東京工業大学
Tokyo Institute of Technology

PointGuard:

Provably Robust 3D Point Cloud Classification

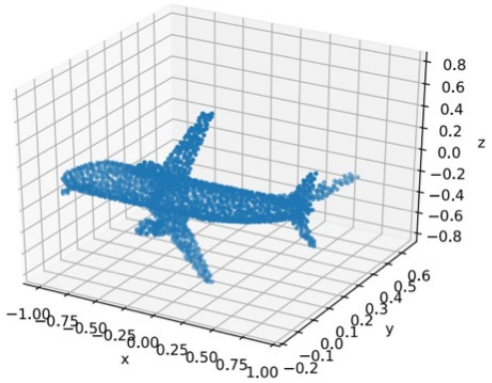
Hongbin Liu, Jinyuan Jia, Neil Zhenqiang Gong

Tian Xiao (23R51042)

26 Jan 2024

3D Point Cloud Classification

Point Cloud



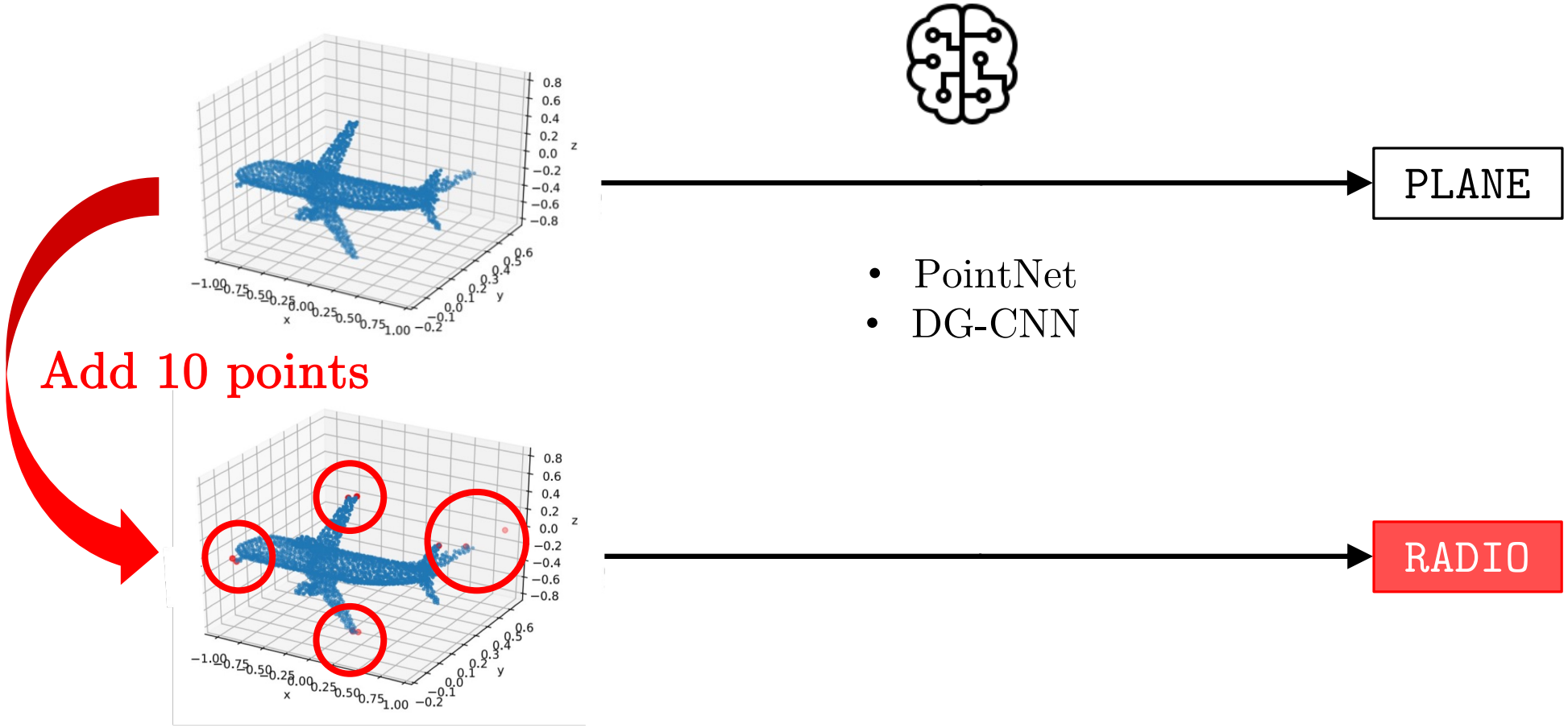
Model

Output

PLANE

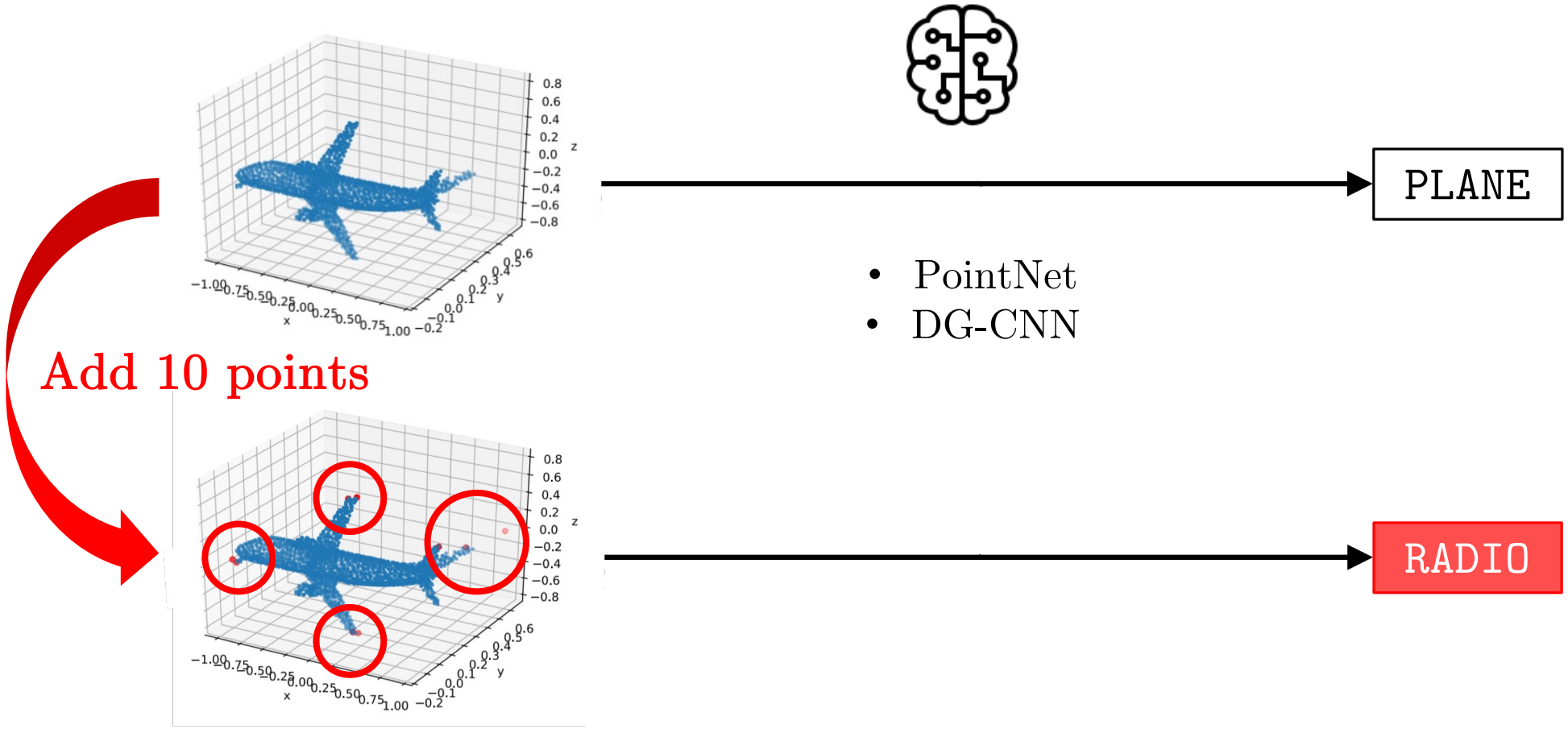
- PointNet
- DG-CNN

3D Point Cloud Classification



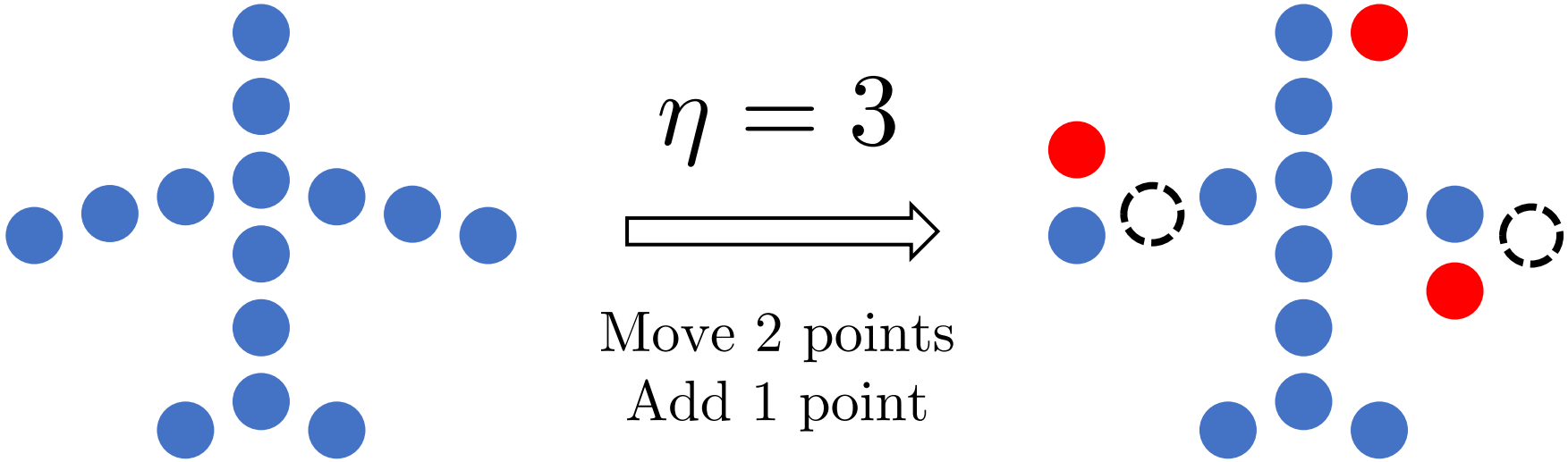
Adversarial Attack

PointNet/DG-CNN is **NOT** robust to adversarial attack.



