Hiding Imperceptible Noise in Curvature-Aware Patches for 3D Point Cloud Attack

Mingyu Yang[1*], Daizong Liu[2*], Keke Tang[3], Pan Zhou[1†], Lixing Chen[4†], and Junyang Chen[5]

[1] Huazhong University of Science and Technology, Wuhan, China
[2] Peking University, Beijing, China
[3] Guangzhou University, Guangzhou, China
[4] Shanghai Jiao Tong University, Shanghai, China
[5] Shenzhen University, Shenzhen, China
{mingyu_yang,panzhou}@hust.edu.cn,dzliu@stu.pku.edu.cn,
tangbohutbh@gmail.com,lxchen@sjtu.edu.cn,junyangchen@szu.edu.cn

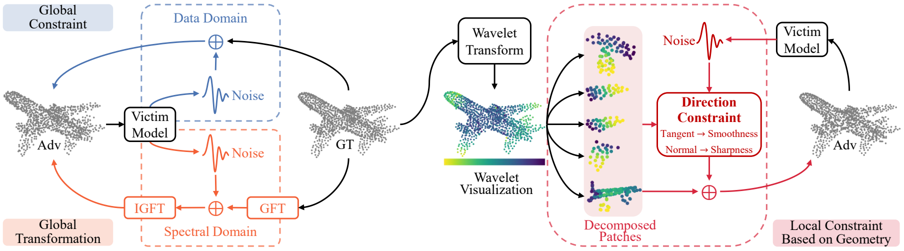
Abstract. With the maturity of depth sensors, point clouds have received increasing attention in various 3D safety-critical applications, while deep point cloud learning models have been shown to be vulnerable to adversarial attacks. Most existing 3D attackers rely on implicit global distance losses to perturb whole points, failing to restrict the proper 3D geometry as point clouds are highly structured. To this end, in this paper, we propose a novel Wavelet Patches Attack (WPA), which leverages local spectral attributes to identify curvature-aware patches for hiding imperceptible perturbations aligned with their local geometric characteristics. Specifically, WPA first transforms the point cloud into the spectral domain using a wavelet operator, obtaining potential geometric structures in different local regions. Each wavelet corresponds to different curvature contexts of local points. Then, by decomposing the 3D object with different curvature-aware levels through the wavelet coefficients, we can perceive the local geometric characteristics and get various curvature-consistent patches. At last, based on the curvature variations of patches, WPA introduces two-type perturbations along the tangent plane and normal vector direction to hide imperceptible noise in slow- and fast-variation patches for preserving the geometric-sensitive local characteristics of smoothness and sharpness, respectively. Experiments demonstrate the superior imperceptibility of our attack method, achieving favorable results on existing 3D classification models while exhibiting robust resistance to various defense mechanisms.

Keywords: Point Cloud · Adversarial Attack · Wavelet Transform

1 Introduction

Deep neural networks have shown to be vulnerable to adversarial examples [10, 35], which add visually indistinguishable perturbations to network inputs but

[*] Mingyu Yang and Daizong Liu contributed equally to this paper. [†] Corresponding authors: Pan Zhou and Lixing Chen



(a) Existing two types of methods generating global perturbation.

(b) Utilizing wavelet tools to introduce imperceptible noise onto geometric-sensitive points.

Fig. 1: Motivation of our proposed attack. Different from perturbing whole points with global constraint or transformation, we utilize local spectral filters to perceive detailed local structures, then hide imperceptible noise in a subset of geometric-sensitive points, to preserve local geometric contexts for generating high-quality adversarial examples.

lead to incorrect prediction results. In the realm of 2D image, researches [4,18,26,39] on adversarial attacks have achieved considerable progress, employing methods that add pixel-wise noise in the spatial or feature domain. However, as the 3D point cloud plays an important role in autonomous driving [3], robotics [24], healthcare [34], etc., the study of robustness in 3D models becomes increasingly critical. Yet, investigations into attacks on 3D deep models are still relatively underexplored. Moreover, unlike 2D images [6–9], the unordered 3D data presents more challenges for adversarial attacks on deep-learning models.

Most existing 3D adversarial attack methodologies [11, 38, 42, 46, 49, 52–54] generally adapt 2D adversarial techniques to the 3D scenario. Some of them [42, 46,49,53] adhere to the point addition/dropping framework, which identifies and modifies critical points within the point cloud to distort the whole representative features. Other works [1, 11, 19, 23, 25, 36, 38, 41, 51] follow the C&W framework [10] to globally perturb the point clouds' Euclidean coordinates through the optimization of the gradients in end-to-end. As illustrated in the upper of Fig. 1(a), these approaches generally employ global distance constraints, such as Chamfer and Hausdorff distances in the data domain, to add noise while implicitly preserving the original shape. Although they achieve high attack success rates, the noise applied under global constraints fails to restrict the proper 3D geometry as point clouds are highly structured, easily disrupting the structural dependencies among neighboring points. Recently, some methods [14, 20, 21, 37] have tried to utilize spectral tools to explicitly preserve the 3D geometries via frequency analysis. As demonstrated in the lower of Fig. 1(a), these methods utilize the Graph Fourier Transform(GFT) [12] to transform point clouds into the graph spectral domain, introducing noise to specific frequency bands to generate adversarial samples. However, GFT is a global transformation, which leads to changes across all points in the data domain, still failing to preserve the detailed point-to-point geometric dependency of distinct local regions.

To alleviate the above issues, we endeavor to first decompose the point cloud into local regions with distinct geometric structures, then introduce to preserve the topology of geometric-sensitive regions to achieve heightened imperceptibility. Specifically, we incorporate the Spectral Graph Wavelet Transform from graph spectral tools [12,15,30] for analyzing the local geometric structures of the point cloud data. As depicted in Fig. 1(b), given that the wavelet kernel predominantly affects a point and its neighbor area, it can decompose the 3D object into curvature-aware patches according to the local geometries. By integrating these patch-level geometric contexts, we selectively add noise to geometric-sensitive patches with distinctive curvature characteristics, such as smooth or sharp regions. Then, we introduce to minimize the perturbation size of these noises to preserve corresponding smoothness and sharpness. In this way, we can generate adversarial samples by perturbing only a subset of points, while perceiving the local geometric information for completely preserving the original 3D shape.

To this end, in this paper, we propose a novel Wavelet Patches Attack (WPA) method, which adeptly utilizes local spectral properties to identify curvature-aware patches, thereby hiding imperceptible perturbations that align with their local geometric characteristics. WPA starts by transforming the point cloud into the spectral domain via the wavelet operator, thus unveiling the potential geometric structures within various local points. Each wavelet corresponds to different curvature contexts of local points. Following this, through the wavelet coefficients, the 3D object can be decomposed into patches with different curvature-aware levels, enabling perceiving the local geometric characteristics and getting various curvature-consistent patches. Ultimately, based on the curvature variations of different patches, WPA introduces two-type perturbations along the tangent plane and normal vector direction to hide imperceptible noise in slow- and fast-variation patches for preserving the patch-wise local characteristics of smoothness and sharpness, respectively.

