Adversarial Geometric Transformations of Point Clouds for Physical Attack

Jingyu Xiang[1], Xuanxiang Lin[1], Ke Chen[2], and Kui Jia[3]

[1] South China University of Technology, Guangzhou, China
[2] Peng Cheng Laboratory, Shenzhen, China
chenk02@pcl.ac.cn
[3] The Chinese University of Hong Kong, Shenzhen, China
kuijia@cuhk.edu.cn

Abstract. Towards adversarial physical attack in real world, we argue that the main challenge lies in discounting adversarial effects by changes of point density along object surface. Most of existing point-wise perturbation based attackers concern on suppressing geometric irregularities, but it remains challenging to produce adversarial shape with geometric smoothness. Adversarial attack via the isometry transformation can alleviate irregular geometries but suffer from its rotation-sensitive nature, so its impractical assumption of category-level pre-alignment on benign object point clouds cannot be relaxed. In light of this, we explore non-rigid geometric transformations for geometry-aware adversaries with a flexible density-aware transformation on the whole point sets, which can thus impose constraints of global and local surface properties when adversarially deforming points. Experiment results on publicly benchmarking ModelNet40 and ScanObjectNN datasets verify the effectiveness of our transformation-based generation algorithms for adversarial shape and physical attack against both rotation sensitive and agnostic point classifiers, significantly outperforming existing adversarial point attackers under diverse recent defenses and the state-of-the-art physical attack methods.

Keywords: Deep Learning · Adversarial Attack · Geometry-aware Transformation · Physical Attack.

1 Introduction

Deep learning based algorithms have been widely adopted for semantic analysis on object shape such as autonomous driving [2, 12, 14] and augmented reality [1, 18, 21], which are verified their vulnerability against adversarial examples on point-based shape representations [3, 40, 53] to leave security issues of neural perception systems. Adversarial attack on point clouds aims to 1) fool the point classifiers of interest; and 2) achieve visual imperceptibility of adversaries to humans.

Adversarial effects in terms of mis-classification can be objectively measured by a series of well-defined performance metrics (*e.g.* classification accuracy),

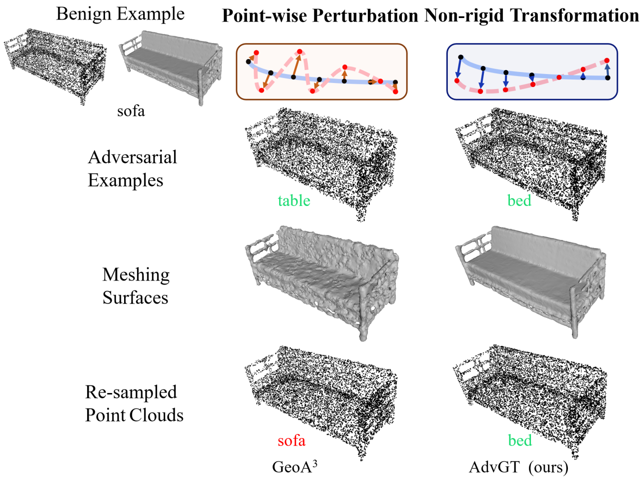

Fig. 1. Comparison with existing point-wise perturbation based physical attack, *i.e.* the GeoA^3 [49], and our non-rigid transformation based AdvGT against the victim PointNet. Class predictions (in color) of adversaries and re-sampled point clouds from reconstructed meshes are provided, which verify the superiority of our AdvGT on generating adversarial shape.

while visual imperceptibility, *i.e.* whether humans can distinguish adversarial examples from benign ones, can only be measured by an approximation surrogate (*e.g.* the PSNR [34]). Despite its insensitive nature to texture changes, humans' visual perception system can capture subtle irregularities on the geometric shape. Therefore, unlike the case of 2D image adversaries where the less perceptible high-frequency and low-magnitude noises are added to pixel grids' intensities, adversarial examples on point sets in 3D domain desire for geometric fairness and smoothness on the shape representations for visual imperceptibility.

Beyond early exploration via attaching and dropping a set of points to benign point clouds, existing adversarial point clouds are mainly obtained via optimization-based point perturbation, resulting in geometric irregularities (*e.g.* point outliers and bumpy surface) that are prone to recent defense strategies utilizing geometric properties of object surface [52, 63] and the more important physical attack setting.

Although a number of recent geometry-aware adversarial attackers [42, 49] are proposed to overwhelm those defenses and survive under physical attack, those point-based adversarial examples still suffer from irrational shape changes and the consequent geometric perceptibility. Fig. 1 shows an example from the ModelNet40 dataset, which is under attack by existing point-wise perturbation based physical attack, *i.e.* the GeoA^3 [49], and our non-rigid transformation based attack. The adversaries by point-wise perturbation based attack can be easily spotted due to its more uneven and irregular shape (especially the meshing surfaces), although it can cheat the victim classifier successfully.

More importantly, perturbation-based adversarial point clouds cannot ensure that all adversarial effects are from the more vital shape changes rather than trivial changes of point density, which therefore leads to the decrease of the attack success rate in adversarial physical attacks. This is demonstrated by the fact that the victim classifier can correctly classify the re-sampled point clouds from reconstructed meshes under the perturbation-based adversarial attack in Fig. 1. Recently, the isometry transformation [60] is proposed to gain adversarial impact mainly dependent on unnoticeable changes of object rotation, which introduces an alternative to generate adversaries by transformation operations on the whole point sets. The advantage of such transformation-based attack method lies in approximately preserving geometric properties of object surface, but its rotation-sensitive nature makes it less feasible to practical scenarios. In view of this, we introduce a simple yet effective attack method − adversarial geometric transformation (AdvGT), via optimization of a flexible density-aware transformations on all points' coordinates for adversarial shapes.

Intuitively, the effectiveness of our density-aware transformation can be attributed to the non-linear transformation applied to the object surface. Additionally, this transformation can induce modifications in the distribution of points, but due to its density-aware projection, it can maintain the geometric properties of continuous 2D manifolds that are embedded in 3D space (i.e., the object surface). As a result, this transformation can reduce the visual perceptibility of irregular geometries. Experiment results on public benchmarks verify the effectiveness of our proposed transformation on generation of adversarial point clouds, significantly outperforming existing attack methods under diverse state-of-the-art defense algorithms. More importantly, our method becomes the new state-of-the-art for the more challenging adversarial physical attack, which can demonstrate the superiority of our AdvGT method on generating adversarial shape, as shown in Fig. 1.

The main contributions of our paper are summarized as the following.

- A novel adversarial attack method is proposed to maintain surface properties yet impose adversarial effects via a non-rigid transformation on all points, which can favor for regular shape deformations to survive in the challenging physical attack and also the ordinary point-based attack against diverse state-of-the-art defenses.
- Technically, this paper proposes adversarial geometric transformation in a density-aware style, which concerns on deformation of global and local geometries rather than point-wise perturbation typically adopted in existing 3D adversarial generators.
- Extensive experimental results can demonstrate our motivation and adversarial effects of obtained point clouds by our AdvGT and re-scanned point clouds from adversarial shape, achieving remarkably superior physical attack success rate (at least 25.60% performance gain) to the state-of-the-art methods.

Source codes will be released after acceptance[1].

[1] https://github.com/starry1010/AdvGT

2 Related Works

Rotation-Sensitive Point Classification Existing deep classifiers on point clouds can be divided into three groups: multi-view image based [11,27,39,47,56]; voxel-based [22, 29, 38, 44, 51]; geometric deep learning based [5, 35, 36, 45, 54, 55, 65]. Both multi-view based and voxel-based algorithms rely on the Euclidean convolution operations on the regular grids-based approximation of object shape projected or converted from point clouds, which can suffer from information loss for transformation between non-Euclidean and Euclidean data space. Geometric deep learning-based methods such as PointNet [35] and DGCNN [45] concerns on the data-specific challenges of irregular structured and orderless point sets via multi-layer perceptrons (MLPs). Yet they require category-level pre-aligned input and thus malperform on point clouds of arbitrarily poses that are widely encountered in real world applications.

Rotation-Agnostic Point Classification Recently, point cloud classification with rotation robustness has received a lot of attention, which can be divided into three main groups – weakly supervision on spatial transformation [35, 58], learning with rotation invariant features [24, 37, 59], and achieving invariance via learning rotation equivariant features [6, 9, 26]. The first group of algorithms cannot guarantee canonical pose transformation as they lack explicit pose supervision signals, while methods falling into the second group depend on rotation invariant quantities that only partially capture the geometric information as feature input, which would not be optimal for semantic classification. In this work, we focus on the rotation-agnostic classifiers via learning rotation-equivariant features, which guarantees that rigid transformations of objects in the Euclidean space can lead to an equivalent transformation of features in feature space. The works [6,9,26] first convert the point clouds to spheric signals and utilize convolution on the rotation group, *i.e.* spherical convolutions, to achieve rotation equivariance, while Weiler *et al.* [48] introduce 3D steerable convolution for rotation-equivariant features by a set of vector-form and scalar-form fields. Thomas *et al.* [41] propose a tensor-field representation to achieve $SE(3)$ equivariance on irregular point clouds. Chen *et al.* [4] decouple the rotation-equivariant convolutions in the $SE(3)$ space into two separable convolution operations in the 3D Euclidean and $SO(3)$ space, which is further combined into rotation invariant features via attention weights. Deng *et al.* [7] introduce a generic concept of Vector Neurons (VN) to extend element-wise neurons with vector-formed directions, which can be easily adopted in existing point classifiers, to address the challenge of non-differentiable characteristics of existing rotation equivariant operation groups on irregular point clouds. In view of its generality and superiority in point-based rotation equivariant feature encoding, our paper adopts the Vector Neurons with two representative backbones – PointNet [35] and DGCNN [45] as the rotation-agnostic victim classifiers to be attacked.