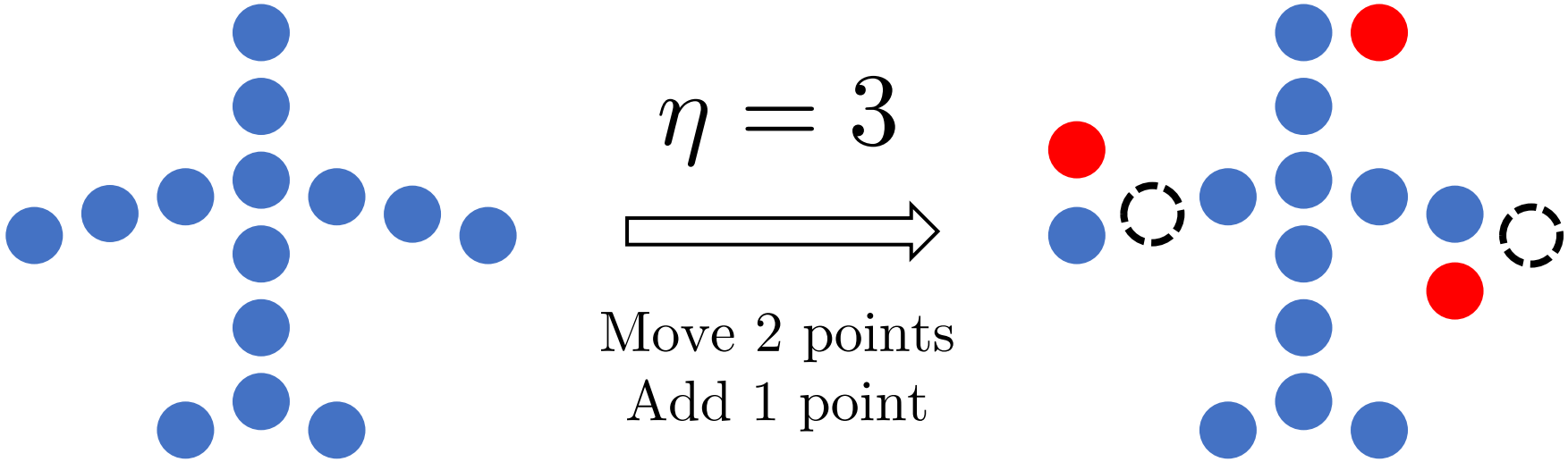
Perturbation Size of an Attack, η

- Number of points perturbed (i.e., moved, added, deleted).