Our main contributions are summarized as follows:

- We introduce a novel Wavelet Patches Attack (WPA), a technique capable of analyzing and capturing the geometric characteristics of various local regions within a point cloud, and adeptly hiding perturbations within areas of specific geometric structures. Unlike previous attacks that employed global constraints or transformations, our method better preserves the local geometric context, achieving superior imperceptibility.
- By perturbing only a subset of points, we have achieved an exceptionally high success rate of attack, while maintaining perturbation sizes that are equal to or even lower than those of other methods. Distinct from approaches that utilize deep learning to identify critical points for noise addition, our strategy employs traditional signal processing and geometric methods to locate sensitive points, offering greater interpretability.
- We conduct extensive experiments using popular 3D classification models on the ModelNet40 and ShapeNetPart datasets, validating the effectiveness of WPA. Moreover, we demonstrate WPA's robustness against current point cloud defense mechanisms and its superior performance compared to other attacks.

2 Related Work

Adversarial attacks on 3D point cloud. Deep neural networks are vulnerable to adversarial examples, which has been extensively explored in the 2D image domain [27–29]. Recently, many works [11, 22, 37, 38, 42, 46, 49, 52–54] have been adapted to 2D adversarial attacks in the 3D vision community, which can be mainly divided into two categories: 1) point-addition/dropping attack: Xiang *et al.* [46] proposed point generation attacks by adding a limited number of synthesized points/clusters/objects to a point cloud, and showed its effectiveness in attacking the PointNet model [31]. Recently, more works [42, 49, 53] utilize gradient-guided attack methods to identify critical points in point clouds for modification, addition, and deletion. Their goal is to add or remove key points that can be identified by calculating the label-dependent importance score referring to the calculated gradient. 2) point perturbation attack: Previous point-wise perturbation attacks [38, 41] learn to perturb xyz coordinates of each point by adopting the C&W framework [1] based on the Chamfer and Hausdorff distances with additional consideration of the benign distribution of points. The subsequent works [11, 23, 25, 51] further applied the iterative gradient method to achieve more fine-grained adversarial perturbation. Since 3D objects generally contain complex and diverse geometric characteristics in different regions, previous works fail to perceive corresponding different types of point-to-point dependencies and directly perturb the whole point clouds, resulting in low imperceptibility. In this paper, we make the first attempt to distinguish the local geometric characteristics of different regions of a 3D object, and only hide the imperceptible noise in certain regions for preserving corresponding geometrics.

Spectral methods for 3D point cloud. Numerous methodologies have been developed that use spectral information to understand point clouds. For instance, several 3D denoising techniques [33, 50] convert the input point cloud into the graph spectral domain, wherein the rough shape of a point cloud is encapsulated within low-frequency components. Consequently, the spectral filter facilitates the reconstruction of the point cloud's original, noise-free structure. Additionally, various applications [2, 32] capture the intricate details of point clouds via transformed high-frequency components, employing these to identify contours or eliminate redundant information. Recently, [14,22,37] has employed the Graph Fourier Transform (GFT) to transpose point clouds into the spectral domain for generating adversarial examples. Specifically, [14] perturbs specific frequency bands to add perturbation that preserve geometric characteristics. [37] fuses two point clouds from different classes in the spectral domain for obtaining the decision boundary. [22] adds global noise to all points to preserve geometric shapes globally, perturbing sharp or smooth regions in random directions. However, as the GFT is a type of global transformation, [14, 22, 37] are incapable of aligning perturbations with the local geometric structures. In contrast, our utilization of wavelets' pronounced locality aids in identifying the local regions corresponding to each point, thereby enabling us to hide imperceptible noise in specific points.

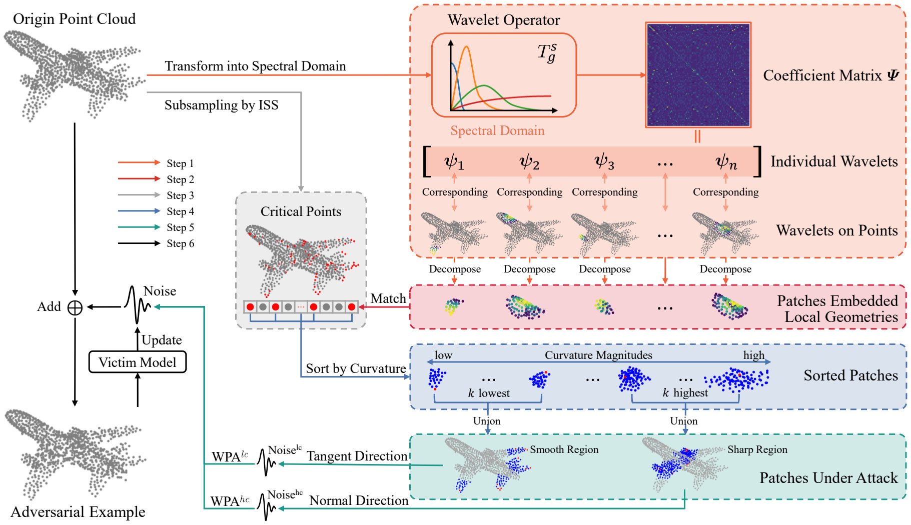


Fig. 2: Pipeline of our proposed WPA attack. To capture the local geometric structures of 3D objects, we first utilize the wavelet operator to transform the point cloud into the spectral domain with different filters for obtaining detailed and distinct local representations. Then, we apply the ISS subsampling algorithm on the 3D shape with previously obtained local wavelet coefficients to locate critical points and decompose object into curvature-consistent patches. Based on these patches, we develop two-type attack methods WPA^{hc} and WPA^{lc} to selectively perturb the points on the imperceptible regions (i.e., characteristics of smoothness and sharpness) of low visual sensitivity.

3 Methodology

3.1 Overview

Problem formulation. Point cloud data represents the collection of surface points sampled from a target object or scene within the selected three-dimensional coordinate system. In general, a point cloud $\mathcal{P}$ comprises an unordered set of points $\{\boldsymbol{p}_i\}_{i=1}^n$, where $\boldsymbol{p}_i \in \mathbb{R}^3$ signifies a coordinate vector, n denotes the number of points encompassed by the point cloud. Since our paper primarily focuses on the task of 3D point cloud classification, we denote the 3D classification model as $f(\cdot) : \mathbb{R}^{n\times 3} \rightarrow \mathbb{R}^C$. For each input $\mathcal{P}$, the classifier's objective is to yield a correct prediction $y = F(\mathcal{P}) = argmax_{i\in[C]} f(\mathcal{P})_i \in Y$. $Y = \{1, 2, 3, \ldots, C\}$ represents the authentic class of the point cloud, and C denotes the number of classes. Generally, to achieve adversarial attacks on point cloud classification models, the attackers aim to add adversarial noise $\boldsymbol{\Delta} \in \mathbb{R}^{n\times 3}$ to the original point cloud $\mathcal{P}$, so that the well-trained classifier is misled to make wrong predictions for the adversarial point cloud $\mathcal{P}' = \mathcal{P} + \boldsymbol{\Delta}$.

