

Supplementary Materials

Appendix A: Implementation Details

Similar to PointNetVLAD, the number of the input points is 4096. We set the number of clusters K in the NetVLAD layer to 64. The margins α and β for the lazy quadruplet loss are set to 0.5 and 0.2 respectively. The learning rate is set to 0.0005. We use a batch size of 1 tuple in each training iteration. Each training tuple consists of 1 positive and 9 negative point clouds. Note that instead of learning the descriptors with PointNetVLAD and then learning the PCAN with the fixed descriptors, we train the PointNet, PCAN, NetVLAD layer jointly for convergence.

Appendix B: Parameters of PAN

Since we compare our PCAN with PAN, an attention network which dose not aggregate the contextual information, we show the parameters for each layer in PAN in Fig. 1. FC layers(l_1, \dots, l_d) represents that it includes n shared MLP layers with width l_i ($i = 1, \dots, d$). The training settings are the same as what we used in PCAN.

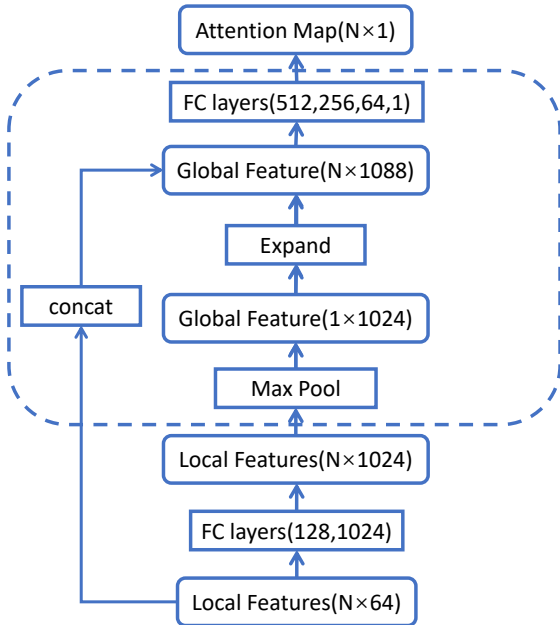


Figure 1. Details of PAN.

Appendix C: Robustness to rotation and jittering

We randomly rotate the points and jitter the position of each points by a Gaussian noise to evaluate the robustness with results shown in Table 1. It shows our network is robust both on rotation and jittering to some extent. Note

that we augment the training data by random rotation while without jittering. We think our attention module can help prevent the noises caused by jittering.

	Origin	Add Rotation	Add Jittering
Oxford	83.81	81.78	83.69
U.S.	79.05	78.86	77.37
R.A.	71.18	69.75	71.04
B.D.	66.82	65.30	66.64

Table 1. Top 1% average recall with rotation and jittering.

Appendix D: Can the proposed attention module be applied to a general network for 3D points?

Our proposed attention module is task-relevant and can be easily integrated to a general network with point-wise feature operation. For instance, in point cloud completion tasks which usually use the encoded global feature obtained from the partial model to produce a completed model, we consider our attention module will contribute to find the more important points in the partial model when encoding. Since this paper focus on point cloud based retrieval, we consider applying attention module on other 3D points applications is appealing and can be leaved for reaseachers to explore.