Generation of Adversarial Point Clouds Deep neural networks are verified their vulnerability to *Adversarial sample*, which was first pointed out by Szegedy *et al.* [40]. Conceptually, by adding slight but intentionally perturbations to inputs that were originally correctly classified, an adversarial sample can mislead classifiers into making incorrect decisions. The generation of adversarial samples is refer to adversarial attack. There are two common modes of adversarial attack: white-box attack and black-box attack. White-box attack and black-box attack differ in the attacker's knowledge of the model's algorithm and parameters: the former has full access, while the latter has none. In this paper, we use the setting of white-box attack.

Adversarial attack has been widely investigated in the 2D image domain [32,33]. The goal of adversarial attack on point sets attempts to alter the benign point clouds with human-unnoticeable changes yet with adversarial effects to fool the classifier of interests, which attract a recent surge of attention in the field of 3D semantic analysis. Existing adversarial generation algorithms can be divided into point attaching/detaching based [50,53,57,61] and point perturbation based [42,49,53]. Point attaching algorithms such as the methods in [53,57] gain adversarial point clouds via attaching a set of independent points or point clusters to the "vulnerable" regions of benign point clouds, which are hard to be ignored by humans. Dropping essential points for object classification [50,61] can be more imperceptible, but cannot avoid dropping points with high-frequency (*i.e.* with higher curvatures) leading to less smooth geometries. Previous point perturbation based attackers [42, 49, 53] optimize point-wise coordinate offsets with regard to promoting mis-classification, additionally using geometry-aware regularization on adversarial generation such as the curvature-consistency objective in [49] and explicit constraints to enforce deformation along the surface [17]. Beyond the above algorithms, Hamdi *et al.* [15] propose a transferable adversarial perturbation attacker that can capture data distribution. Zhou *et al.* [62] incorporate label encoding of target predictions into feature encoding of benign examples to generate adversarial point clouds, in a generative adversarial network (GAN) structure. Hu *et al.* [16] vary certain geometric structures in the graph spectral domain for adversarial effects, while Liu *et al.* [25] craft adversarial samples from the low-frequency component of point clouds. The work [60] demonstrates the vulnerability of main-stream 3d models under global isometric transformation, which share similar scripts of transformation based adversarial attack as our AdvGT. However, their method mainly concerns rigidly rotating objects to fool the victim classifiers, which can be less effective on rotation agnostic classifiers, while ours non-rigid transformation method on points' coordinates can encourage adversarial shape against both rotation sensitive and agnostic classifiers.

3 Methodology

3.1 Preliminaries

Given $\mathcal{X}$ and $\mathcal{Y}$ denoting the input and output space respectively, training samples for supervised semantic classification on point clouds consist of $(\mathcal{P}, y)$, where $\mathcal{P} = \{\boldsymbol{p}_i\}_{i=1}^{\mathrm{n}} \in \mathcal{X}$ denotes a point cloud and $y \in \mathcal{Y}$ represents its corresponding semantic class label. Each point $\boldsymbol{p}_i \in \mathbb{R}^3$ is depicted by its 3D coordinates and the size of $\mathcal{P}$ is n. Such a problem aims to learn a mapping function $\Phi_\theta : \mathcal{X} \to \mathcal{Y}$ that classifies any point cloud $\mathcal{P}$ into one of the M object categories in $\mathcal{Y}$ (*i.e.* $|\mathcal{Y}| = \mathrm{M}$), where θ is a set of model parameters to be optimized.

In the context of geometric deep learning on point clouds, the mapping function $\Phi_\theta(\mathcal{P})$ can be made up of a cascade of a feature encoding module $\Phi_{\mathrm{fea}} : \mathcal{X} \to \mathcal{F}$ and a classification module $\Phi_{\mathrm{cls}} : \mathcal{F} \to \mathcal{Y}$ as follows:

$$\Phi_\theta(\mathcal{P}) = \Phi_{\mathrm{cls}} \circ \Phi_{\mathrm{fea}}(\mathcal{P}), \tag{1}$$

where $\mathcal{F}$ denotes the feature space and the feature encoder can consist of multi-layer perceptrons (MLPs) such as PointNet [35] and DGCNN [45] or convolution based on rotation equivariant rotation groups such as the SphericalCNNs [6] and 3D Steerable CNNs [48]. The classification module Φ_{cls} predicts the probabilities $\boldsymbol{p}$ of $\mathcal{P}$ belonging to object categories, where $\boldsymbol{p} = \Phi_\theta(\mathcal{P}) = [p_1, \cdots, p_{\mathrm{M}}]$ is subjected to $\sum_{i=1}^{\mathrm{M}} p_i = 1$. The deep model Φ_θ is trained by adjusting its parameters θ to minimize the cross-entropy loss $J(\Phi_\theta(\mathcal{P}), y)$, for each sample $\mathcal{P}$ and its one-hot label y that are sampled from data distribution $\mathcal{D}$:

$$\min_\theta \ \mathbb{E}_{(\mathcal{P},y)\sim\mathcal{D}} J(\Phi_\theta(\mathcal{P}), y). \tag{2}$$

For rotation-sensitive classifiers, the representative PointNet [35], and DGCNN [45] are selected. Since our adversarial attack is not limited to rotation equivariant point classifiers, we select the Vector Neurons [7] for its generality to existing point classifiers and differentiable characteristics on irregular points. Note that, the main difference between rotation-sensitive and rotation-equivariant classifiers lies in layer-wise feature encoding, and therefore both groups of popular point classification algorithms are adopted in our experiments.

To attack a trained classifier Φ_θ, an adversary $\tilde{\mathcal{P}} = \{\tilde{\boldsymbol{p}}_i\}_{i=1}^{\mathrm{n}} \in \tilde{\mathcal{X}}$ is generated from a benign point cloud $\mathcal{P}$ such that the obtained $\tilde{\mathcal{P}}$ would be mis-classified by Φ_θ, where the $\tilde{\mathcal{X}}$ denotes the adversarial space. The adversary $\tilde{\mathcal{P}}$ can be obtained via point-wise perturbation [42,49,53], attaching/detaching points [50,53,57,61] or global transformation [60]. Specifically, adversarial generation of point clouds can be carried out via optimizing the following object function:

$$\min \quad L_{\mathrm{mis}}(\tilde{\mathcal{P}}) + \lambda L_{\mathrm{imp}}(\tilde{\mathcal{P}}, \mathcal{P}) \tag{3}$$

$$\text{s.t.} \quad y \neq \arg\max_j \Phi_\theta(\tilde{\mathcal{P}})_j, \tag{4}$$

where λ is a trade-off parameter between the loss term $L_{\mathrm{mis}}(\tilde{\mathcal{P}})$ measuring adversarial attack success rate, and the distance metric $L_{\mathrm{imp}}(\tilde{\mathcal{P}}, \mathcal{P})$ penalizing geometric dissimilarity between the benign $\mathcal{P}$ and the resulting $\tilde{\mathcal{P}}$ in terms of visual imperceptibility to humans.

Under the setting of the white-box attack, *i.e.* the attacker has full access to the architecture and parameters θ of a neural classifier Φ_θ, adversarial point clouds $\tilde{\mathcal{P}}$ of an untargeted attack can be mis-classified from the class y of $\mathcal{P}$ to one of the other classes (*i.e.* $y \neq \arg\max \ \Phi_\theta(\mathcal{P})$). To promote attacking effects on the victim classifier,we use the C&W loss function introduced by [3], which has been adopted in recent works [42, 49, 53] as follows:

$$L_{\mathrm{mis}}(\tilde{\mathcal{P}}) = \max\{-\kappa, \max_{j \neq t} \Phi_\theta(\tilde{\mathcal{P}})_j - \Phi_\theta(\tilde{\mathcal{P}})_t\}, \tag{5}$$

where $\kappa \geqslant 0$ is a margin threshold, and t is the class entry with the highest confidence except for the true label in $\Phi_\theta(\tilde{\mathcal{P}})$ under the untargeted attack setting.

3.2 Adversarial Geometric Transformations

Existing adversarial generation on perturbing individual points $\{\boldsymbol{p}_i\}_{i=1}^{\mathrm{n}}$ of $\mathcal{P}$ can enforce geometry-aware constraints on local regions for point-based adversaries but still cannot avoid geometric irregularities due to failure of imposing unified changes on global shape, which can thus lead to unsatisfactory adversarial shape and physical attack. Inspired by [20, 62], this paper explores geometrically smooth shape transformations on the points of $\mathcal{P}$ when generating adversaries. Adversarial effects of such transformation can be contributed to all the points' deformation, rather than individual ones or point clusters as in previous works. Without loss of generality, geometric transformation operations on 3D shapes have been well investigated in mathematics and computer vision, including rigid ones such as the isometry transformation and non-rigid ones such as the affine and the projective transformations. Since rigid deformations do not change geometric patterns of the object surface, adversarial effects are only triggered by changes in the location and pose of point clouds. As a result, their resulting point adversarial examples [60] cannot survive in the wild due to their impractical rotation-sensitive nature, *e.g.* when attacking rotation-agnostic classifiers or for the more challenging physical attack on arbitrarily posed benign point clouds. In order to generate practical adversarial examples, we propose a straightforward yet effective algorithm that utilizes non-rigid transformations to perturb all points in the point cloud $\mathcal{P}$ in a globally correlated manner, with the aim of preserving geometric fairness and smoothness.

