ART-Point: Improving Rotation Robustness of Point Cloud Classifiers via Adversarial Rotation

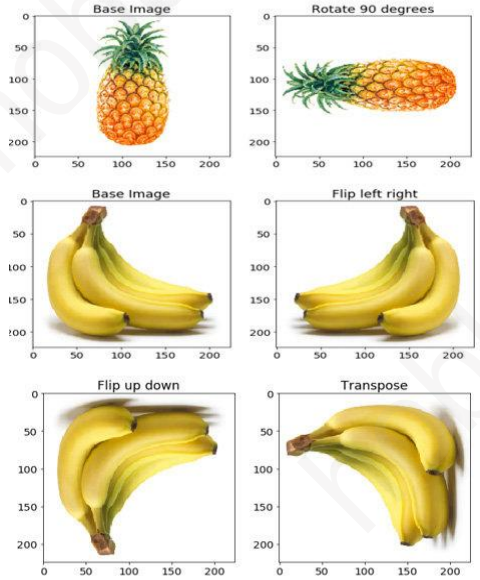
Huo Mingda

Jinan University, Guangzhou

May 11, 2023

研究内容：旋转增强数据

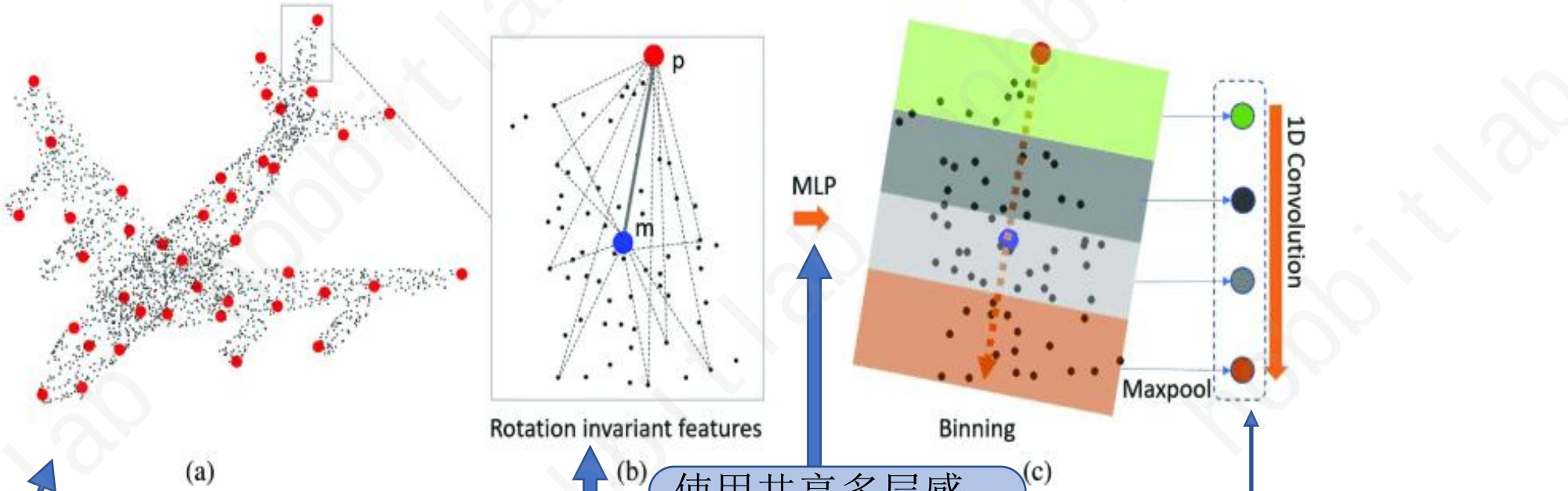
无法保证任意旋转鲁棒性



通过旋转增强获得：旋转来增加数据

天文数字，难以企及

研究内容：旋转不变性



对于具有/不具有关联特征的输入点云，通过最远点采样对代表点（红点）进行采样。

查询K个邻居产生参考点P的局部点集（P→M为参考方向）

使用共享多层感知器扩展到高维空间

最大值池化总结出每个箱子点集的特征

全局坐标系下点对关于主方向的角度和距离判定描述符

研究创新：对抗性训练优势

旋转增强数据

旋转不变性

依赖于特定的描述符和网络架构，限制了分类器在对齐数据集上的性能

如何在不改变输入空间和网络架构的前提下，减少点云旋转对分类结果的影响？

通过对抗训练获得

$$\min_{\theta} \rho(\theta), \quad \text{where} \quad \rho(\theta) = \mathbb{E}_{(p,q)\sim\mathcal{D}}[\max_{\delta\in\mathcal{S}} L(\theta, p+\delta, q)].$$