Attack pipeline. To generate high-quality adversarial samples, we propose a novel Wavelet Patches Attack (WPA), which aims to hide adversarial noise in specific regions of a point cloud while preserving corresponding local geometric characteristics. As shown in Fig. 2, WPA first transforms the point cloud into the spectral domain using wavelet filters to generate a wavelet coefficients matrix that represents the geometric dependency across local points. Then, to generate curvature-aware patches according to the coefficients matrix, we employ an Intrinsic Shape Signature (ISS) strategy to uniformly sample critical points, and utilize them to meticulously carve out wavelet patches with local characteristics of wavelets at each point. Combining these feature points with their corresponding patches, WPA proceeds to rank them based on curvature magnitude. To make the attack more natural and imperceptible, we only select two specific patches with sensitive geometric curvatures (lowest curvature of smoothness and highest curvature of sharpness) for perturbations. Two distinct attack variants, i.e., WPA^{hc} and WPA^{lc}, are further devised to attack these selected patches, respectively. In this manner, our approach is able to perceive and preserve distinct local geometries for generating high-quality adversarial samples.

3.2 Exploring Local Geometric Characteristics of Point Clouds through Wavelets Processing

Most existing 3D attack methods are developed in the data domain, which perturb points through implicit global distance constraints. Although some methods try to add noise in the spectral domain, all of these methods globally perturb the whole points of 3D objects, failing to perceive and preserve the distinctive types of point-to-point dependencies in local regions (such as smoothness and sharpness). To perceive local geometric characteristics for better preserving the topology of 3D objects, we endeavor to utilize the wavelet operator, the pronounced locality of which enables explicit preservation of local geometric characteristics across various parts of the point cloud, while retaining interaction information between points in the spectral domain.

Preliminary of spectral transformation. To transform the unordered point cloud into spectral domain, point cloud $\mathcal{P}$ is commonly represented by a symmetric (*i.e.* undirected) graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}, \boldsymbol{A}\}$ containing $n = |\mathcal{V}|$ nodes. $\mathcal{V}$ represents the node set, $\mathcal{E}$ denotes the edge set, and $\boldsymbol{A}$ signifies the weighted adjacency matrix of the graph. To capture local correspondence within the point cloud, we construct an undirected K-nearest-neighbor graph (K-NN graph), subsequently completing the $\mathcal{V}$ of $\mathcal{G}$. The graph weights $w_{i,j} = w_{j,i} \in \boldsymbol{A}$ are then determined by calculating the Euclidean distance between points $\boldsymbol{p}_i$ and $\boldsymbol{p}_j$, *i.e.*, $w_{i,j} = w_{j,i} = ||\boldsymbol{p}_j - \boldsymbol{p}_j||_2$. Moreover, each node $v_i \in \mathcal{V}$ is assigned a graph signal $\boldsymbol{h}_i$. Different from previous work [14], we treat the graph signal $\boldsymbol{h}_i$ as $\boldsymbol{\delta}_i$, where $\boldsymbol{\delta}_i \in \mathbb{R}^n$ is a one-hot vector with 1 on v_i and zeros elsewhere.

In general Graph Fourier Transformation (GFT), the combinatorial graph Laplacian operator is defined as $\boldsymbol{L} := \boldsymbol{D} - \boldsymbol{A}$, where $\boldsymbol{D}$ represents the degree matrix, with each element $d_{i,j} = \sum_{j=1}^{n} w_{i,j}$. As the constructed $\mathcal{G}$ has non-negative real values for $w_{i,j}$, the resulting matrix $\boldsymbol{L}$ is real, symmetric, and positive semi-definite. It admits an eigenvalue-decomposition $\boldsymbol{L} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^T$, where $\boldsymbol{U} = [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n]$ comprises orthonormal eigenvectors $\boldsymbol{u}_i$, and $\boldsymbol{\Lambda} = diag(\lambda_1, \ldots, \lambda_n)$

consists of eigenvalues $\{\lambda_1 = 0 \leq \lambda_2 \leq \ldots \leq \lambda_n\}$. In the realm of graph signal processing, the aforementioned eigenvalues λ are referred to as the graph spectrum.

Transform point clouds through wavelet operator. Diverging from global transformations like the aforementioned GFT, the Spectral Graph Wavelet Transform [12] was the first to combine a spectral design with spatial domain localization. This wavelet transform presents a sparser encoding for graph signals, with the wavelet kernel acting within its local region, equivalent to the neighborhood of the central node. Therefore, wavelet transform demonstrates higher encoding efficiency, particularly in undirected graphs formed by point clouds, enabling a more precise capture of local characteristics and geometric structures for each point. To perceive and preserve the local geometries, we introduce to utilize wavelet transform for spectral transformation in this section.

Specifically, the wavelet kernel is constructed by designing a real-valued function $g(\cdot) : \mathbb{R}^n \to \mathbb{R}^n$. In the spectral domain of $\boldsymbol{L}$, $g(\cdot)$ manifests as a band-pass filter, akin to the Fourier transform of the "mother wavelet" for the continuous wavelet transform. Therefore, we can define the wavelet operator/transformation as:

$$T_g^s = g_s(\boldsymbol{L}) = \boldsymbol{U}\boldsymbol{\Lambda}(g_s(\lambda))\boldsymbol{U}^T, \tag{1}$$

where s denotes the scale of wavelet operator. By applying the T_g^s to $\boldsymbol{h}_i$ of each node, we can obtain corresponding individual spectral graph wavelet coefficients centered at node i:

$$\boldsymbol{\psi}_{s,i} = T_g^s \boldsymbol{h}_i = \boldsymbol{U}\boldsymbol{\Lambda}(g_s(\lambda))\boldsymbol{U}^T \boldsymbol{h}_i. \tag{2}$$

It is worth noting that due to the band-pass filter nature of the wavelet kernel function $g(\cdot)$, $g(0) = 0$, and $\lim\limits_{\lambda \to 0} g(\lambda) \to 0$, wavelet transform typically introduces a set of scaling functions to serve as a low-pass filter during signal decomposition. To enhance the representation of low-frequency signals, the same goes for our work. Thereby, $T_g^s \boldsymbol{h}$ gives the wavelet coefficients $\boldsymbol{\psi}_{s,i}$ for the graph signal $\boldsymbol{h}$ at scale s.

Analysis on wavelet outputs. By applying the wavelet transformation to the undirected K-NN graph $\mathcal{G}$, we can obtain the wavelet coefficients matrix as:

$$\boldsymbol{\Psi} = \phi_{WT}(\mathcal{P}, s) = [\boldsymbol{\psi}_{s,1}, \ldots, \boldsymbol{\psi}_{s,n}], \tag{3}$$

where ϕ_{WT} denotes the wavelet transformation. From the perspective of graph signal processing, the j-th coefficient in the vector $\boldsymbol{\psi}_{s,i} \in \mathbb{R}^n$ represents the contribution of node j concerning the central node i or, in other words, the energy diffused from the node i to node j. In terms of the practical significance of point cloud data, the wavelet $\boldsymbol{\psi}_{s,i}$ for each point characterizes the local geometric structures and semantic context within the local points, centered around node i. By comparing the coefficient of each individual wavelet with ϵ, we can utilize a threshold ϵ to define geometry-sensitive patch $\boldsymbol{P}_i \subseteq \mathcal{P}$ at $\boldsymbol{p}_i$. We decompose these patches from the whole point cloud via the following formulation:

$$\text{Decompose}(\mathcal{P}, \boldsymbol{\Psi}; i, \epsilon) = \boldsymbol{P}_i = \{\boldsymbol{p}_j\}_i, \quad s.t. \quad |\boldsymbol{\psi}_{s,i}(j)| > \epsilon, \quad \boldsymbol{p}_i, \boldsymbol{p}_j \in \mathcal{P}. \tag{4}$$