Density-Aware Transformations This paper explores a novel density-aware transformation, which combines a number of locally anchored geometric transformations to increase flexibility and diversity of adversarial point clouds. Technically, a set of anchor points $\mathcal{P}^{\mathcal{A}} = \{\boldsymbol{p}_j^{\mathcal{A}}\}_{j=1}^{\mathrm{m}} \in \mathcal{P}$ are sampled from the benign point cloud, on which geometric transformations are conducted to produce adversaries. To obtain the anchor points, the Farthest Point Sampling (FPS) algorithm is adopted for its simplicity. However, other sampling strategies such as geometry-aware sampling [30] could also be considered. To ensure spatial continuity and fairness of adversarial shape, a set of local anchor points in 3D space are converted into an anchor density map on 2D object surface manifolds. A typical option is to use the Gaussian kernel. Given the anchor points $\mathcal{P}^{\mathcal{A}}$, the Gaussian kernel is defined as follows:

$$\mathcal{K}_\sigma(\boldsymbol{p}_i, \boldsymbol{p}_j^{\mathcal{A}}) = \exp(\frac{-\|\boldsymbol{p}_i - \boldsymbol{p}_j^{\mathcal{A}}\|_2^2}{2\sigma^2}), \tag{6}$$

where $\mathcal{K}_\sigma(\cdot,\cdot)$ denotes the Gaussian kernel with a hyperparameter σ and $\|\cdot\|_2$ denotes the Euclidean norm. Therefore, any point $\boldsymbol{p}_i$ sampled from object surface can be measured by its density to the anchor $\boldsymbol{p}_j^{\mathcal{A}} \in \mathcal{P}^{\mathcal{A}}$.

We employ the Nadaraya-Watson kernel regression [46] to incorporate the Gaussian anchor density into the transformation, which results in the following density-aware transformation on local shape as:

$$\tilde{\boldsymbol{p}_i} = \mathcal{T}(\boldsymbol{p}_i) = \frac{\sum_{j=1}^{\mathrm{m}} \mathcal{K}_\sigma(\boldsymbol{p}_i, \boldsymbol{p}_j^{\mathcal{A}})\mathcal{T}_j}{\sum_{j=1}^{\mathrm{m}} \mathcal{K}_\sigma(\boldsymbol{p}_i, \boldsymbol{p}_j^{\mathcal{A}})}, i = 1, 2, \ldots, \mathrm{n}, \tag{7}$$

where $\mathcal{T}$ is the density-aware transformation for the benign point clouds, and $\mathcal{T}_j$ is the transformation centered at anchor point $\boldsymbol{p}_j^{\mathcal{A}}$. For the input point $\boldsymbol{p}_i, i = 1, 2, \ldots, \mathrm{n}$, $\mathcal{T}_j$ includes several transformations (*i.e.* scaling, rotation, and translation) that can be written as:

$$\mathcal{T}_j(\boldsymbol{p}_i) = \boldsymbol{S}_j \boldsymbol{R}_j(\boldsymbol{p}_i - \boldsymbol{p}_j^{\mathcal{A}}) + \boldsymbol{T}_j + \boldsymbol{p}_j^{\mathcal{A}}, \tag{8}$$

where $\boldsymbol{S}_j \in \mathbb{R}^{3\times3}$, $\boldsymbol{R}_j \in \mathbb{R}^{3\times3}$, and $\boldsymbol{T}_j \in \mathbb{R}^3$ respectfully denotes a scaling matrix, a rotation matrix, and a translation vector of $\mathcal{T}_j$ to be optimized to generate $\tilde{\mathcal{P}}$ that would be misclassified.

To avoid the overall surface destruction of adversarial examples, the scaling matrix $\boldsymbol{S}$ should be as close to the unit matrix $\boldsymbol{I}$ as possible; the rotation angle A of rotation matrix $\boldsymbol{R}$ and all entries of the translation vector $\boldsymbol{T}$ are also supposed to approach 0, which results in the following object function of implicit regularization of geometric imperceptibility as follows:

$$L_{\mathrm{def}}(\mathcal{T}) = \frac{1}{\mathrm{m}} \sum_{j=1}^{\mathrm{m}} (\|\boldsymbol{S}_j - \boldsymbol{I}\|_F^2 + \|\boldsymbol{A}_j\|_2^2 + \|\boldsymbol{T}_j\|_2^2), \tag{9}$$

where $\|\cdot\|_F$ denotes the Frobenius norm, $\boldsymbol{A}_j \in \mathbb{R}^3$ denotes the rotation angles at anchor points $\boldsymbol{p}_j^{\mathcal{A}}$ along three axes.

Loss Functions To encourage geometric similarity between $\mathcal{P}$ and the resulting $\tilde{\mathcal{P}}$, we consider two objective terms of explicit constraints for geometric imperceptibility – the Chamfer distance [10] and the consistency of local curvature [49] on similarities of global and local geometries respectively between benign and adversarial examples. The Chamfer distance takes the average of the distances of all nearest point pairs. Specifically, given two point sets $\mathcal{P}$ and $\tilde{\mathcal{P}}$ both having n points, the chamfer distance can be obtained as follows:

$$\begin{aligned} L_{\mathrm{Cha}}(\mathcal{P}, \tilde{\mathcal{P}}) = \frac{1}{n} \sum_{\tilde{\boldsymbol{p}} \in \tilde{\mathcal{P}}} \min_{\boldsymbol{p} \in \mathcal{P}} \|\tilde{\boldsymbol{p}} - \boldsymbol{p}\|_2^2 \\ + \frac{1}{n} \sum_{\boldsymbol{p} \in \mathcal{P}} \min_{\tilde{\boldsymbol{p}} \in \tilde{\mathcal{P}}} \|\boldsymbol{p} - \tilde{\boldsymbol{p}}\|_2^2, \end{aligned} \tag{10}$$

The loss term L_{Cur} that encourages the consistency of local geometries can be depicted as [49]:

$$\begin{aligned} L_{\text{Cur}}(\mathcal{P}, \tilde{\mathcal{P}}) = & \frac{1}{n} \sum_{\tilde{\boldsymbol{p}} \in \tilde{\mathcal{P}}} \|\zeta(\tilde{\boldsymbol{p}}, \tilde{\boldsymbol{n}}_{\boldsymbol{p}}; \tilde{\mathcal{P}}) - \zeta(\boldsymbol{p}, \boldsymbol{n}_{\boldsymbol{p}}; \mathcal{P})\|_2^2 \\ \text{s.t.} \quad & \boldsymbol{p} = \arg\min_{\boldsymbol{p} \in \mathcal{P}} \|\tilde{\boldsymbol{p}} - \boldsymbol{p}\|_2, \end{aligned} \tag{11}$$

where the function $\zeta(\boldsymbol{p}, \boldsymbol{n}_{\boldsymbol{p}}; \mathcal{P})$ is to describe local geometries of neighbourhood $\mathcal{N}_{\boldsymbol{p}}$ anchored on the point $\boldsymbol{p}$ with the following definition

$$\zeta(\boldsymbol{p}, \boldsymbol{n}_{\boldsymbol{p}}; \mathcal{P}) = \frac{1}{k} \sum_{\boldsymbol{q} \in \mathcal{N}_{\boldsymbol{p}}} | < (\boldsymbol{q} - \boldsymbol{p}) / \|\boldsymbol{q} - \boldsymbol{p}\|_2, \boldsymbol{n}_p > |,$$

where k denotes the number of points falling into $\mathcal{N}_{\boldsymbol{p}}$ and $\boldsymbol{n}_{\boldsymbol{p}}$ is the pre-computed normal vector of $\boldsymbol{p}$ based on the eigen decomposition as [49]. Note that, $\tilde{\boldsymbol{n}}_{\boldsymbol{p}}$ in $\zeta(\tilde{\boldsymbol{p}}, \tilde{\boldsymbol{n}}_{\boldsymbol{p}}; \tilde{\mathcal{P}})$ can be approximated by $\boldsymbol{n}_{\boldsymbol{p}}$, *i.e.* the normal vector in $\mathcal{P}$ closest to $\tilde{\boldsymbol{p}}$ as a surrogate.

As a result, the imperceptible loss $L_{\text{imp}}(\mathcal{P}, \tilde{\mathcal{P}})$ in equation (3) can be substituted by the following formulation:

$$L_{\text{imp}}(\mathcal{P}, \tilde{\mathcal{P}}) = L_{\text{def}}(\mathcal{T}) + \alpha L_{\text{Cha}}(\mathcal{P}, \tilde{\mathcal{P}}) + \beta L_{\text{Cur}}(\mathcal{P}, \tilde{\mathcal{P}}), \tag{12}$$

where α and β are the trade-off parameters between loss terms. Here, the two loss terms of L_{Cha} and L_{Cur} can avoid suffering from dis-alignment caused by severe geometrical deformation between $\mathcal{P}$ and $\tilde{\mathcal{P}}$.