内部最大：被判定模型遭遇的最坏情况

外部最小：解决矛盾的最小调优方法

最终目的：寻找最坏情况下的最优解，鲁棒性角度，解决本质问题

研究创新：对抗性训练优势

旋转增强数据

旋转不变性

依赖于特定的描述符和网络架构，限制了分类器在对齐数据集上的性能

如何在不改变输入空间和网络架构的前提下，减少点云旋转对分类结果的影响？

通过对抗训练获得

$$\min_{\theta} \rho(\theta), \quad \text{where} \quad \rho(\theta) = \mathbb{E}_{(p,q)\sim\mathcal{D}}[\max_{\delta\in\mathcal{S}} L(\theta, p+\delta, q)].$$

内部最大：被判定模型遭遇的最坏情况

外部最小：解决矛盾的最小调优方法

最终目的：寻找最坏情况下的最优解，鲁棒性角度，解决本质问题

Huo Mingda, Jinan University

内部最大：轴向旋转攻击
（Axis-Wise Rotation Attack）

$$\frac{\partial L}{\partial \phi_z} = \sum_{i=1}^{n} \left(\frac{\partial x_i}{\partial \phi_z} \frac{\partial L}{\partial x_i} + \frac{\partial y_i}{\partial \phi_z} \frac{\partial L}{\partial y_i} + \frac{\partial z_i}{\partial \phi_z} \frac{\partial L}{\partial z_i} \right)$$

$$= \sum_{i=1}^{n} \left(-y_i \frac{\partial L}{\partial x_i} + x_i \frac{\partial L}{\partial y_i} \right),$$

通过梯度下降迭代优化找到高损失对抗性旋转的角度

$$\xi^* = \text{argmax}_\xi \left| \frac{\partial L}{\partial \phi_\xi} \right|, \xi \in [x, y, z],$$

细分为三个轴后选择最激进的旋转轴

$$\phi_{\xi^*}^{(t+1)} = \text{Proj}_{[-\pi,\pi]} \left(\phi_{\xi^*}^{(t)} + \alpha \text{sign} \left(\frac{\partial L}{\partial \phi_{\xi^*}} \right) \right).$$