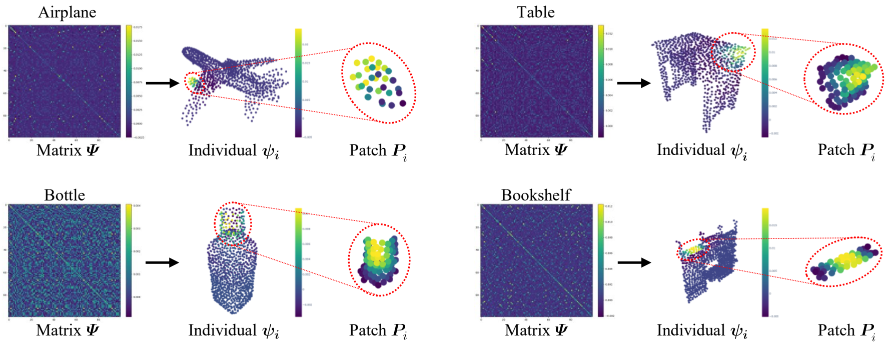


Fig. 3: Visualization on matrix $\boldsymbol{\Psi}$ generating by the wavelet operator, individual wavelet $\boldsymbol{\psi}_i$ at point $\boldsymbol{p}_i$, and its attended patch $\boldsymbol{P}_i$. For clearly reading, we only demonstrate a part of $\boldsymbol{\Psi}$.

Based on this, we can obtain n number of patches $\{\boldsymbol{P}_i\}_{i=1}^{n}$ for each point in $\mathcal{P}$. For example, as illustrated in Fig. 3, we acquire the wavelet coefficient matrix $\boldsymbol{\Psi}$ for each object through the wavelet transformation. This matrix's sparsely highlighted sections correspond to the wavelet's localized effects. Consequently, from the individual wavelets $\boldsymbol{\psi}_i$ in the matrix (center of the subplot), we can obtain the attended patch $\boldsymbol{P}_i$ acquired from Eq. (4), with each patch encompassing the geometric context of the local region.

3.3 Crafting Imperceptible Perturbations in Curvature-aware Patches

By employing the above spectral methods, we are able to perceive point-based regions corresponding to different local geometries. To scrutinize the variances among different patches and their significance in point cloud classification tasks, we analyze the patches by calculating their curvature and selectively add noise only on the geometric-sensitive patches. To minimize the number of perturbed points and imperceptibly hide noise, we devise distinct attack methodologies for selected patches with varying curvatures.

Select specific patches for decomposing point clouds. Using all patches constructed on each point is complicated since there are significant overlaps between adjacent patches. Therefore, we propose to employ a subsampling algorithm to to decompose the whole point cloud into different local patches, i.e., it suffices to select a subset from $\{\boldsymbol{P}_i\}_{i=1}^{n}$, one that can approximate the entire original point cloud.

To be specific, we utilize the Intrinsic Shape Signature (ISS) algorithm, which is a traditional point cloud processing method calculating the curvature of the point cloud and the normal vectors. It employs curvature and normal vector to construct an intrinsic shape signature, capturing local shape details. The analysis of the ISS yields a set of feature points $\mathcal{P}^{ISS} \subset \mathcal{P}$. By leveraging the wavelet outputs and ISS, the point cloud can be decomposed into distinct patches:

$$\{\mathrm{Decompose}(\mathcal{P}, \boldsymbol{\Psi}; i, \epsilon)\}_{i=1}^{m} = \{\boldsymbol{P}_i\}^{ISS}, \quad s.t. \quad \boldsymbol{p}_i \in \mathcal{P}^{ISS} = \phi_{ISS}(\mathcal{P}), \tag{5}$$

where m is the number of points in $\mathcal{P}^{ISS}$, ϕ_{ISS} denotes the ISS subsampling algorithm and each single patch $\boldsymbol{P}_i$ is centering at $\boldsymbol{p}_i$.

Add curvature-aware noise to specific patches. In order to hide adversarial perturbations within specific patches more covertly, we conduct curvature calculations and subsequent sorting for each wavelet patch in $\{\boldsymbol{P}_i\}^{ISS}$. For the highest or lowest curvatures, the adversarial perturbation is applied along the direction of the normal vectors or tangent plane at the patch points, respectively.

In accordance with the chosen patch curvature magnitudes, two types of attacks are constructed as WPA^{hc} and WPA^{lc}. For patch curvatures characterized by highest magnitudes denoted as $\boldsymbol{P}^{high}$ and lowest magnitudes denoted as $\boldsymbol{P}^{low}$, we introduce perturbations $\boldsymbol{\Delta}_{high}, \boldsymbol{\Delta}_{low} \in \mathbb{R}^{n\times 1}$, respectively. These perturbations are added to the normal vectors and tangent plane directions of the patch points:

$$\boldsymbol{P}^{high} = \boldsymbol{P}_{m-k+1} \cup \boldsymbol{P}_{m-k+2} \cup \cdots \cup \boldsymbol{P}_m, \tag{6}$$

$$\boldsymbol{P}^{low} = \boldsymbol{P}_1 \cup \boldsymbol{P}_2 \cup \cdots \cup \boldsymbol{P}_k, \tag{7}$$

where m is the number of patches contained in $\{\boldsymbol{P}_i\}^{ISS}$. $\boldsymbol{P}^{high}$ and $\boldsymbol{P}^{low}$ respectively represent the sets of patches with the k highest or lowest curvatures. To selectively perturb points, diagonal mask matrices are constructed as:

$$\boldsymbol{M}_{high} = diag(r_i^{high}) \in \mathbb{R}^{n\times n} \quad s.t. \quad r_i^{high} = \begin{cases} 0, \boldsymbol{p}_i \notin \boldsymbol{P}^{high} \\ 1, \boldsymbol{p}_i \in \boldsymbol{P}^{high} \end{cases}, \tag{8}$$

$$\boldsymbol{M}_{low} = diag(r_i^{low}) \in \mathbb{R}^{n\times n} \quad s.t. \quad r_i^{low} = \begin{cases} 0, \boldsymbol{p}_i \notin \boldsymbol{P}^{low} \\ 1, \boldsymbol{p}_i \in \boldsymbol{P}^{low} \end{cases}, \tag{9}$$

which ensures the perturbations are zero for non-selected points. The final adversarial perturbation $\boldsymbol{\Delta}$ can be obtained through the following:

$$\boldsymbol{\Delta} = \begin{cases} \boldsymbol{\Delta}_{high} \cdot \boldsymbol{M}_{high}\boldsymbol{N}_{normal}, \; for \text{ WPA}^{hc} \\ \boldsymbol{\Delta}_{low} \cdot \boldsymbol{M}_{low}\boldsymbol{N}_{tangent}, \quad for \text{ WPA}^{lc} \end{cases}, \tag{10}$$