3.3 Optimization

Optimization of the object function (3) and (4) is carried out by the Stochastic Gradient Descent (SGD). We have $\boldsymbol{S}, \boldsymbol{A}$ and $\boldsymbol{T}$ to be optimized according to the following rule:

$$\begin{aligned} \boldsymbol{S}_j^{t+1} &\leftarrow \boldsymbol{S}_j^t - \eta \nabla (L_{\text{mis}}(\tilde{\mathcal{P}}) + \lambda L_{\text{imp}}(\mathcal{P}, \tilde{\mathcal{P}})) \\ \boldsymbol{A}_j^{t+1} &\leftarrow \boldsymbol{A}_j^t - \eta \nabla (L_{\text{mis}}(\tilde{\mathcal{P}}) + \lambda L_{\text{imp}}(\mathcal{P}, \tilde{\mathcal{P}})) \\ \boldsymbol{T}_j^{t+1} &\leftarrow \boldsymbol{T}_j^t - \eta \nabla (L_{\text{mis}}(\tilde{\mathcal{P}}) + \lambda L_{\text{imp}}(\mathcal{P}, \tilde{\mathcal{P}})) \end{aligned} \tag{13}$$

where ∇ denotes the gradients of the loss with respect to the parametric matrices or vectors that are going to be updated; η is the learning rate and t denotes the t-th iteration.

4 Experiments

4.1 Dataset and Settings

Dataset We use both synthetic (ModelNet40) and real-world (ScanObjectNN) datasets for comparative evaluation on adversarial attack algorithms. ModelNet40 [51] is adopted as existing works [42, 49], which is consisted of 12,311

CAD objects from the 40 common object categories. We take 9,843 CAD models for training and the remaining 2,468 for testing. Due to the imbalanced sample distribution across classes, for the generation of adversarial point clouds, we follow the work [49, 53] to randomly select 25 instances that can be correctly classified by the victim networks from each of the 10 categories in the test set, including airplane, bed, bookshelf, bottle, chair, monitor, sofa, table, toilet, and vase. ScanObjectNN [43] is a recent point cloud object dataset that consists of 2,890 objects from 15 classes. These objects were extracted from real-world indoor scenes. We adopt the **OBJ_ONLY** version of the dataset, which splits 2,309 objects for training and 581 objects for validation. ScanObjectNN poses significant challenges for 3D deep learning due to the presence of background clutter, missing parts and deformations commonly found in the data.

Comparative Attack and Defense Methods We qualitatively and quantitatively compare our attacks with a number of algorithms, including FGM [31], IFGM [13], MIFGM [8], PGD [28], KNN [42], 3d-Adv [53], GeoA^3 [49], and SI-Adv [17]. All adversarial attackers are under the identical untargeted setting [3] (*i.e.* the class having the second-largest logit predicted by victim networks is chosen as the attacking target) for a fair comparison. Diverse recent defenders for adversarial attacks are adopted to verify adversarial robustness, including SOR [63], SRS [64], DUP-Net [64] and IF-defense (ConvONet) [52].

Performance Metrics The attack success rate (ASR) is adopted to evaluate the adversarial effectiveness of comparative attack methods, which is measured by the ratio of adversaries successfully fooling a victim point classifier. To quantitatively measure the visual imperceptibility in terms of geometric smoothness and uniformity of adversarial point clouds, we adopt three performance metrics (*i.e.* the k-Nearest Neighbor Distance $L_{k\mathrm{NN}}$ [42], the Uniform Metric L_{uni} [23] and the consistency of local curvature L_{Cur} [49]) to assess the alteration brought by attack methods.

Implementation Details We uniformly sample 1024 points from the surface for each CAD model via the Farthest Point Sampling, which is normalized into a unit ball. Two rotation-sensitive classifiers (PointNet, and DGCNN) and two rotation equivariant VN-based classifiers (VN-PointNet and VN-DGCNN) are attacked with the maximum gradient iteration of 1000 and a learning rate of 0.01. The batch size is set to 1. Besides, an early stop strategy is adopted, *i.e.* , when our attack succeeded before reaching the pre-defined maximum iteration, the attacker would stop.

4.2 Evaluation on Adversarial Point Clouds

Evaluation on Defense-Free Adversarial Attackers In order to conduct a fair comparison our AdvGT attack with existing methods, we executed various baseline attacks on both synthetic (ModelNet40) and real-world (ScanObjectNN) datasets. The results of these experiments, conducted under a defense-free setting, are presented in TABLE 1 and 2. It is observed that our AdvGT achieves 100% ASR and the least the Uniform Metric L_{uni}. Although the k-Nearest Neighbor Distance $L_{k\mathrm{NN}}$ may not always be the optimal choice, it exhibits superior performance in comparison to a majority of attackers.

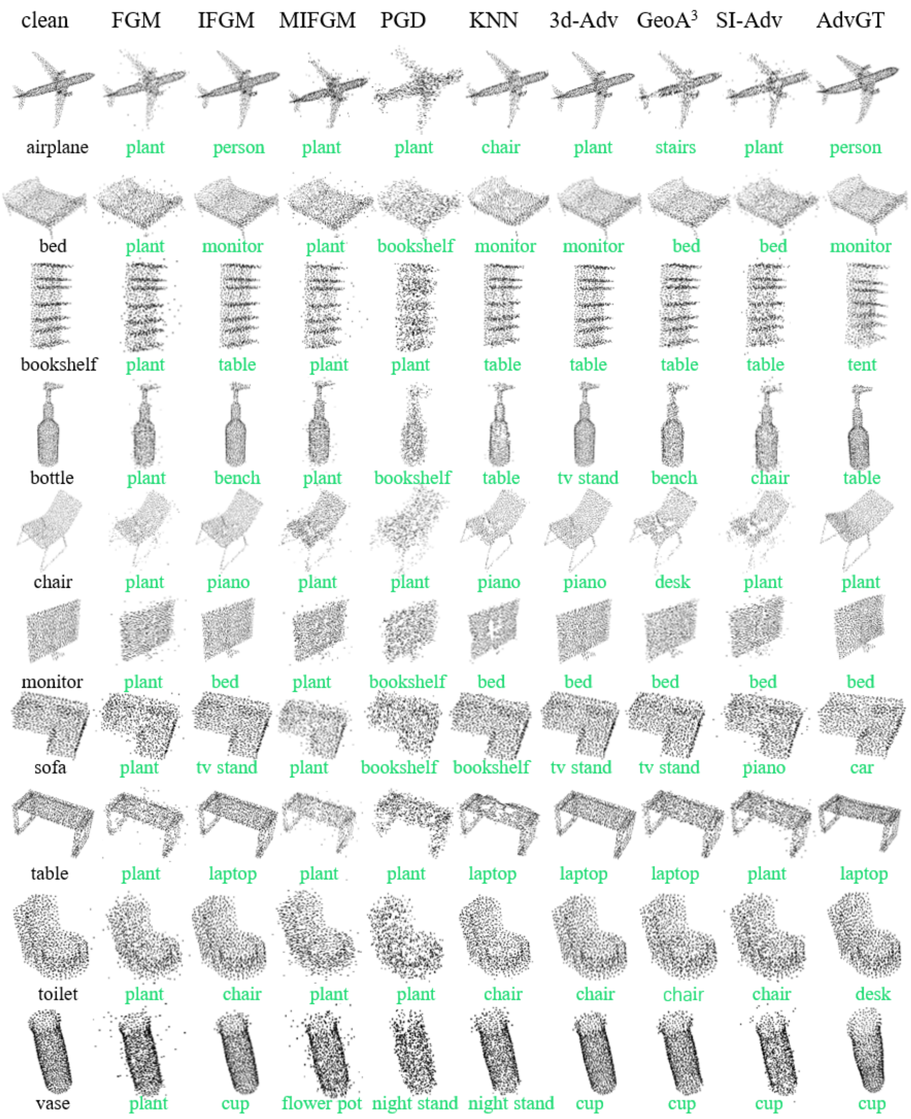

Fig. 2. Adversarial point clouds generated by our proposed methods and other attackers under the defense-free setting against PointNet. All the shown examples here are successful attacks on the victim model with true labels in black, and wrong predicted labels given by the victim highlighted in green.

The visualization of adversarial examples is shown in Fig. 2. It can be observed that our AdvGT induces subtle changes that are imperceptible to the humans, in contrast to the majority of comparative adversarial attackers that exhibit geometric irregularities that result in perceptible changes to humans. These results suggest that our proposed AdvGT algorithm is capable of achieving commendable performance in terms of geometric imperceptibility, as evidenced by its ability to generate adversaries with high level of smoothness and uniformity. This achievement is attributed to our algorithm's ability to preserve the surface properties of point clouds. It can reduce damage during the process of deceiving victim classifiers, owing to the use of highly flexible and diverse transformations.

Evaluation on Attacking under Diverse Defenses To demonstrate the robustness of adversarial attacks against recent defenses, a series of experiments are conducted to compare our attacks with four powerful attack methods that show better performance on the uniform and the k-NN metrics in TABLE 1 and 2. Firstly, a comparative evaluation is carried out under the defense of SOR and SRS, which statistically or randomly drop a subset of points from adversarial point sets. Results of experiments conducted on ModelNet40 and ScanObjectNN respectively reported in Fig. 3 and 4. Our AdvGT algorithm

Table 1. Comparison between our AdvGT and comparative attackers without any defense on attack success rate, the uniform metric L_{uni}, and the k-NN distance $L_{k\text{NN}}$ ($\times 10^{-3}$). The experiments are conducted on ModelNet40.