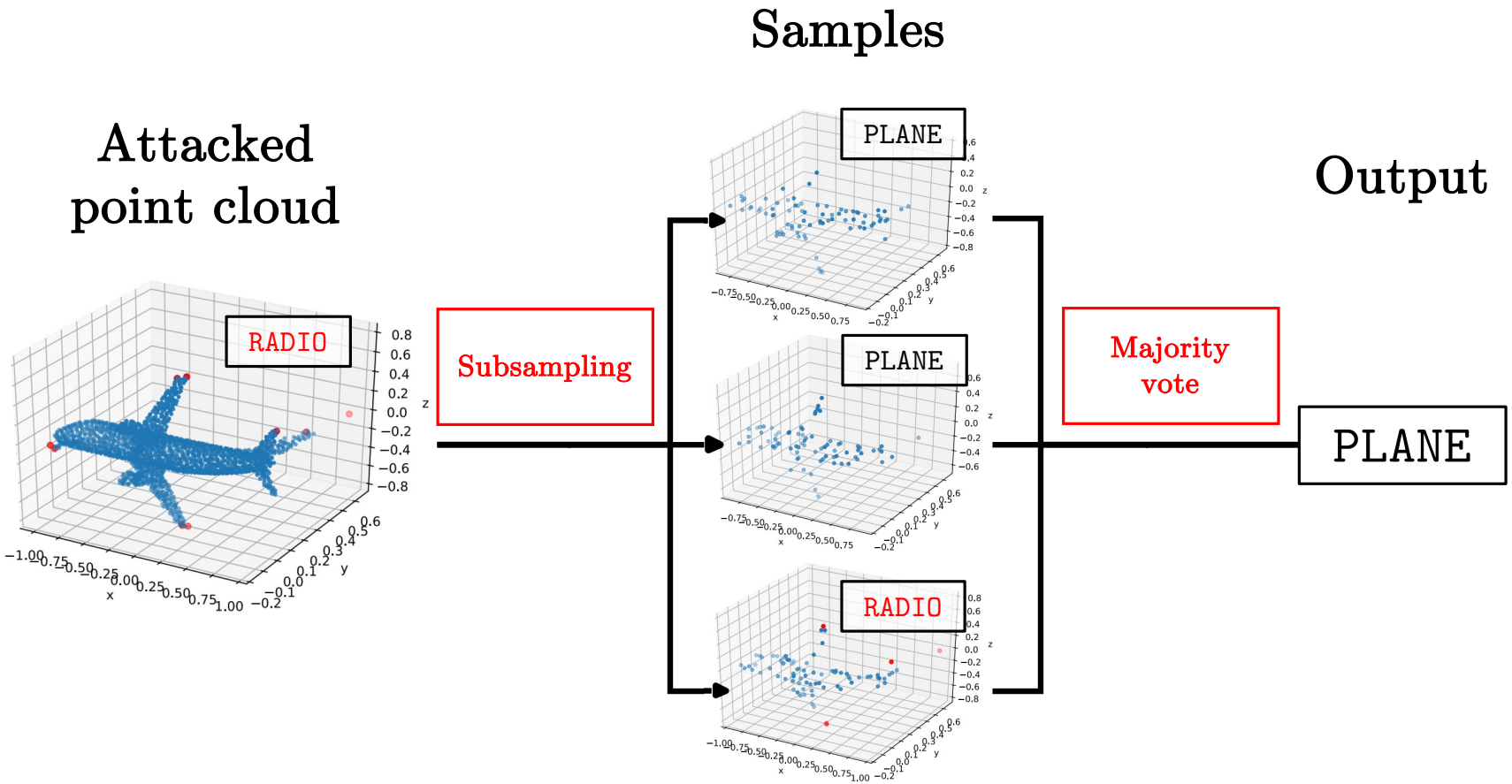
Perturbation Size of an Attack, η

- Number of points perturbed (i.e., moved, added, deleted).