where $\boldsymbol{N}_{normal} = [\boldsymbol{n}_1, \cdots, \boldsymbol{n}_n]^T$ and $\boldsymbol{N}_{tangent} = [\boldsymbol{t}_1, \cdots, \boldsymbol{t}_n]^T$ are matrices composed of the normal vectors $\boldsymbol{n} \in \mathbb{R}^3$ or tangent vectors $\boldsymbol{t} \in \mathbb{R}^3$ for each point in $\mathcal{P}$. $\boldsymbol{n}$ and $\boldsymbol{t}$ can be calculated for point $\boldsymbol{p}_i$ using traditional methods introduced in [13]. In particular, the 3×3 positive semi-definite covariance matrix $\boldsymbol{C}$ is constructed as:

$$\boldsymbol{C} = \sum_{i\neq j} (\boldsymbol{p}_j - \boldsymbol{p}_i) \otimes (\boldsymbol{p}_i - \boldsymbol{p}_j), \quad s.t. \quad \boldsymbol{p}_i, \boldsymbol{p}_j \in \boldsymbol{P}^{high} \quad or \quad \boldsymbol{p}_i, \boldsymbol{p}_j \in \boldsymbol{P}^{low}, \tag{11}$$

where $\otimes$ denotes the operator of outer product. By performing eigenvalue decomposition on $\boldsymbol{C}$, $\boldsymbol{n}_i$ will be obtained as the eigenvector corresponding to the smallest eigenvalue, and the vector resulting from the addition of the other two eigenvectors as $\boldsymbol{t}_i$.

3.4 Generating Adversarial Examples.

The task of generating the adversarial sample $\mathcal{P}'$ can be formulated as the following optimization problem:

$$\min_{\boldsymbol{\Delta}} \mathcal{L}_{mis}(f(\mathcal{P}'), y) + \alpha \cdot \mathcal{L}_{reg}(\mathcal{P}', \mathcal{P}), \quad s.t. \quad \mathcal{P}' = \mathcal{P} + \boldsymbol{\Delta}, \tag{12}$$

where $\mathcal{L}_{mis}$ is the cross-entropy loss used to enhance the misclassification of the model, $\mathcal{L}_{reg}$ is the regularizer penalizing the difference between $\mathcal{P}'$ and $\mathcal{P}$, and α is the penalty parameter. Here, we employ Chamfer distance(CD) [5] and Hausdorff distance(HD) [17] as the regularizer loss $\mathcal{L}_{reg}(\mathcal{P}', \mathcal{P})$. The final formulation for adversarial sample generation can be represented by:

$$\min_{\boldsymbol{\Delta}} \mathcal{L}_{mis}(f(\mathcal{P}'), y) + \alpha \cdot (\beta_1 \cdot \mathcal{L}_{CD}(\mathcal{P}', \mathcal{P}) + \beta_2 \cdot \mathcal{L}_{HD}(\mathcal{P}', \mathcal{P})), \tag{13}$$

where

$$\mathcal{P}' = \mathcal{P} + \boldsymbol{\Delta}, \quad s.t. \quad \boldsymbol{\Delta} = \begin{cases} \boldsymbol{\Delta}_{high} \cdot \boldsymbol{M}_{high} \boldsymbol{N}_{normal}, \ for\ \mathrm{WPA}^{hc} \\ \boldsymbol{\Delta}_{low} \cdot \boldsymbol{M}_{low} \boldsymbol{N}_{tangent}, \quad for\ \mathrm{WPA}^{lc} \end{cases}. \tag{14}$$

4 Experiments

4.1 Attack Setup

Datasets and baselines. We employ ModelNet40 [44] and ShapeNetPart [48] to evaluate the performance of various adversarial attack methods on point clouds. Specifically, ModelNet40 consists of 9,843 CAD models designated for training and 2,468 for testing. ShapeNetPart consists of 12,137 shapes allocated for training, with 2,874 set aside for testing. Notably, we exclusively selected objects from the test set that were accurately classified by benign classifiers to generate adversarial samples, which are ensured to uniformly sample to 1,024 points. Our approach is comprehensively evaluated on the four most popular models, *i.e.*, PointNet [31], DGCNN [40], PointConv [43], and CurveNet [47].

Implementation details. The B-spline wavelet kernel was applied as T_g in Eq. (1). We use Adam optimizer to optimize the objective in Eq. (13). A steadfast training regimen of 200 iterations is implemented, wherein the learning rate and momentum are established at 0.01 and 0.9, respectively. The penalty coefficients $\beta_1 = 1$ and $\beta_2 = 0.1$ are adopted as the default values in Eq. (13). The initial penalty coefficient α in Eq. (12) is set to 10, subject to automatic adjustment via a binary search conducted over 20 iterations, which follows [1]. In Eq. (4), we set the threshold ϵ to 10^{-5}, and the proportion of points under attack that determined by k in Eqs. (6) and (7) is established at 0.5. We select $K = 10$ for the building a K-NN Graph and $s = 2$ as the scale in Eq. (3) for the wavelet operator. All experiments are conducted on a single NVIDIA Quadro RTX 5000 GPU.

Table 1: Quantitative comparison on the perturbation size generated by different attack methods on ModelNet40 and ShapeNetPart dataset. The bold numbers denote the most imperceptible attacks, and the underscored numbers denote the second-best.

Dataset	Method	PointNet			DGCNN			PointConv			CurveNet		
		ASR↑	CD↓	HD↓	ASR↑	CD↓	HD↓	ASR↑	CD↓	HD↓	ASR↑	CD↓	HD↓
ModelNet40	PGD [26]	100%	0.0018	0.0302	100%	0.0019	0.0202	100%	0.0019	0.0129	100%	0.0015	0.0286
	AdvPC [11]	100%	0.0013	0.0346	100%	0.0012	0.0186	100%	0.0010	0.0127	100%	0.0018	0.0278
	SI-Adv [16]	100%	**0.0002**	0.0205	100%	**0.0004**	0.0054	100%	**0.0003**	0.0116	100%	**0.0006**	0.0199
	GeoA [41]	100%	0.0064	0.0175	100%	0.0176	0.0402	99%	0.0014	**0.0062**	100%	0.0012	0.0067
	GSDA [14]	100%	0.0007	0.0031	100%	0.0104	0.1401	100%	0.0017	0.0218	100%	0.0040	0.0142
	ITA [19]	100%	0.0037	0.0052	100%	0.0058	0.0066	-	-	-	-	-	-
	GSDA++ [21]	100%	0.0006	0.0028	100%	0.0072	0.0135	-	-	-	-	-	-
	WPA^{hc}	100%	0.0004	**0.0020**	100%	0.0008	0.0069	**100%**	0.0012	0.0075	100%	0.0007	0.0057
	WPA^{lc}	100%	0.0004	**0.0020**	100%	0.0006	**0.0043**	**100%**	0.0010	**0.0062**	100%	**0.0006**	**0.0047**
ShapeNetPart	PGD [26]	100%	0.0017	0.0434	100%	0.0019	0.0628	100%	0.0018	0.0442	98%	0.0016	0.0377
	AdvPC [11]	100%	0.0019	0.0543	100%	0.0029	0.0649	100%	0.0015	0.0404	69%	0.0016	0.0374
	SI-Adv [16]	96%	0.0010	0.0433	95%	**0.0009**	0.0418	95%	**0.0008**	**0.0125**	91%	**0.0008**	0.0358
	GeoA [41]	100%	0.0013	0.0358	100%	0.0025	0.0272	99%	0.0026	0.0251	100%	0.0022	0.0457
	GSDA [14]	95%	0.0023	**0.0257**	98%	0.0035	0.0388	94%	0.0026	0.0240	100%	0.0526	0.1809
	WPA^{hc}	**100%**	**0.0006**	0.0301	**100%**	0.0019	0.0306	**100%**	0.0025	0.0337	**100%**	0.0020	0.0445
	WPA^{lc}	**100%**	**0.0006**	0.0299	**100%**	0.0018	**0.0250**	**100%**	0.0019	0.0234	**100%**	0.0019	**0.0353**