Attacks	Rotation-Sensitive						Rotation-Agnostic					
	PointNet			DGCNN			VN-PointNet			VN-DGCNN		
	ASR (%)↑	$L_{\text{uni}}\downarrow$	$L_{k\text{NN}}\downarrow$	ASR (%)↑	$L_{\text{uni}}\downarrow$	$L_{k\text{NN}}\downarrow$	ASR (%)↑	$L_{\text{uni}}\downarrow$	$L_{k\text{NN}}\downarrow$	ASR (%)↑	$L_{\text{uni}}\downarrow$	$L_{k\text{NN}}\downarrow$
FGM [31]	99.60	0.46	5.70	**100.00**	0.49	2.10	**100.00**	0.32	7.97	95.60	0.46	3.86
IFGM [13]	**100.00**	0.19	0.57	98.40	0.19	**0.60**	99.60	0.17	0.66	98.40	0.17	0.66
MIFGM [8]	**100.00**	0.45	6.80	**100.00**	0.48	2.30	**100.00**	0.37	5.44	99.60	0.48	2.58
PGD [28]	**100.00**	0.31	1.20	**100.00**	0.30	1.21	**100.00**	0.30	1.21	**100.00**	0.30	1.22
KNN [42]	96.80	0.21	**0.52**	99.60	0.25	0.66	**100.00**	0.25	0.60	**100.00**	0.28	**0.59**
3d-Adv [53]	**100.00**	0.18	0.61	**100.00**	0.19	0.62	**100.00**	0.17	0.64	**100.00**	0.18	0.66
GeoA^3 [49]	**100.00**	0.21	0.69	**100.00**	0.24	0.91	**100.00**	0.19	0.71	**100.00**	0.21	0.73
SI-Adv [17]	**100.00**	0.32	0.90	**100.00**	0.37	1.23	**100.00**	0.27	0.64	**100.00**	0.37	1.18
CTRI [60]	99.60	-	-	95.60	-	-	-	-	-	-	-	-
AdvGT (ours)	**100.00**	**0.15**	0.65	**100.00**	**0.16**	0.69	**100.00**	**0.12**	**0.58**	**100.00**	**0.14**	0.64

Table 2. Comparison between our AdvGT and comparative attackers without any defense on attack success rate, the uniform metric L_{uni}, and the k-NN distance $L_{k\text{NN}}$ ($\times 10^{-3}$). The experiments are conducted on ScanObjectNN.

Attacks	Rotation-Sensitive						Rotation-Agnostic					
	PointNet			DGCNN			VN-PointNet			VN-DGCNN		
	ASR (%)↑	$L_{\text{uni}}\downarrow$	$L_{k\text{NN}}\downarrow$	ASR (%)↑	$L_{\text{uni}}\downarrow$	$L_{k\text{NN}}\downarrow$	ASR (%)↑	$L_{\text{uni}}\downarrow$	$L_{k\text{NN}}\downarrow$	ASR (%)↑	$L_{\text{uni}}\downarrow$	$L_{k\text{NN}}\downarrow$
FGM [31]	90.72	0.42	6.23	84.44	0.46	2.00	80.80	0.45	2.75	86.00	0.42	3.99
IFGM [13]	**100.00**	0.17	0.43	**100.00**	0.17	0.46	**100.00**	0.17	0.48	99.60	0.17	0.48
MIFGM [8]	**100.00**	0.41	6.45	96.36	0.44	2.16	96.80	0.44	2.03	91.20	0.42	2.67
PGD [28]	**100.00**	0.29	1.08	**100.00**	0.28	1.11	**100.00**	0.29	1.11	**100.00**	0.29	1.12
KNN [42]	97.35	0.18	**0.25**	**100.00**	0.27	**0.36**	**100.00**	0.17	**0.42**	99.60	0.18	**0.42**
3d-Adv [53]	**100.00**	0.17	0.43	**100.00**	0.17	0.46	**100.00**	0.17	0.49	**100.00**	0.17	0.48
GeoA^3 [49]	**100.00**	0.20	0.60	**100.00**	0.21	0.71	**100.00**	0.19	0.59	**100.00**	0.19	0.59
SI-Adv [17]	**100.00**	0.30	1.00	99.67	0.38	1.39	99.60	0.36	1.28	98.80	0.34	1.14
CTRI [60]	**100.00**	-	-	99.73	-	-	-	-	-	-	-	-
AdvGT (ours)	**100.00**	**0.15**	0.45	**100.00**	**0.14**	0.44	**100.00**	**0.15**	0.45	**100.00**	**0.14**	0.45

performs similarly well under the two defense methods. Specifically, our method consistently outperforms other competitors with significantly large margins using different ratios of dropping points. These results confirm our motivation that assigning adversarial effects to all points by the transformation operation can significantly improve the robustness of our adversarial point clouds.

Other defense algorithms (*i.e.* SOR*, DUP-Net, and IF-Defense) via shape smoothing and recovery are also employed to prevent adversarial attacks as

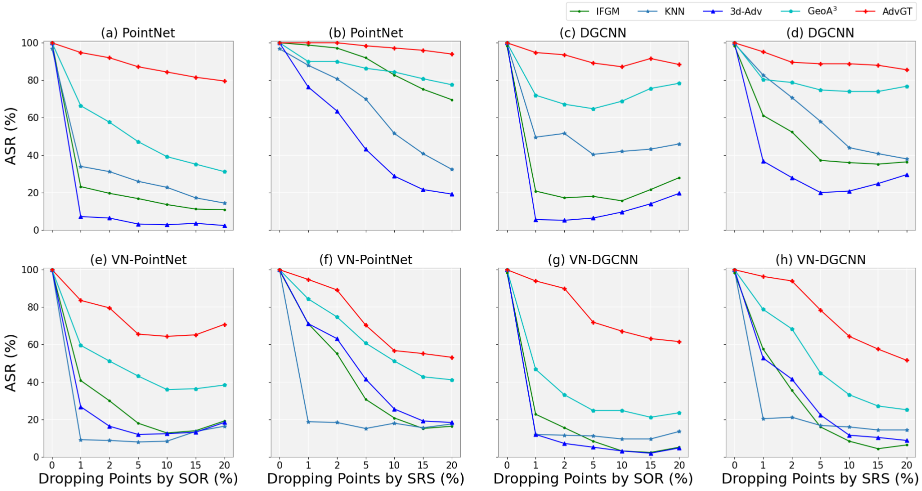

Fig. 3. Attack success rate (%) of IFGM [13], KNN [42], 3d-Adv [53], GeoA^3 [49], and our attack under two defense methods by dropping a range of ratios of points. The experiments are conducted on MoedelNet40.

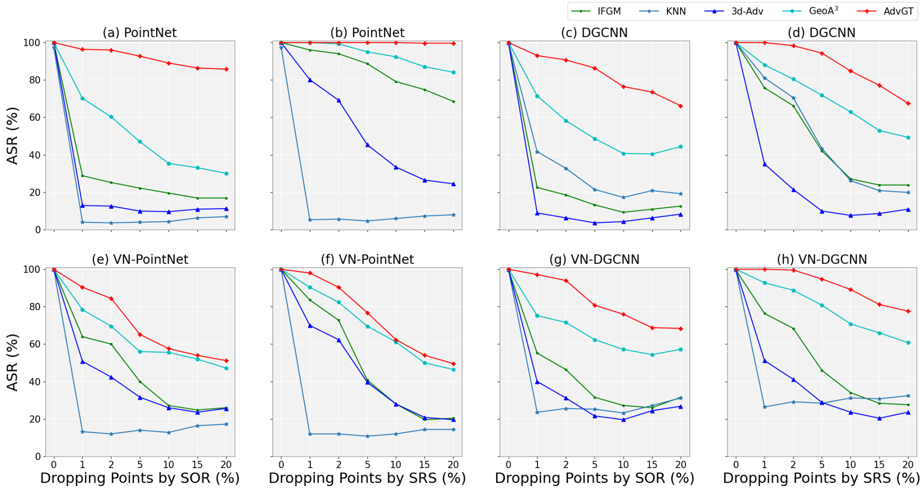

Fig. 4. Attack success rate (%) of IFGM [13], KNN [42], 3d-Adv [53], GeoA^3 [49], and our attack under two defense methods by dropping a range of ratios of points. The experiments are conducted on ScanObjectNN.

presented in Fig. 5 and 6. It is noted that SOR* is another version of SOR. In particular, it improves geometric smoothness of point clouds by computing the mean μ and standard deviation σ of nearest neighbor distances and removing the points which fall outside the $\mu \pm \alpha \times \sigma$, where α (is set to 1.1) determines the size of the analyzed neighborhood. Similar results to those obtained by dropping points are observed, which again verify the effectiveness of our attack algorithms in suppressing diverse state-of-the-art defenses. To conclude, the experiments demonstrate the robustness and effectiveness of our adversarial attack methods against a wide range of defense algorithms and classifiers.

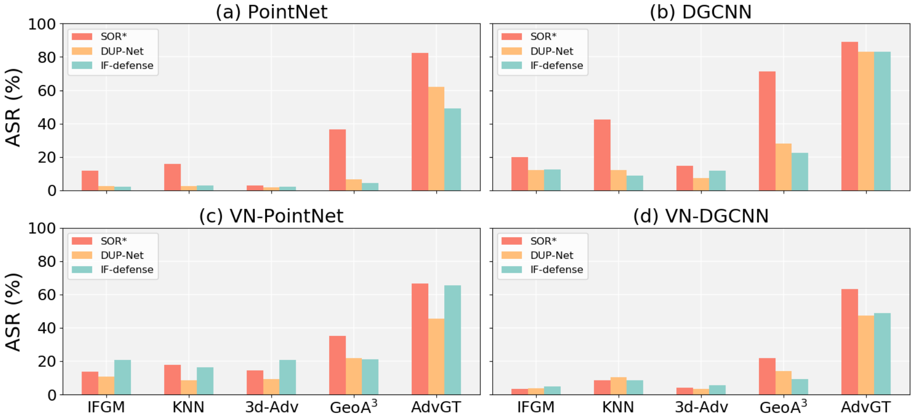

Fig. 5. Attack success rate (%) of IFGM [13], KNN [42], 3d-Adv [53], GeoA^3 [49], and our attack method under SOR*, DUP-Net, and IF-Defense. The experiments are conducted on MoedelNet40.