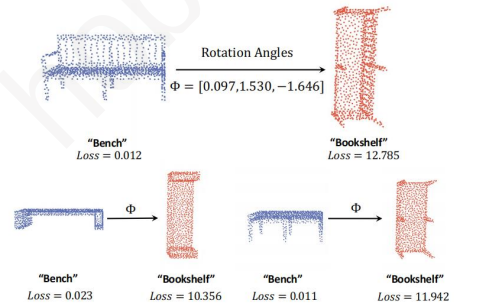
投影梯度下降（PGD）选择攻击角度

外部最小：一步优化
（one-step optimization）

旋转池

标签泄露引起的过拟合问题

分类保存对抗样本的旋转角度



对抗性旋转的可转移性

强化旋转池，
增效对抗样本，
提升模型鲁棒性素质

实验对比

模型梯度信息击穿旋转增强（RA）

用预训练的模型找到分类模型在旋转上的弱点，
然后生成新点云作为数据，去训练模型。

旋转不变（ID）/等变（EA）依赖
输入空间姿态信息分离

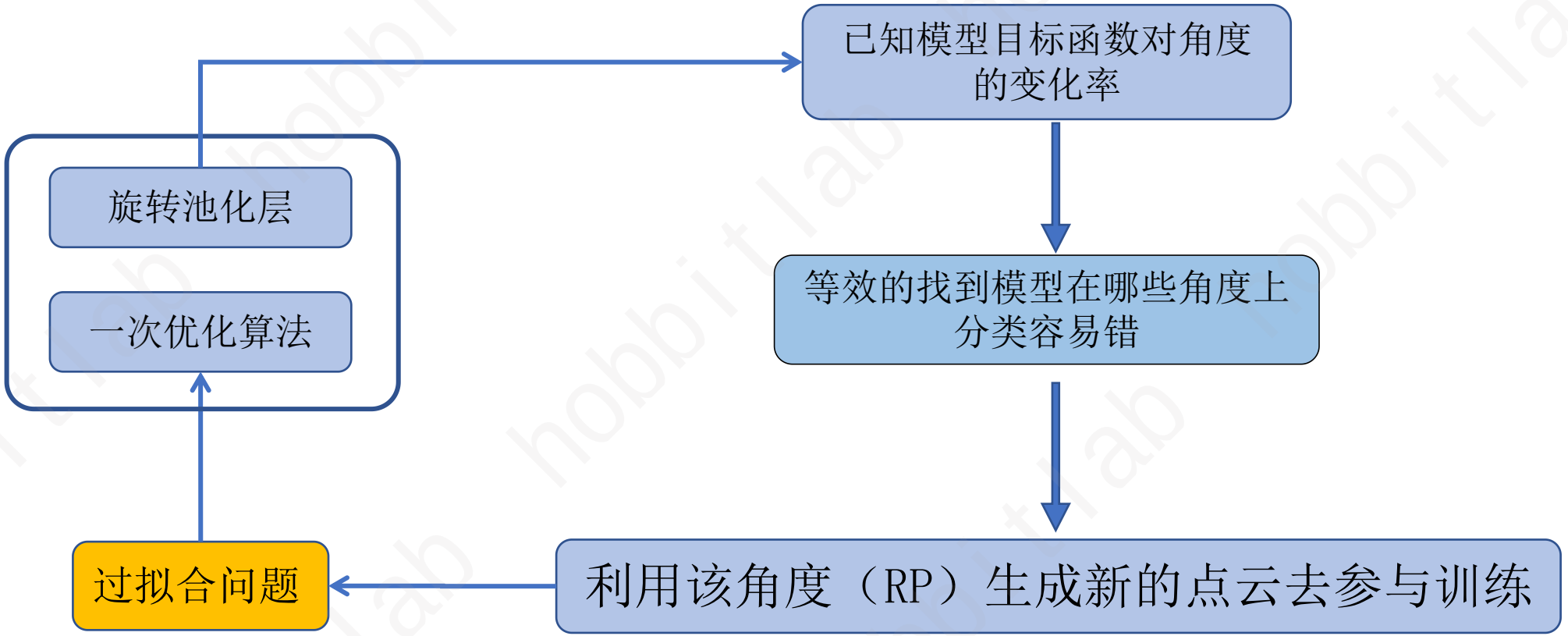
Method	ModelNet40		
	Attack	Random	Clean
PointNet [29] (RA)	55.6	74.4	76.7
PointNet++ [30] (RA)	58.9	80.1	82.3
DGCNN [39] (RA)	65.6	85.7	87.6
ART-PointNet (**Ours**)	85.6(30.0↑)	84.3(9.9↑)	85.5(8.8↑)
ART-PointNet++ (**Ours**)	90.1(31.2↑)	87.5(7.4↑)	88.6(6.3↑)
ART-DGCNN (**Ours**)	**91.5**(25.9↑)	**90.5**(4.8↑)	**91.3**(3.7↑)

Method	ShapeNet16		
	Attack	Random	Clean
PointNet [29] (RA)	66.4	87.3	89.5
PointNet++ [30] (RA)	70.5	89.7	92.1
DGCNN [39] (RA)	74.4	90.5	94.3
ART-PointNet (**Ours**)	96.9(30.5↑)	95.1(7.8↑)	96.2(6.7↑)
ART-PointNet++ (**Ours**)	97.8(27.3↑)	96.3(6.6↑)	97.5(5.4↑)
ART-DGCNN (**Ours**)	**98.4**(24.0↑)	**97.7**(7.2↑)	**98.1**(3.8↑)

Method	ModelNet40		
	Attack	Random	Clean
Classifiers Using Invariant Descriptors			
SFCNN [31]	90.1	90.1	90.1
RI-Conv [48]	86.5	86.4	86.5
ClusterNet [4]	87.1	87.1	87.1
RI-Framework [17]	89.4	89.3	89.4
Classifiers with Equivariant Architectures			
TFN [37]	87.6	87.6	87.6
REQNN [34]	74.4	74.1	74.4
VN-PointNet [7]	77.2	77.2	77.2
VN-DGCNN [7]	90.2	90.2	90.2
EPN [5]	88.3	88.3	88.3
Ours			
ART-PointNet	85.6	84.3	85.5
ART-PointNet++	90.1	87.5	88.6
ART-DGCNN	**91.5**	**90.5**	**91.3**

鲁棒依赖：
空间姿态信息
与点云分离

总结


已知模型目标函数对角度的变化率
等效的找到模型在哪些角度上分类容易错
利用该角度（RP）生成新的点云去参与训练
过拟合问题
旋转池化层
一次优化算法

Thanks!