4.2 Attack Performance

Quantitative comparison. To fairly evaluate the efficacy of our proposed WPA, we conducted a comparative analysis against five other methods, namely PGD [26], AdvPC [11], SI-Adv [16], GeoA [41], GSDA [14], ITA [19], GSDA++ [21]. The results of this comparison are shown in Tab. 1. It is observed that our WPA method achieves a 100% attack success rate(ASR) across four different victim 3D models, while generating adversarial samples with almost the lowest perturbation size according to evaluation metrics. This substantiates the effectiveness of WPA, demonstrating its capability to hide imperceptible noise within curvature-aware patches, thereby reducing the number of perturbed points to achieve comparable, if not superior, adversarial attack impact.

Visualization results. We present the visualization results of our proposed WPA alongside those of GeoA [41] and GSDA [14] in Fig. 4. It can be observed that our adversarial examples are more imperceptible than other attacks. This is attributed to the fact that, unlike GeoA and GSDA which employ global constraints and transformations, our approach selectively perturbs geometric-sensitive patches and introduces curvature-wise noise. Consequently, we achieve remarkably effective attacks by only perturbing a subset of points. Additionally, the right side of Fig. 4 also presents the visualization of the wavelet coefficients of the original point clouds.

Evaluation on Robustness. To demonstrate the efficacy of attacks on well-defended 3D models, we conducted experiments with various defense methods, including Statistical Outlier Removal(SOR) [55], Simple Random Sampling(SRS) [49], Dup-Net [55], and IF-Defense [45]. As illustrated in Tab. 2, the success rate of PGD is relatively low across all defenses because it often results in an uneven local distribution and outliers. Conversely, SI-Adv and GeoA achieved higher success rates as they global constrain the geometric shape of

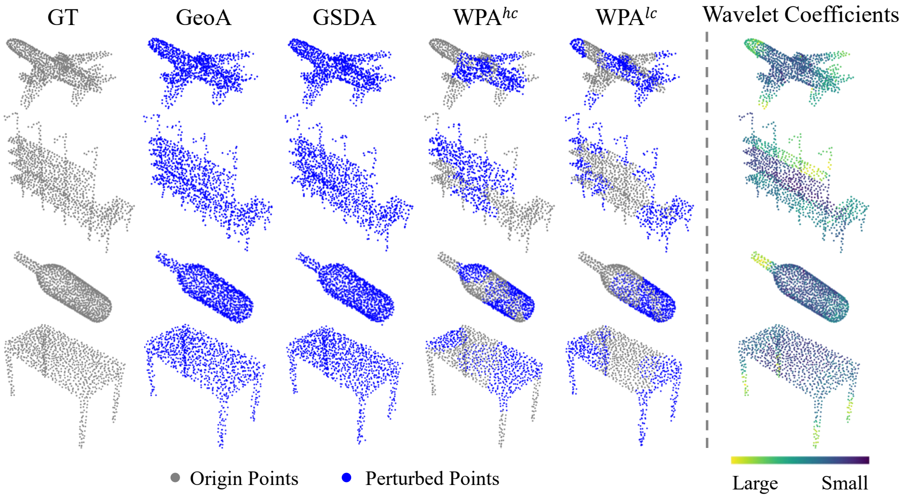


Fig. 4: Visualization of the adversarial examples and wavelet coefficients.

Table 2: ASR of different attacks on PointNet on ModelNet40 equipped with various defenses. The bold numbers denote the highest rate, and the underscored numbers denote the second-highest.

Defense	PGD	SI-Adv	GeoA	GSDA	**WPA^{hc}**	**WPA^{lc}**
No Defense	100.0%	99.8%	100.0%	100.0%	**100.0%**	**100.0%**
SOR [55]	52.3%	97.4%	62.5%	81.0%	**98.9%**	96.2%
SRS [49]	51.4%	**85.6%**	67.6%	81.0%	80.4%	83.9%
DUP-Net [55]	49.5%	95.8%	59.2%	68.9%	**98.1%**	95.5%
IF-Defense [45]	37.1%	61.2%	38.7%	50.1%	**72.3%**	68.2%

the object's surface with geometric-aware constraints, thereby reducing outliers. GSDA introduces perturbations in the spectral domain, corresponding to smaller noise in the data domain, and likewise achieves commendable attack outcomes. In comparison, our WPA secured the highest attack success rates against nearly all defensive methods, indicating that the noise we introduce aligns with the geometric characteristics of the object's local regions, achieving effective results through perturbation in a fewer patches.

4.3 Ablations

Effectiveness of various components in WPA pipeline. We firstly conduct ablation experiments on the components within our WPA pipeline, applying various attacks on point clouds transformed by the wavelet operator. Tab. 3 demonstrates the experiment results. In the attacks without using the ISS subsampling algorithm, we randomly select points equivalent in number to the ISS critical points as substitutes. In the attacks that do not utilize "HC" or "LC",

Table 3: Effectiveness of various components in WPA pipeline. The bold numbers denote the most imperceptible attacks, and the underscored numbers denote the second-best. "ISS?" denotes wheather we apply the ISS subsampling algorithm. "HC?"/"LC?" denote we perturb the patches with the k highest/lowest curvature and project the noise to normal/tangent direction. Victim model: PointNet.

ISS?	HC?	LC?	ModelNet40			ShapeNetPart		
			ASR	CD	HD	ASR	CD	HD
×	×	×	100%	0.00125	0.00695	87%	**0.00054**	0.03237
✓	×	×	100%	0.00064	0.00437	99%	0.00059	0.03777
✓	✓	×	100%	**0.00044**	**0.00197**	**100%**	0.00058	0.03008
✓	×	✓	100%	0.00045	0.00204	**100%**	0.00059	**0.02986**

Table 4: Investigation on different types of wavelet kernel. Victim model: PointNet.

Wavelet Kernel	Method	ModelNet40			ShapeNetPart		
		ASR	CD	HD	ASR	CD	HD
Meyer	WPA^{hc}	100%	0.0005	**0.0020**	100%	**0.0005**	0.0312
	WPA^{lc}	100%	0.0005	0.0022	100%	**0.0005**	**0.0293**
Mexican Hat	WPA^{hc}	100%	**0.0004**	0.0021	100%	0.0007	0.0434
	WPA^{lc}	100%	**0.0004**	0.0023	100%	0.0006	0.0334
B-spline	WPA^{hc}	100%	**0.0004**	**0.0020**	100%	0.0006	0.0301
	WPA^{lc}	100%	**0.0004**	**0.0020**	100%	0.0006	0.0299

we randomly select a same number of patches to perturb without changing the generated noise's direction. The results indicate that on the ModelNet40 dataset, the adversarial samples crafted using ISS exhibit a lower perturbation size. This is attributable to the fact that the local regions where the critical points obtained from ISS subsampling play a more pronounced role in the feature of the whole shape. Conversely, on the ShapeNetPart dataset, attacks applying ISS achieved a higher ASR, but requiring a larger perturbation budget. On the other hand, attack strategies that targeted specific curvatures yielded superior experimental outcomes. This is attributable to the strategic concealment of noise within the local geometric structures, thereby achieving enhanced imperceptibility.