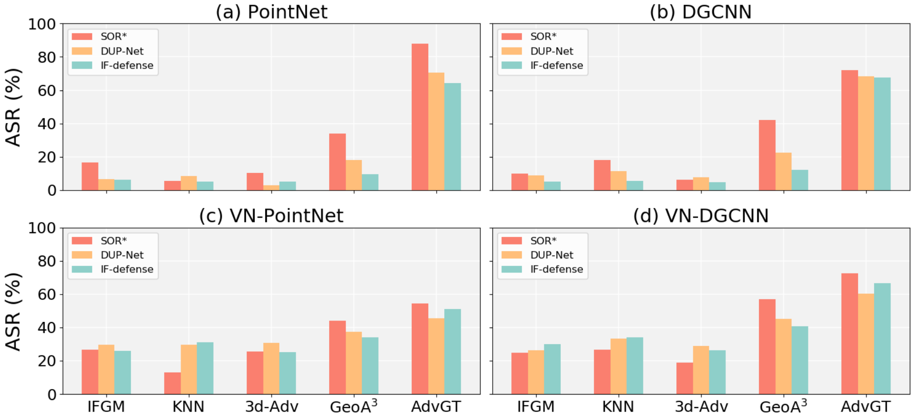

Fig. 6. Attack success rate (%) of IFGM [13], KNN [42], 3d-Adv [53], GeoA^3 [49], and our attack method under SOR*, DUP-Net, and IF-Defense. The experiments are conducted on ScanObjectNN.

Table 3. Attack success rate (%) of re-sampled point clouds of reconstructed shape from adversarial point clouds by comparative methods.

Victims	Rotation-Sensitive		Rotation-Agnostic	
	PointNet	DGCNN	VN-PointNet	VN-DGCNN
KNN [42]	6.00	8.80	18.40	30.00
3d-Adv [53]	3.20	3.60	11.60	16.40
GeoA^3 [49]	18.40	5.20	16.00	23.60
AdvGT (ours)	**92.40**	**90.40**	**41.60**	**57.20**

4.3 Evaluation on Shape and Physical Attack

From a practical perspective, adversarial shapes are more favorable than adversarial point clouds, which inspires the challenging physical attack [49] on re-sampled or re-scanned point clouds from the reconstructed mesh-based surfaces, as surface reconstruction can discount irregular geometries in local regions and cause reduction of adversarial effects. In our experiments, we uniformly sample 10,000 points from each CAD model in ModelNet40 to obtain the input, benign point clouds for the convenience of surface reconstruction, which are optimized under adversarial attack algorithms to generate point-based adversaries. Given adversarial point clouds by comparative attack methods, the Screened Poisson Surface Reconstruction algorithm [19] is adopted to reconstruct mesh-based surface, from which we re-sample 10,000 points as an approximation of adversarial shape for evaluation.

We compare our AdvGT with the state-of-the-art methods [42, 49, 53], in which [42] and [49] studied the physical attack effects. For a fair comparison, the same procedure of meshing and point re-sampling is adopted for all competitors as our AdvGT. Results of generated adversarial shape are shown in TABLE 3. We can find out that the attack success rate dramatically dropped for all

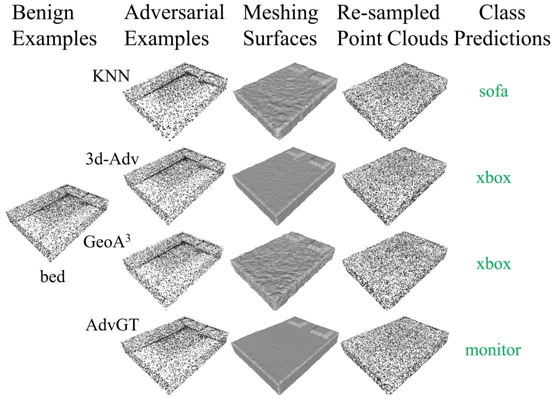

Fig. 7. Visualization of benign examples, adversarial point sets by our AdvGT and three competitors – KNN [42], 3d-Adv [53] and GeoA^3 [49], reconstructed meshes and re-sampled point clouds against VN-PointNet. The classification predictions of re-sampled point sets are also illustrated.

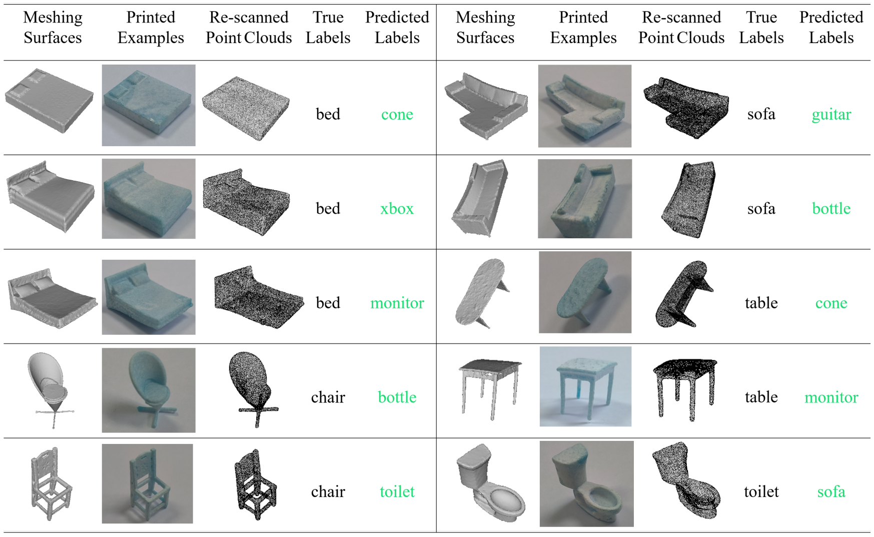

Fig. 8. Visualization of reconstructed surfaces in meshes, 3D-printed examples, and re-scanned point clouds sets, true labels and class predictions against our AdvGT. The objects are generated by VN-PointNet.

the comparative methods, while our method consistently performs better on the attack success rate (at least 25.60% higher) when attacking four point classifiers. Visualization of an adversarial example is given in Fig. 7, where the reconstructed meshes of our AdvGT possess the best quality with more regular surfaces than the other three point-wise deformation based attacks.

For a real test of physical attack, among all the successfully attacked testing instances from our adversaries, we choose some examples to reconstruct meshes via 3D printing and then re-scan each printed object with a 3D scanner to mimic the procedure of acquiring point clouds from adversarial shape in real world. Considering that the process of printing and re-scanning the adversarial samples changes their pose, and therefore may affect the classification accuracy of the rotation-sensitive classifiers, we select 20 adversarial samples for both rotation-equivariant classifiers respectively, *i.e.* VN-PointNet and VN-DGCNN. All re-scanned point clouds can be mis-classified to other classes (*i.e.* achieving 100% on the ASR in the real test), which can verify the effectiveness of our algorithm for adversarial physical attack. The visualization of reconstructed surfaces in meshes, 3D-printed examples, and re-scanned point clouds is reported in Fig. 8.

4.4 Ablation Studies

We conduct an ablation study about the effects of implicit and explicit similarities between adversarial and benign point clouds in geometric imperceptible objectives L_{imp}. The results are reported in TABLE 4. In most instances, using

Table 4. Attack success rate (%) and L_{Cur} ($\times 10^{-2}$) of ablation studies with AdvGT.

Metrics	PointNet				DGCNN			
	w/o $L_{\mathcal{T}}$	w/o L_{Cha}	w/o L_{Cur}	with all	w/o $L_{\mathcal{T}}$	w/o L_{Cha}	w/o L_{Cur}	with all
ASR↑	**100.00**	99.60	99.60	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
L_{Cur} ↓	2.19	2.20	2.19	**1.72**	1.25	1.31	1.25	**1.24**
Metrics	VN-PointNet				VN-DGCNN			
	w/o $L_{\mathcal{T}}$	w/o L_{Cha}	w/o L_{Cur}	with all	w/o $L_{\mathcal{T}}$	w/o L_{Cha}	w/o L_{Cur}	with all
ASR↑	99.60	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
L_{Cur} ↓	2.84	2.84	2.83	**2.70**	2.83	**2.77**	2.80	2.85

all three loss terms performs better than those degenerated ones removing any objective on the ASR and L_{Cur}.

Besides, we conduct one more experiment about the effects of varying m of pre-defined anchors, whose results are shown in TABLE 5. We observed that as m increases, achieving a 100% success rate becomes increasingly challenging for all victim networks. Meanwhile, the value of L_{Cur} exhibits a trend of first decreasing and then increasing with the growth of m. Interestingly, this turning point occurs earlier in the PointNet and VN-PointNet networks compared to the DGCNN and VN-DGCNN networks. The reason for this discrepancy may lie in the different network structures of PointNet and DGCNN, where the former only considers global information, while the latter takes into account both local and global information. In the attack method we proposed, a smaller m, representing fewer transformation anchor points, is more likely to disrupt the global shape of the point cloud, while more anchor points are more likely to damage its local detail information.

To verify the effect of hyper-parameters σ, attack results with different σ is adopted during adversarial attacking are shown in TABLE 6. Insights drawn

Table 5. Attack success rate (%) and L_{Cur} ($\times 10^{-2}$) of ablation studies with AdvGT, the experiments are conducted with different anchor points m.

