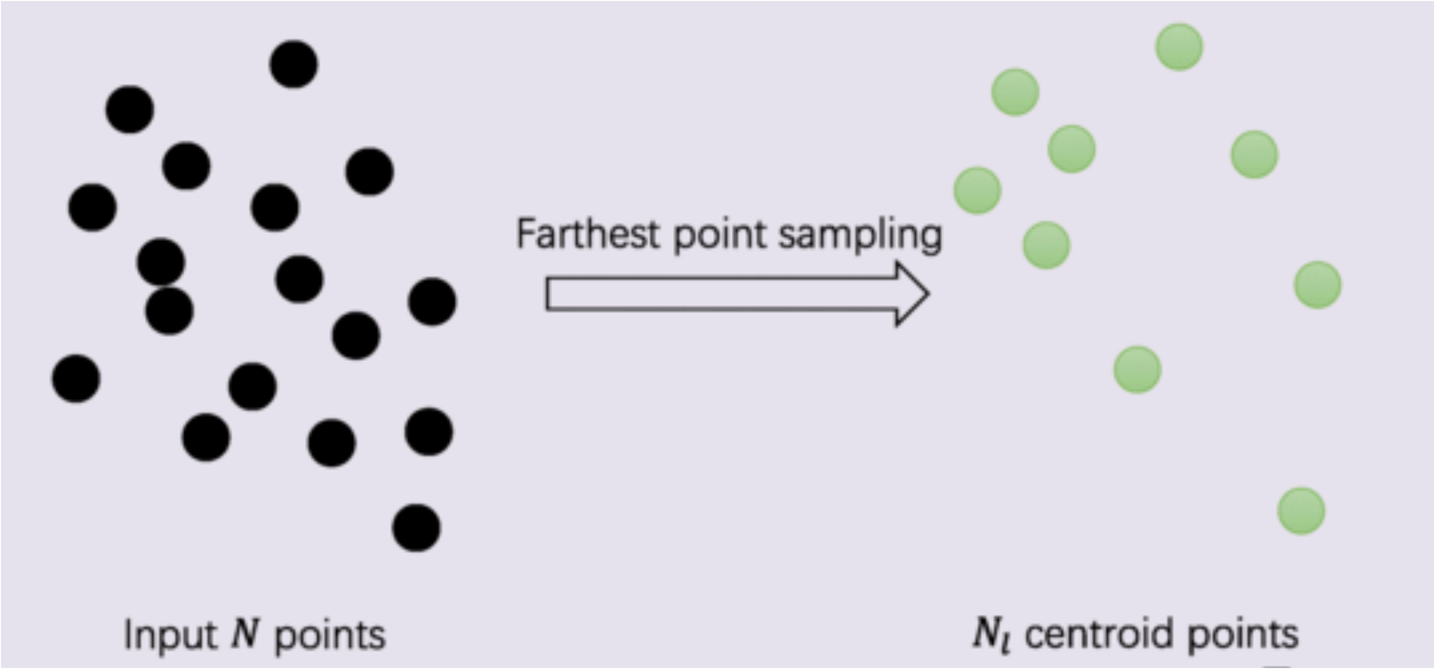
$$\begin{aligned}\bar{\alpha}_{i,j} &= \frac{\tilde{\alpha}_{i,j}}{\sqrt{d_a}},\\ \alpha_{i,j} &= \mathrm{softmax}(\bar{\alpha}_{i,j}) = \frac{\exp\left(\bar{\alpha}_{i,j}\right)}{\sum\limits_{k}\exp\left(\bar{\alpha}_{i,k}\right)},\end{aligned} \tag{4}$$

Local Neighbor Embedding layer

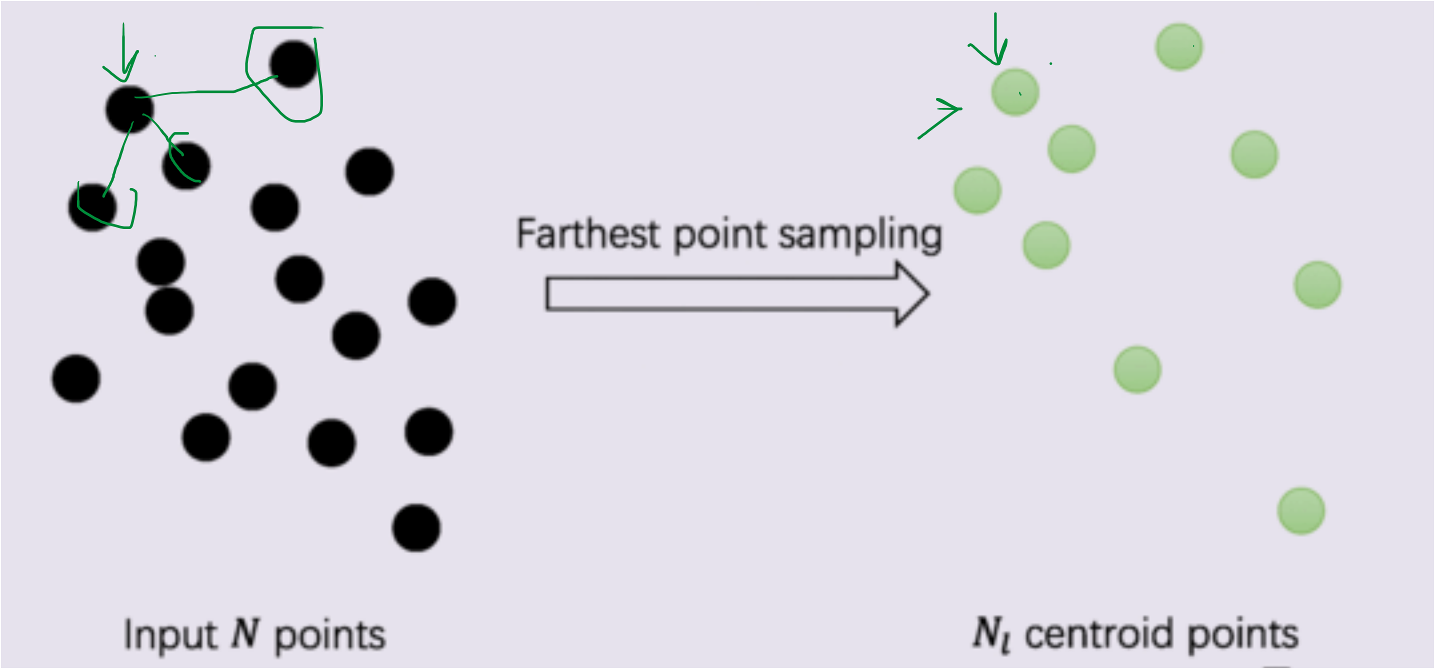
- Grouping and Sampling operations to augment PCT’s ability of local feature extraction

- Most of the code for this layer is already implemented, you have to implement small pieces of the code though

Sampling



Grouping



Notes about Grading

- (90 points in total) The maximum points for each sub-task are provided
- (10 points):
  - (5 points) Obtain a validation accuracy of more than or equal to 85% on at least two of the above models
  - (5 points) Obtain a minimum of 80% of validation accuracy on all the three models