Investigation on different types of wavelet kernel. To investigate the impact of different wavelet kernels within the wavelet operator on attack performance, we conducted experiments utilizing a variety of wavelet kernels. As shown in Tab. 4, our WPA remains relatively insensitive to the choice of wavelet kernel, provided all other experimental conditions are kept constant. This insensitivity arises because, although various wavelet kernels transform point cloud data into the spectral domain in distinct manners, the spectral information post-transformation encompasses the intrinsic geometric structural features of the point cloud. Consequently, WPA yields comparable perturbation sizes in the data domain for patches corresponding to each point. Based on these findings, we select the B-spline wavelet kernel, which demonstrated the most favorable experimental outcomes, as the default setting for all subsequent experiments.

Investigation on the proportion of perturbed points determined by k. As illustrated in Fig. 5, we conducted evaluations on WPA^{hc} and WPA^{lc} under

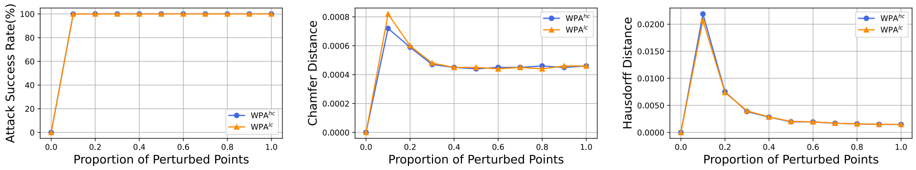


Fig. 5: WPA^{hc} and WPA^{lc} attack performance when setting different proportion of perturbed points determined by k, in terms of ASR, CD and HD results between adversarial examples and the origins.

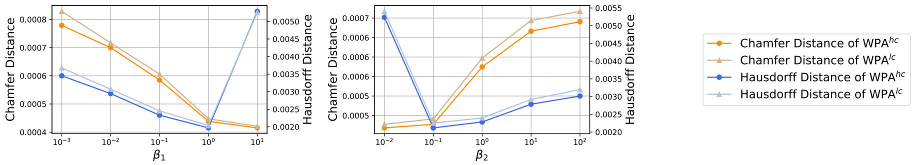


Fig. 6: WPA^{hc} and WPA^{lc} attack performance when applying varied β_1,β_2. All Attacks achieved 100% success rate. Experiments are conducted on ModelNet40 using PointNet as the model of classifier.

various k settings in Eqs. (6) and (7). By adjusting the k for patches with k highest/lowest curvature, WPA^{hc}/WPA^{lc} introduces imperceptible noise to different proportion of points. The results indicate that the adversarial attacks are successful across all values, demonstrating our method's efficacy in hiding noise within the point cloud's geometric structure. When the proportion of perturbed points is set to 0.5, the adversarial samples exhibit superior and stable perturbation sizes, achieving commendable attack performance with only partial points perturbed.

Investigation on penalty parameters β_1 and β_2. To investigate the impact of the regularization constraints β_1 and β_2 in Eq. (13), we conducted a study wherein one value was held constant at its default setting while another was varied. The experimental findings revealed that varying β_1 and β_2 yielded a consistent attack success rate of 100%. Moreover, as demonstrated in Fig. 6, the default settings of $\beta_1 = 1$ and $\beta_2 = 0.1$ delivered the most optimal results in terms of perturbation size.

5 Conclusion

In this paper, we propose a novel Wavelet Patches Attack(WPA), which leverages local spectral attributes to identify curvature-aware patches for hiding imperceptible perturbations aligned with their local geometric characteristics. By utilizing the wavelet transform, we are able to capture local regions embedded with the local geometric context, subsequently decomposing the point cloud into geometric-sensitive patches. Then, we introduced imperceptible noise to different patches based on curvature variations, thereby generating adversarial examples with minimized perturbation size by perturbing only a subset of points. Experiments validate both effectiveness and robustness of our WPA.

References

1. Carlini, N., Wagner, D.: Towards evaluating the robustness of neural networks. In: 2017 IEEE Symposium on Security and Privacy (SP). pp. 39–57 (2017)
2. Chen, S., Tian, D., Feng, C., Vetro, A., Kovačević, J.: Fast resampling of three-dimensional point clouds via graphs. IEEE Transactions on Signal Processing **66**(3), 666–681 (2017)
3. Chen, X., Ma, H., Wan, J., Li, B., Xia, T.: Multi-view 3d object detection network for autonomous driving. In: Proceedings of the IEEE conference on Computer Vision and Pattern Recognition (CVPR). pp. 1907–1915 (2017)
4. Dong, Y., Liao, F., Pang, T., Su, H., Zhu, J., Hu, X., Li, J.: Boosting adversarial attacks with momentum. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 9185–9193 (2018)
5. Fan, H., Su, H., Guibas, L.J.: A point set generation network for 3d object reconstruction from a single image. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 605–613 (2017)
6. Fang, X., Liu, D., Fang, W., Zhou, P., Xu, Z., Xu, W., Chen, J., Li, R.: Fewer steps, better performance: Efficient cross-modal clip trimming for video moment retrieval using language. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 38, pp. 1735–1743 (2024)
7. Fang, X., Liu, D., Zhou, P., Hu, Y.: Multi-modal cross-domain alignment network for video moment retrieval. IEEE Transactions on Multimedia **25**, 7517–7532 (2022)
8. Fang, X., Liu, D., Zhou, P., Nan, G.: You can ground earlier than see: An effective and efficient pipeline for temporal sentence grounding in compressed videos. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 2448–2460 (2023)
9. Fang, X., Xiong, Z., Fang, W., Qu, X., Chen, C., Dong, J., Tang, K., Zhou, P., Cheng, Y., Liu, D.: Rethinking weakly-supervised video temporal grounding from a game perspective. In: European Conference on Computer Vision. Springer (2024)
10. Goodfellow, I.J., Shlens, J., Szegedy, C.: Explaining and harnessing adversarial examples. arXiv preprint arXiv:1412.6572 (2014)
11. Hamdi, A., Rojas, S., Thabet, A., Ghanem, B.: Advpc: Transferable adversarial perturbations on 3d point clouds. In: European Conference on Computer Vision (ECCV). pp. 241–257 (2020)
12. Hammond, D.K., Vandergheynst, P., Gribonval, R.: Wavelets on graphs via spectral graph theory. Appl. Comput. Harmonic Anal. **30**(2), 129–150 (2011)
13. Hoppe, H., DeRose, T., Duchamp, T., McDonald, J., Stuetzle, W.: Surface reconstruction from unorganized points. In: Proceedings of the 19th Annual Conference on Computer Graphics and Interactive Techniques. pp. 71–78 (1992)
14. Hu, Q., Liu, D., Hu, W.: Exploring the devil in graph spectral domain for 3d point cloud attacks. In: European Conference on Computer Vision (ECCV) (2022)
15. Hu, W., Pang, J., Liu, X., Tian, D., Lin, C.W., Vetro, A.: Graph signal processing for geometric data and beyond: Theory and applications. IEEE Transactions on Multimedia **24**, 3961–3977 (2021)
16. 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)
17. Huttenlocher, D.P., Klanderman, G.A., Rucklidge, W.J.: Comparing images using the hausdorff distance. IEEE Transactions on Pattern Analysis and Machine Intelligence **15**(9), 850–863 (1993)