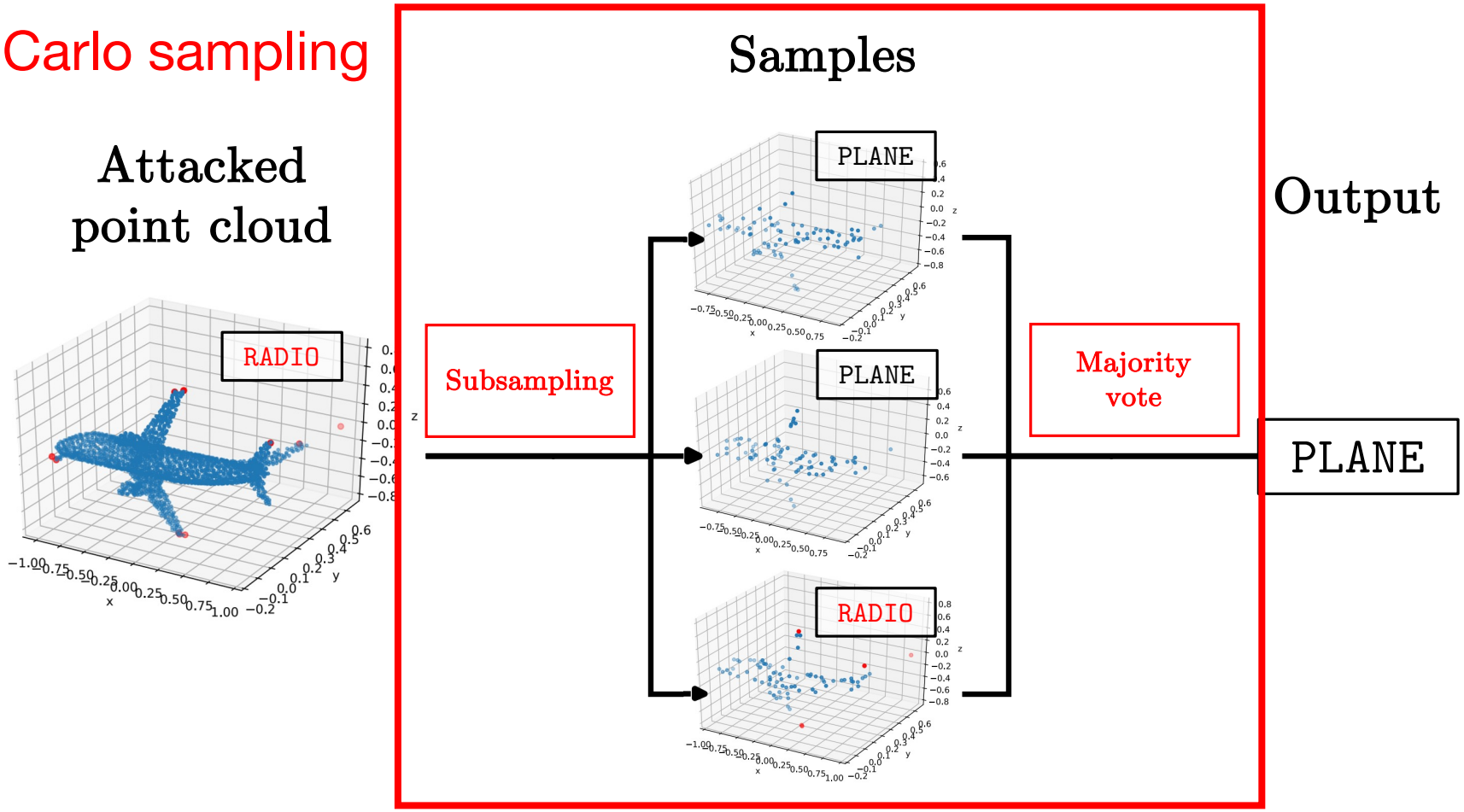
- Model robustness is assessed through **certified perturbation size** (max η s.t. the model always outputs correctly).