Metrics	PointNet							
	m=2	m=4	m=6	m=8	m=10	m=12	m=14	m=16
ASR↑	99.20	**100.00**	99.20	99.20	99.60	99.20	99.60	99.60
L_{Cur} ↓	2.81	**1.72**	2.07	2.01	2.07	2.53	2.54	2.54
Metrics	VN-PointNet							
	m=2	m=4	m=6	m=8	m=10	m=12	m=14	m=16
ASR↑	99.60	**100.00**	99.60	98.80	99.20	94.80	96.00	96.80
L_{Cur} ↓	3.45	2.70	2.60	**2.41**	2.48	2.46	2.49	2.54
Metrics	DGCNN							
	m=2	m=4	m=6	m=8	m=10	m=12	m=14	m=16
ASR↑	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	94.00	92.80	94.00
L_{Cur} ↓	1.82	1.24	1.19	1.15	**1.14**	3.88	3.81	3.78
Metrics	VN-DGCNN							
	m=2	m=4	m=6	m=8	m=10	m=12	m=14	m=16
ASR↑	**100.00**	**100.00**	**100.00**	**100.00**	99.60	99.20	99.60	98.80
L_{Cur} ↓	3.76	2.85	2.48	2.38	**2.37**	2.39	2.42	2.41

Table 6. Attack success rate (%) and L_{Cur} ($\times 10^{-2}$) of ablation studies with AdvGT, the experiments are conducted with different σ.

Metrics	PointNet									
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.00
ASR↑	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	99.60	98.00	97.20	96.40	96.00
L_{Cur} ↓	2.25	2.24	1.93	1.78	1.72	**1.67**	1.69	1.74	1.80	2.39
Metrics	VN-PointNet									
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.00
ASR↑	**100.00**	**100.00**	**100.00**	99.60	**100.00**	96.00	90.80	91.20	90.40	89.20
L_{Cur} ↓	2.82	3.22	3.28	2.85	2.70	2.42	**1.91**	1.93	1.93	1.96
Metrics	DGCNN									
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.00
ASR↑	**100.00**	**100.00**	99.20	99.20	**100.00**	99.60	99.60	**100.00**	99.20	**100.00**
L_{Cur} ↓	3.74	3.18	2.35	2.19	**1.24**	1.25	1.30	1.34	1.45	1.37
Metrics	VN-DGCNN									
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.00
ASR↑	**100.00**	**100.00**	**100.00**	99.60	**100.00**	97.60	92.80	88.40	90.00	93.20
L_{Cur} ↓	4.01	4.27	3.67	3.24	2.85	**2.35**	2.37	2.46	2.45	2.49

from the table indicate that when the value of σ is small, the attack is more effective. However, as σ increases, the effectiveness of the attack diminishes. Interestingly, L_{Cur} exhibits a trend of initial decrease followed by an increase with σ. This can be attributed to the characteristics of the Gaussian curve. When σ is small, the Gaussian curve is steep, leading to prominent local deformations in the adversarial point cloud. These deformations result in a larger L_{Cur} and are easily detected by the victim classifiers, thereby facilitating a successful attack. Conversely, when σ is large, the Gaussian curve is smoother, making the deformations in the adversarial point cloud less noticeable to the victim classifier. This necessitates more deformations for a successful attack, resulting in a larger L_{Cur}, and hence, larger σ values also result in less effective attacks.

In an effort to investigate the distinct impacts of our chosen three transformation methods on the attack performance, and to ascertain the necessity of each, we conducted an ablation study on the process of point cloud transformation. The results are presented in 7. It can be observed from the table that

Table 7. Attack success rate (%) and L_{Cur} ($\times 10^{-2}$) of ablation studies with AdvGT, the experiments are conducted with different transformation. R, S and T respectively represent rotation, scaling, and translation.

Metrics	PointNet				DGCNN			
	w/o $\boldsymbol{R}$	w/o $\boldsymbol{S}$	w/o $\boldsymbol{T}$	with all	w/o $\boldsymbol{R}$	w/o $\boldsymbol{S}$	w/o $\boldsymbol{T}$	with all
ASR↑	98.80	91.60	92.40	**100.00**	88.80	71.20	79.20	**100.00**
L_{Cur} ↓	1.78	1.87	1.73	**1.72**	4.13	3.84	2.92	**1.24**
Metrics	VN-PointNet				VN-DGCNN			
	w/o $\boldsymbol{R}$	w/o $\boldsymbol{S}$	w/o $\boldsymbol{T}$	with all	w/o $\boldsymbol{R}$	w/o $\boldsymbol{S}$	w/o $\boldsymbol{T}$	with all
ASR↑	94.40	81.20	88.40	**100.00**	97.60	79.20	84.40	**100.00**
L_{Cur} ↓	2.75	2.89	2.86	**2.70**	2.88	2.88	3.05	**2.85**

the attack performance using all transformation methods surpasses that of any other combination where one or more transformations are omitted. One possible interpretation of these results is that the combination of the three transformation methods leads to a more diversified overall transformation of the adversarial point clouds, thereby enhancing the attack capability.

5 Conclusions

In this paper, we propose a more practical setting on generating adversaries in the physical adversarial attack. Different from existing point-wise deformation based attackers, our AdvGT can favour more rational and imperceptible adversarial shape via non-rigid transformation, whose effectiveness can be verified in our experiments. More importantly, owing to the transformation-based nature of our AdvGT, adversarial effects can be shared among all points and good geometric smoothness and uniformity of adversaries can be achieved, which thus ensures the survival of our AdvGT under recent defenses and adversarial physical attacks. However, we acknowledge that our study is primarily focused on the task of point cloud classification. It does not extend to other critical tasks such as point cloud recognition and semantic segmentation, thereby presenting potential limitations. Nevertheless, these limitations also highlight potential directions for future research.

Acknowledgement

This work is supported in part by the Guangdong Pearl River Talent Program (Introduction of Young Talent) under Grant No. 2019QN01X246, the Guangdong Basic and Applied Basic Research Foundation under Grant No. 2022A1515011549 and the Major Key Project of Peng Cheng Laboratory under Grant No. PCL2023A08.

References

1. Akgul, O., Penekli, H.I., Genc, Y.: Applying deep learning in augmented reality tracking. In: 2016 12th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS). pp. 47–54. IEEE (2016)
2. Al-Qizwini, M., Barjasteh, I., Al-Qassab, H., Radha, H.: Deep learning algorithm for autonomous driving using googlenet. In: 2017 IEEE Intelligent Vehicles Symposium (IV). pp. 89–96. IEEE (2017)
3. Carlini, N., Wagner, D.: Towards evaluating the robustness of neural networks. In: 2017 IEEE Symposium on Security and Privacy (SP). pp. 39–57. IEEE Computer Society (2017)
4. Chen, H., Liu, S., Chen, W., Li, H., Hill, R.: Equivariant point network for 3d point cloud analysis. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 14514–14523 (2021)

5. Chen, Y., Wang, Z., Zou, L., Chen, K., Jia, K.: Quasi-balanced self-training on noise-aware synthesis of object point clouds for closing domain gap. In: Computer Vision–ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part XXXIII. pp. 728–745. Springer (2022)
6. Cohen, T.S., Geiger, M., Köhler, J., Welling, M.: Spherical cnns. In: International Conference on Learning Representations (2018)
7. Deng, C., Litany, O., Duan, Y., Poulenard, A., Tagliasacchi, A., Guibas, L.J.: Vector neurons: A general framework for so (3)-equivariant networks. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 12200–12209 (2021)
8. Dong, Y., Liao, F., Pang, T., Su, H., Zhu, J., Hu, X., Li, J.: Boosting adversarial attacks with momentum. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 9185–9193. IEEE (2018)
9. Esteves, C., Allen-Blanchette, C., Makadia, A., Daniilidis, K.: Learning so (3) equivariant representations with spherical cnns. In: Proceedings of the European Conference on Computer Vision (ECCV). pp. 52–68 (2018)
10. Fan, H., Su, H., Guibas, L.: A point set generation network for 3d object reconstruction from a single image. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 2463–2471. IEEE Computer Society (2017)
11. Feng, Y., Zhang, Z., Zhao, X., Ji, R., Gao, Y.: Gvcnn: Group-view convolutional neural networks for 3d shape recognition. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 264–272 (2018)
12. Fujiyoshi, H., Hirakawa, T., Yamashita, T.: Deep learning-based image recognition for autonomous driving. IATSS research **43**(4), 244–252 (2019)
13. Goodfellow, I.J., Shlens, J., Szegedy, C.: Explaining and harnessing adversarial examples. stat **1050**, 20 (2015)
14. Grigorescu, S., Trasnea, B., Cocias, T., Macesanu, G.: A survey of deep learning techniques for autonomous driving. Journal of Field Robotics **37**(3), 362–386 (2020)
15. Hamdi, A., Rojas, S., Thabet, A., Ghanem, B.: Advpc: Transferable adversarial perturbations on 3d point clouds. In: Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XII 16. pp. 241–257. Springer (2020)
16. Hu, Q., Liu, D., Hu, W.: Exploring the devil in graph spectral domain for 3d point cloud attacks. In: Computer Vision–ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part III. pp. 229–248. Springer (2022)
17. Huang, Q., Dong, X., Chen, D., Zhou, H., Zhang, W., Yu, N.: Shape-invariant 3d adversarial point clouds. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 15335–15344 (2022)
18. Kán, P., Kafumann, H.: Deeplight: light source estimation for augmented reality using deep learning. The Visual Computer **35**(6), 873–883 (2019)
19. Kazhdan, M., Hoppe, H.: Screened poisson surface reconstruction. ACM Transactions on Graphics (TOG) **32**(3), 1–13 (2013)
20. Kim, S., Lee, S., Hwang, D., Lee, J., Hwang, S.J., Kim, H.J.: Point cloud augmentation with weighted local transformations. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 548–557 (2021)
21. Lalonde, J.F.: Deep learning for augmented reality. In: 2018 17th Workshop on Information Optics (WIO). pp. 1–3. IEEE (2018)