18. Kurakin, A., Goodfellow, I., Bengio, S.: Adversarial machine learning at scale. arXiv preprint arXiv:1611.01236 (2016)
19. Liu, D., Hu, W.: Imperceptible transfer attack and defense on 3d point cloud classification. IEEE transactions on pattern analysis and machine intelligence **45**(4), 4727–4746 (2022)
20. Liu, D., Hu, W.: Explicitly perceiving and preserving the local geometric structures for 3d point cloud attack. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 38, pp. 3576–3584 (2024)
21. Liu, D., Hu, W., Li, X.: Point cloud attacks in graph spectral domain: When 3d geometry meets graph signal processing. IEEE Transactions on Pattern Analysis and Machine Intelligence (2023)
22. Liu, D., Hu, W., Li, X.: Robust geometry-dependent attack for 3d point clouds. IEEE Transactions on Multimedia (2023)
23. Liu, D., Yu, R., Su, H.: Extending adversarial attacks and defenses to deep 3d point cloud classifiers. In: 2019 IEEE International Conference on Image Processing (ICIP). pp. 2279–2283 (2019)
24. Liu, M., Li, X., Ling, Z., Li, Y., Su, H.: Frame mining: a free lunch for learning robotic manipulation from 3d point clouds. In: Conference on Robot Learning. pp. 527–538. PMLR (2023)
25. Ma, C., Meng, W., Wu, B., Xu, S., Zhang, X.: Efficient joint gradient based attack against sor defense for 3d point cloud classification. In: Proceedings of the 28th ACM International Conference on Multimedia. pp. 1819–1827 (2020)
26. Madry, A., Makelov, A., Schmidt, L., Tsipras, D., Vladu, A.: Towards deep learning models resistant to adversarial attacks. arXiv preprint arXiv:1706.06083 (2017)
27. Moosavi-Dezfooli, S.M., Fawzi, A., Fawzi, O., Frossard, P.: Universal adversarial perturbations. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 1765–1773 (2017)
28. Moosavi-Dezfooli, S.M., Fawzi, A., Frossard, P.: Deepfool: a simple and accurate method to fool deep neural networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 2574–2582 (2016)
29. Mustafa, A., Khan, S.H., Hayat, M., Goecke, R., Shen, J., Shao, L.: Deeply supervised discriminative learning for adversarial defense. IEEE transactions on pattern analysis and machine intelligence **43**(9), 3154–3166 (2020)
30. Ortega, A., Frossard, P., Kovačević, J., Moura, J.M., Vandergheynst, P.: Graph signal processing: Overview, challenges, and applications. Proceedings of the IEEE **106**(5), 808–828 (2018)
31. 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 Conference on Computer Vision and Pattern Recognition (CVPR). pp. 652–660 (2017)
32. Ramasinghe, S., Khan, S., Barnes, N., Gould, S.: Spectral-gans for high-resolution 3d point-cloud generation. in 2020 ieee. In: RSJ International Conference on Intelligent Robots and Systems (IROS). pp. 8169–8176
33. Rosman, G., Dubrovina, A., Kimmel, R.: Patch-collaborative spectral point-cloud denoising. In: Computer Graphics Forum. vol. 32, pp. 1–12. Wiley Online Library (2013)
34. Singh, S.P., Wang, L., Gupta, S., Goli, H., Padmanabhan, P., Gulyás, B.: 3d deep learning on medical images: a review. Sensors **20**(18), 5097 (2020)
35. 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)

36. Tang, K., He, X., Peng, W., Wu, J., Shi, Y., Liu, D., Zhou, P., Wang, W., Tian, Z.: Manifold constraints for imperceptible adversarial attacks on point clouds. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 38, pp. 5127–5135 (2024)
37. Tao, Y., Liu, D., Zhou, P., Xie, Y., Du, W., Hu, W.: 3dhacker: Spectrum-based decision boundary generation for hard-label 3d point cloud attack. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 14340–14350 (2023)
38. 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)
39. Tu, C.C., Ting, P., Chen, P.Y., Liu, S., Zhang, H., Yi, J., Hsieh, C.J., Cheng, S.M.: Autozoom: Autoencoder-based zeroth order optimization method for attacking black-box neural networks. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 33, pp. 742–749 (2019)
40. 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)
41. Wen, Y., Lin, J., Chen, K., Chen, C.P., Jia, K.: Geometry-aware generation of adversarial point clouds. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI) (2020)
42. Wicker, M., Kwiatkowska, M.: Robustness of 3d deep learning in an adversarial setting. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 11767–11775 (2019)
43. Wu, W., Qi, Z., Fuxin, L.: Pointconv: Deep convolutional networks on 3d point clouds. In: Proceedings of the IEEE/CVF Conference on computer vision and pattern recognition. pp. 9621–9630 (2019)
44. 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 Conference on Computer Vision and Pattern Recognition (CVPR). pp. 1912–1920 (2015)
45. 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)
46. Xiang, C., Qi, C.R., Li, B.: Generating 3d adversarial point clouds. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 9136–9144 (2019)
47. Xiang, T., Zhang, C., Song, Y., Yu, J., Cai, W.: Walk in the cloud: Learning curves for point clouds shape analysis. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 915–924 (2021)
48. Yi, L., Kim, V.G., Ceylan, D., Shen, I.C., Yan, M., Su, H., Lu, C., Huang, Q., Sheffer, A., Guibas, L.: A scalable active framework for region annotation in 3d shape collections. ACM Transactions on Graphics (ToG) **35**(6), 1–12 (2016)
49. Zhang, Q., Yang, J., Fang, R., Ni, B., Liu, J., Tian, Q.: Adversarial attack and defense on point sets. arXiv preprint arXiv:1902.10899 (2019)
50. Zhang, S., Cui, S., Ding, Z.: Hypergraph spectral analysis and processing in 3d point cloud. IEEE Transactions on Image Processing **30**, 1193–1206 (2020)
51. Zhang, Y., Liang, G., Salem, T., Jacobs, N.: Defense-pointnet: Protecting pointnet against adversarial attacks. In: 2019 IEEE International Conference on Big Data (Big Data). pp. 5654–5660 (2019)

52. 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 Conference on Computer Vision and Pattern Recognition (CVPR). pp. 1201–1210 (2020)
53. Zheng, T., Chen, C., Yuan, J., Li, B., Ren, K.: Pointcloud saliency maps. In: Proceedings of the IEEE International Conference on Computer Vision (ICCV). pp. 1598–1606 (2019)
54. 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 Conference on Computer Vision and Pattern Recognition (CVPR). pp. 10356–10365 (2020)
55. 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 International Conference on Computer Vision (ICCV). pp. 1961–1970 (2019)