PointGuard: Overview

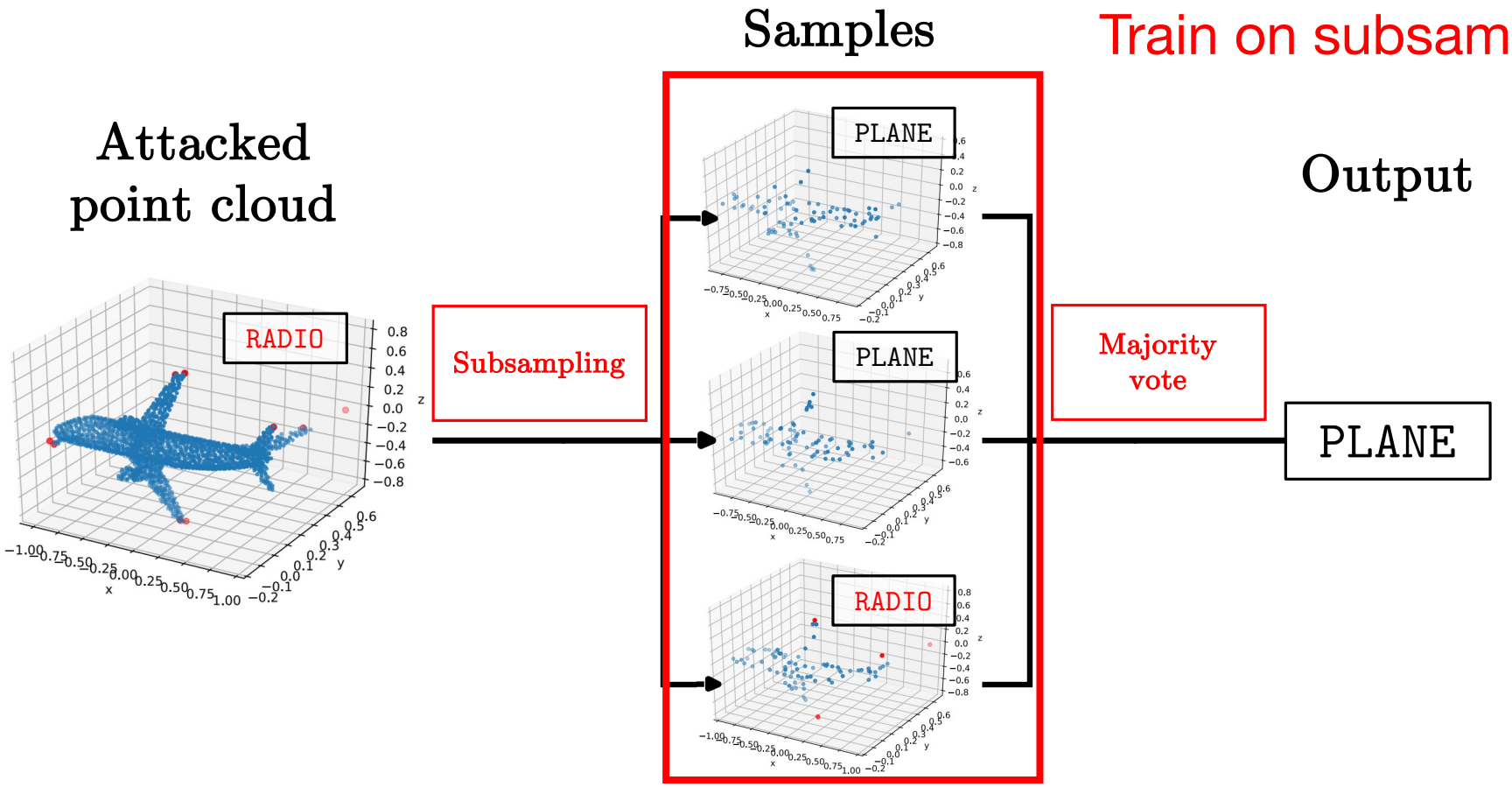


PointGuard: Implementation

Monte-Carlo sampling



PointGuard: Implementation



References

[1] Liu, H., Jia, J., & Gong, N. Z. (2021). PointGuard: Provably Robust 3D Point Cloud Classification. In *Proc. CVPR* (pp. 6186-6195).
[2] Cohen, J., Rosenfeld, E., & Kolter, Z. (2019). Certified Adversarial Robustness via Randomized Smoothing. In *Proc. ICML* (pp. 1310-1320).
[3] Zhou, H., Chen, K., Zhang, W., Fang, H., Zhou, W., & Yu, N. (2019). DUP-Net: Denoiser