22. Le, T., Duan, Y.: Pointgrid: A deep network for 3d shape understanding. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 9204–9214 (2018)
23. Li, R., Li, X., Fu, C.W., Cohen-Or, D., Heng, P.A.: PU-GAN: a point cloud upsampling adversarial network. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 7203–7212 (2019)
24. Li, X., Li, R., Chen, G., Fu, C.W., Cohen-Or, D., Heng, P.A.: A rotation-invariant framework for deep point cloud analysis. IEEE Transactions on Visualization and Computer Graphics (2021)
25. Liu, B., Zhang, J., Zhu, J.: Boosting 3d adversarial attacks with attacking on frequency. IEEE Access **10**, 50974–50984 (2022)
26. Liu, M., Yao, F., Choi, C., Sinha, A., Ramani, K.: Deep learning 3d shapes using alt-az anisotropic 2-sphere convolution. In: International Conference on Learning Representations (2018)
27. Ma, C., Guo, Y., Yang, J., An, W.: Learning multi-view representation with lstm for 3-d shape recognition and retrieval. IEEE Transactions on Multimedia **21**(5), 1169–1182 (2018)
28. Madry, A., Makelov, A., Schmidt, L., Tsipras, D., Vladu, A.: Towards deep learning models resistant to adversarial attacks. In: International Conference on Learning Representations (2018)
29. Maturana, D., Scherer, S.: Voxnet: A 3d convolutional neural network for real-time object recognition. In: 2015 IEEE/RSJ international conference on intelligent robots and systems (IROS). pp. 922–928. IEEE (2015)
30. Miao, Y., Pajarola, R., Feng, J.: Curvature-aware adaptive re-sampling for point-sampled geometry. Computer-Aided Design **41**(6), 395–403 (2009)
31. Miyato, T., Dai, A.M., Goodfellow, I.: Adversarial training methods for semi-supervised text classification. In: International Conference on Learning Representations (2017)
32. Moosavi-Dezfooli, S.M., Fawzi, A., Fawzi, O., Frossard, P.: Universal adversarial perturbations. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 1765–1773 (2017)
33. Moosavi-Dezfooli, S.M., Fawzi, A., Frossard, P.: Deepfool: a simple and accurate method to fool deep neural networks. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 2574–2582 (2016)
34. Ponomarenko, N., Ieremeiev, O., Lukin, V., Egiazarian, K., Carli, M.: Modified image visual quality metrics for contrast change and mean shift accounting. In: 2011 11th International Conference The Experience of Designing and Application of CAD Systems in Microelectronics (CADSM). pp. 305–311 (2011)
35. Qi, C.R., Su, H., Mo, K., Guibas, L.J.: Pointnet: Deep learning on point sets for 3d classification and segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 652–660 (2017)
36. Qi, C.R., Yi, L., Su, H., Guibas, L.J.: Pointnet++ deep hierarchical feature learning on point sets in a metric space. In: Proceedings of the 31st International Conference on Neural Information Processing Systems. pp. 5105–5114 (2017)
37. Rao, Y., Lu, J., Zhou, J.: Spherical fractal convolutional neural networks for point cloud recognition. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 452–460 (2019)
38. Riegler, G., Osman Ulusoy, A., Geiger, A.: Octnet: Learning deep 3d representations at high resolutions. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 3577–3586 (2017)

39. Su, H., Maji, S., Kalogerakis, E., Learned-Miller, E.: Multi-view convolutional neural networks for 3d shape recognition. In: 2015 IEEE International Conference on Computer Vision (ICCV). pp. 945–953 (2015)
40. Szegedy, C., Zaremba, W., Sutskever, I., Bruna, J., Erhan, D., Goodfellow, I., Fergus, R.: Intriguing properties of neural networks. arXiv preprint arXiv:1312.6199 (2013)
41. Thomas, N., Smidt, T., Kearnes, S., Yang, L., Li, L., Kohlhoff, K., Riley, P.: Tensor field networks: Rotation-and translation-equivariant neural networks for 3d point clouds. arXiv preprint arXiv:1802.08219 (2018)
42. Tsai, T., Yang, K., Ho, T.Y., Jin, Y.: Robust adversarial objects against deep learning models. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 34, pp. 954–962 (2020)
43. Uy, M.A., Pham, Q.H., Hua, B.S., Nguyen, T., Yeung, S.K.: Revisiting point cloud classification: A new benchmark dataset and classification model on real-world data. In: Proceedings of the IEEE/CVF international conference on computer vision. pp. 1588–1597 (2019)
44. Wang, P.S., Liu, Y., Guo, Y.X., Sun, C.Y., Tong, X.: O-cnn: Octree-based convolutional neural networks for 3d shape analysis. ACM Transactions On Graphics (TOG) **36**(4), 1–11 (2017)
45. Wang, Y., Sun, Y., Liu, Z., Sarma, S.E., Bronstein, M.M., Solomon, J.M.: Dynamic graph cnn for learning on point clouds. ACM Transactions On Graphics (TOG) **38**(5), 1–12 (2019)
46. Watson, G.S.: Smooth regression analysis. Sankhyā: The Indian Journal of Statistics, Series A pp. 359–372 (1964)
47. Wei, X., Yu, R., Sun, J.: View-gcn: View-based graph convolutional network for 3d shape analysis. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 1850–1859 (2020)
48. Weiler, M., Geiger, M., Welling, M., Boomsma, W., Cohen, T.S.: 3d steerable cnns: Learning rotationally equivariant features in volumetric data. Advances in Neural Information Processing Systems **31** (2018)
49. Wen, Y., Lin, J., Chen, K., Chen, C.P., Jia, K.: Geometry-aware generation of adversarial point clouds. IEEE Transactions on Pattern Analysis & Machine Intelligence **44**(06), 2984–2999 (2022)
50. Wicker, M., Kwiatkowska, M.: Robustness of 3d deep learning in an adversarial setting. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 11767–11775 (2019)
51. Wu, Z., Song, S., Khosla, A., Yu, F., Zhang, L., Tang, X., Xiao, J.: 3d shapenets: A deep representation for volumetric shapes. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 1912–1920 (2015)
52. Wu, Z., Duan, Y., Wang, H., Fan, Q., Guibas, L.J.: If-defense: 3d adversarial point cloud defense via implicit function based restoration. arXiv preprint arXiv:2010.05272 (2020)
53. Xiang, C., Qi, C.R., Li, B.: Generating 3d adversarial point clouds. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 9136–9144 (2019)
54. Xiang, T., Zhang, C., Song, Y., Yu, J., Cai, W.: Walk in the cloud: learning curves for point clouds shape analysis. In: 2021 IEEE/CVF International Conference on Computer Vision (ICCV). pp. 895–904. IEEE Computer Society (2021)
55. Xu, M., Ding, R., Zhao, H., Qi, X.: Paconv: Position adaptive convolution with dynamic kernel assembling on point clouds. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 3173–3182 (2021)

56. Yang, B., Wang, J., Clark, R., Hu, Q., Wang, S., Markham, A., Trigoni, N.: Learning object bounding boxes for 3d instance segmentation on point clouds. In: Proceedings of the 33rd International Conference on Neural Information Processing Systems. pp. 6740–6749 (2019)
57. Yang, J., Zhang, Q., Fang, R., Ni, B., Liu, J., Tian, Q.: Adversarial attack and defense on point sets. arXiv preprint arXiv:1902.10899 (2019)
58. Yuan, W., Held, D., Mertz, C., Hebert, M.: Iterative transformer network for 3d point cloud. arXiv preprint arXiv:1811.11209 (2018)
59. Zhang, Z., Hua, B.S., Chen, W., Tian, Y., Yeung, S.K.: Global context aware convolutions for 3d point cloud understanding. In: 2020 International Conference on 3D Vision (3DV). pp. 210–219. IEEE (2020)
60. Zhao, Y., Wu, Y., Chen, C., Lim, A.: On isometry robustness of deep 3d point cloud models under adversarial attacks. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 1201–1210 (2020)
61. Zheng, T., Chen, C., Yuan, J., Li, B., Ren, K.: Pointcloud saliency maps. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 1598–1606 (2019)
62. Zhou, H., Chen, D., Liao, J., Chen, K., Dong, X., Liu, K., Zhang, W., Hua, G., Yu, N.: Lg-gan: Label guided adversarial network for flexible targeted attack of point cloud based deep networks. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 10356–10365 (2020)
63. Zhou, H., Chen, K., Zhang, W., Fang, H., Zhou, W., Yu, N.: Deflecting 3d adversarial point clouds through outlier-guided removal. arXiv preprint arXiv:1812.11017 (2018)
64. Zhou, H., Chen, K., Zhang, W., Fang, H., Zhou, W., Yu, N.: Dup-net: Denoiser and upsampler network for 3d adversarial point clouds defense. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 1961–1970 (2019)
65. Zou, L., Tang, H., Chen, K., Jia, K.: Geometry-aware self-training for unsupervised domain adaptation on object point clouds. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 6403–6412 (2021)