PV-Ada: Point-Voxel Adaptive Feature Abstraction for Robust Point Cloud Classification.

Lifa Zhu

2022/10/23

Team - DGPC

Lifa Zhu[1], Changwei Lin[1], Chen Zheng[2], and Ninghua Yang[1]

[1] DeepGlint, Beijing, China
[2] China University of Mining and Technology, Beijing, China

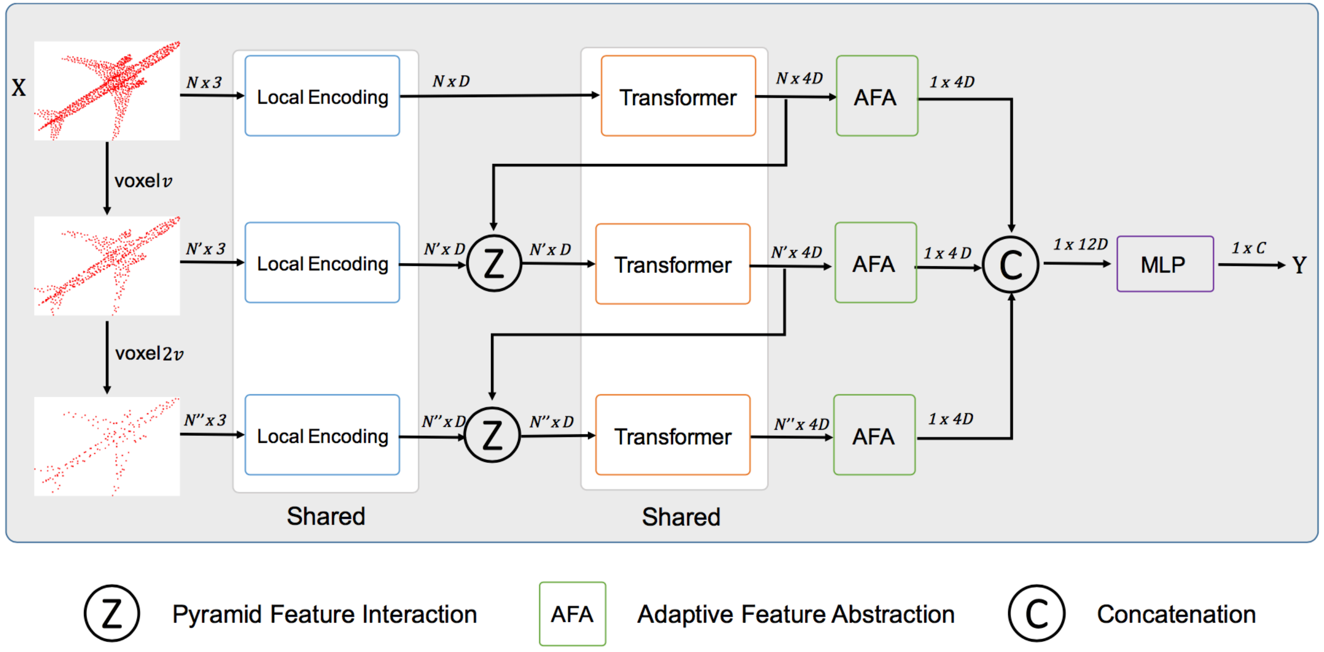
Contents

- Ideas
- Network architecture
- Experiments
- Discussion and Conclusion

Ideas

- The rough contour may benefit point cloud classification under corruptions, instead of paying attention to detailed structures.
- Not all all points in point cloud are equal for individual feature representation, such as plane points and outlier points.

Architecture



Inputted point cloud X, PV-Ada outputs C scores Y for all the C candidate categories.

Loss function: Cross entropy loss with label smoothing.

$$\begin{aligned} z_1(\mathbf{F}^1, \mathbf{F}^2) &= \mathbf{F}^2, \\ z_2(\mathbf{F}^1, \mathbf{F}^2) &= h_1(\mathbf{F}^1) + \mathbf{F}^2, \\ z_3(\mathbf{F}^1, \mathbf{F}^2) &= h_3(\mathrm{cat}([h_2(\mathbf{F}^1), \mathbf{F}^2])), \end{aligned} \tag{3}$$

$$\begin{aligned} \mathbf{S}^t_{\mathbf{X}} &= g(\mathbf{F}^t_{\mathbf{X}}), \\ \mathbf{F}_{\mathbf{X}} &= \mathrm{MaxPooling}(\mathbf{S}^t_{\mathbf{X}} \cdot \mathbf{F}^t_{\mathbf{X}}), \end{aligned} \tag{4}$$

- **Point-voxel encoder**
  - Local encoding groups k nearest neighbors for each point (voxel), then learns D dimensional point (voxel) feature $\mathbf{F}^e_{\mathbf{X}} \in \mathbb{R}^{N\times D}$.
  - Transformer is based on offset-attention(PCT) and residual structure. It generates enhanced point feature $\mathbf{F}^t_{\mathbf{X}} \in \mathbb{R}^{N\times 4D}$.
  - The input point cloud X is progressively voxelized twice with voxel size v and $2v$, generating $\mathbf{X}' \in \mathbb{R}^{N'\times 3}$ and $\mathbf{X}'' \in \mathbb{R}^{N''\times 3}$, respectively.