and Upsampler Network for 3D Adversarial Point Clouds Defense. In *Proc. ICCV* (pp. 1961-1970).

Certified Robust 2D Model

Backup

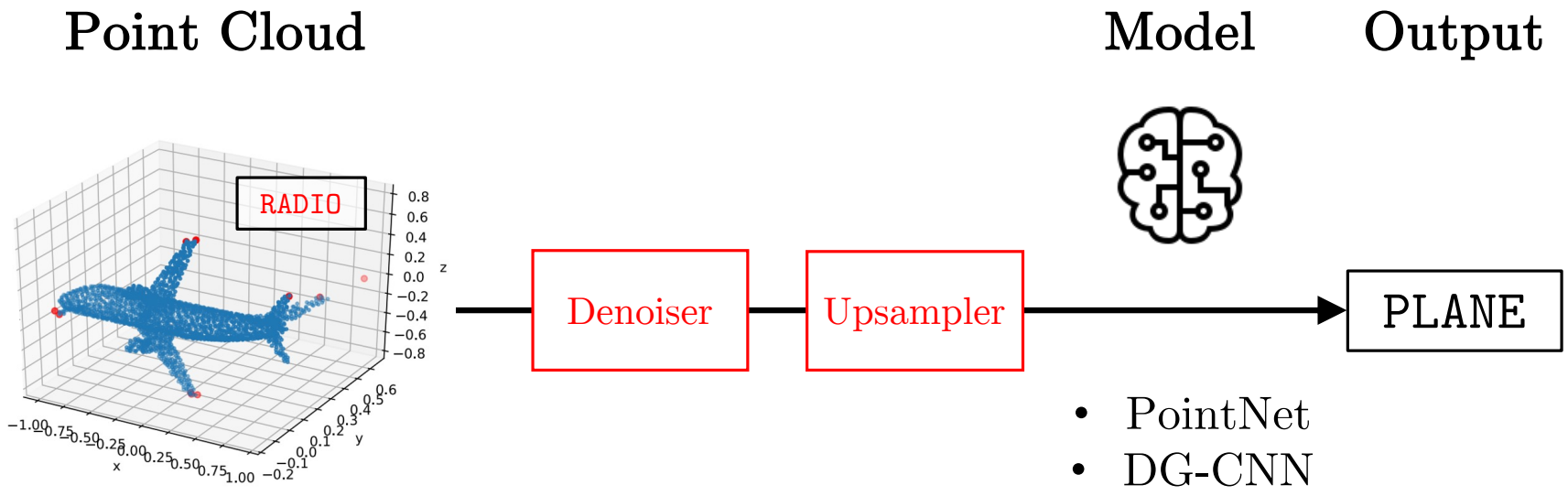
- Randomized smoothing [2]

Add random noises

Empirically Robust 3D Model

Backup

- DUP-Net [3]



Certified Perturbation Size

Backup

$$r^* = \underset{r}{\operatorname{argmax}}\, r$$

$$s.t. \max_{n-r \le t \le n+r} \frac{\binom{t}{k}}{\binom{n}{k}} - 2 \cdot \frac{\binom{\max(n,t)-r}{k}}{\binom{n}{k}} + 1 - \underline{p'_y} + \overline{p}'_e < 0,$$