    Pyramid feature interaction is conducted as Equation (3).
- **Adaptive feature abstraction (AFA)** predicts point weight $\mathbf{S}^t_{\mathbf{X}}$ and conducts feature abstraction as Equation (4).

Experiments

- Training set: ModelNet40 training set.
- Validation set: the public ModelNet-C dataset.
- Testing set: the private ModelNet-C dataset.

Table 1. Full results for Overall Accuracy (OA) on testing set (ExtraOA), validation set (mOA) and clean ModelNet40 *testing set*. W.M. is short for WOLFMix and Tp is short for Tapering. Bold: best in column. Underline: second best in column.

	ExtraOA ↑	Clean ↑	mOA↑	Scale	Jitter	Drop-G	Drop-L	Add-G	Add-L	Rotate
PointNet[15]+W.M.	-	0.884	0.743	0.801	**0.850**	0.857	0.776	0.343	0.807	0.768
PCT[3]+W.M.	-	**0.934**	0.873	0.906	0.730	**0.906**	**0.898**	0.912	0.861	0.895
GDANet[26]+W.M.	0.797	**0.934**	0.871	**0.915**	0.721	0.868	<u>0.886</u>	0.910	0.886	**0.912**
RPC[17]+W.M.	0.739	<u>0.933</u>	0.865	0.905	0.694	0.895	<u>0.894</u>	0.902	0.868	<u>0.897</u>
PV-Ada+W.M.	<u>0.860</u>	0.923	**0.884**	<u>0.911</u>	<u>0.796</u>	<u>0.9</u>	0.88	**0.915**	**0.89**	0.896
PV-Ada+W.M.+Tp	**0.865**	0.911	<u>0.88</u>	0.907	0.792	0.897	0.874	<u>0.914</u>	0.884	0.892

Table 2. Full results for mCE on validation set. Bold: best in column. Underline: second best in column.

	mCE ↓	Scale	Jitter	Drop-G	Drop-L	Add-G	Add-L	Rotate
DGCNN[23]+W.M.	0.590	0.989	0.715	0.698	0.575	**0.285**	<u>0.415</u>	<u>0.451</u>
PCT[3]+W.M.	0.574	1.000	0.854	**0.379**	**0.493**	0.298	0.505	0.488
GDANet[26]+W.M.	<u>0.571</u>	**0.904**	0.883	0.532	0.551	0.305	<u>0.415</u>	**0.409**
RPC[17]+W.M.	0.601	1.011	0.968	0.423	<u>0.512</u>	0.332	0.480	0.479
PV-Ada+W.M.	**0.538**	<u>0.947</u>	**0.652**	<u>0.403</u>	0.58	<u>0.292</u>	**0.4**	0.493

Table 3. Ablation studies on the proposed modules and positive tricks on ModelNet-C.

Base	P-V	Best val	RSMix	AFA	Tapering	clean ↑	mOA ↑	ExtraOA ↑
✓						0.928	0.862	-
✓	✓					0.927	0.879	0.843
✓	✓	✓				0.923	0.881	0.848
✓	✓	✓	✓			0.921	0.882	0.855
✓	✓	✓	✓	✓		0.923	0.884	0.860
✓	✓	✓	✓	✓	✓	0.911	0.88	0.865

Table 4. Ablation studies on the negative designed modules.

Model	clean ↑	mOA ↑
Base + P-V (z_1)	0.927	0.879
Base + P-V (z_2)	0.932	0.862
Base + P-V (z_3)	0.926	0.868
Base + P-V (share)	0.927	0.879
Base + P-V (w.o. share)	0.929	0.855
Base + P-V (#2 voxel)	0.927	0.879
Base + P-V (#3 voxel)	0.914	0.868

Table 5. Model size and inference time

Model	# parameters	inference time
PV-Ada	3.16 M	15.7 ms

Discussion and Conclusion

- Equipped with a stronger backbone, better performance may be achieved in point cloud classification under corruptions.
- PV-Ada is proposed for robust point cloud classification under corruptions, which outperforms the state-of-the-art published models and achieves 0.884 and 0.865 OA on the public and private ModelNet-C test set.
- Point-voxel encoder and adaptive feature abstraction are two effective components for robust point cloud classification.
- Code is available at https://github.com/zhulf0804/PV-Ada.

Thank you.