

## Task 2.3: Document analysis and recognition

## Task 1:

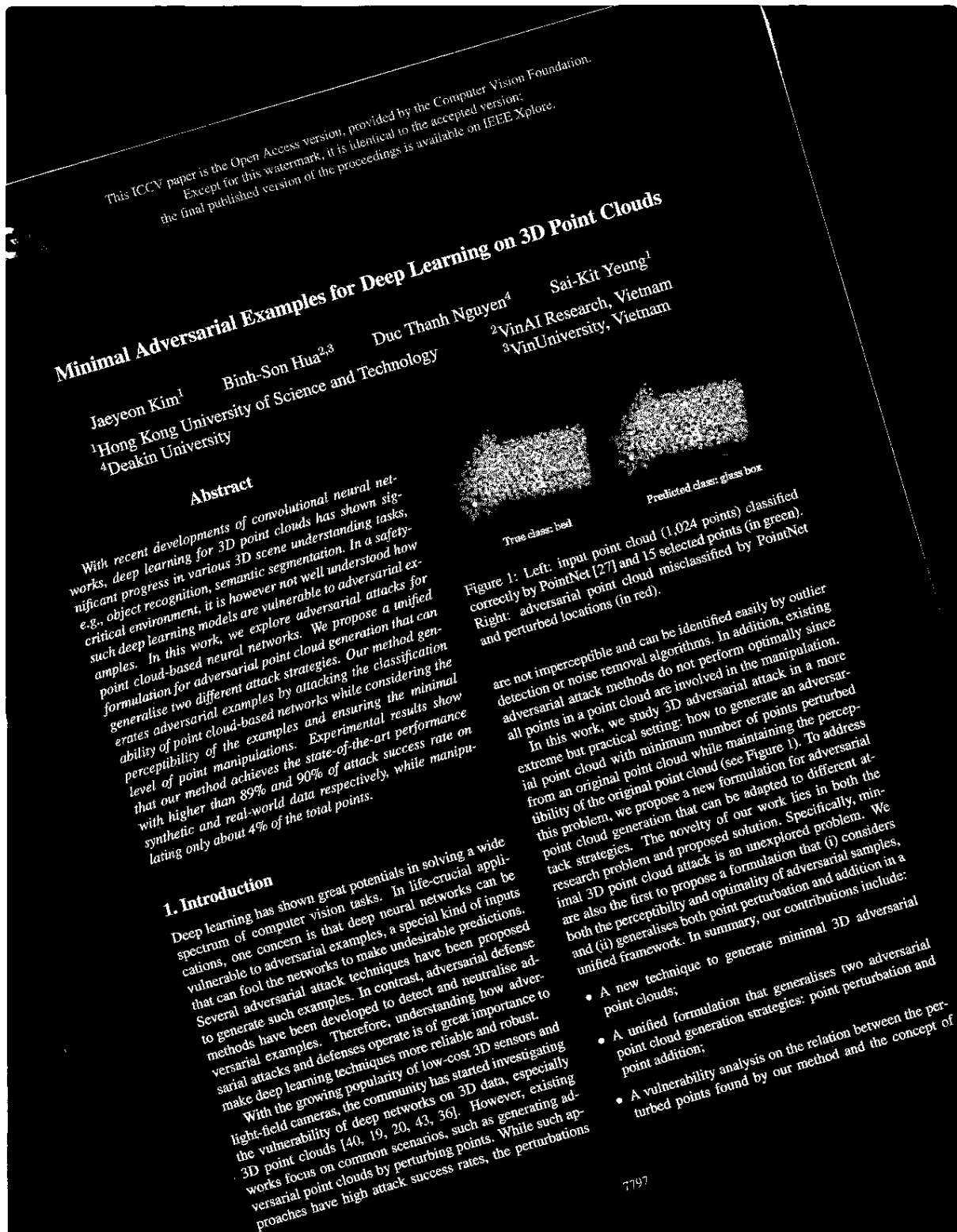
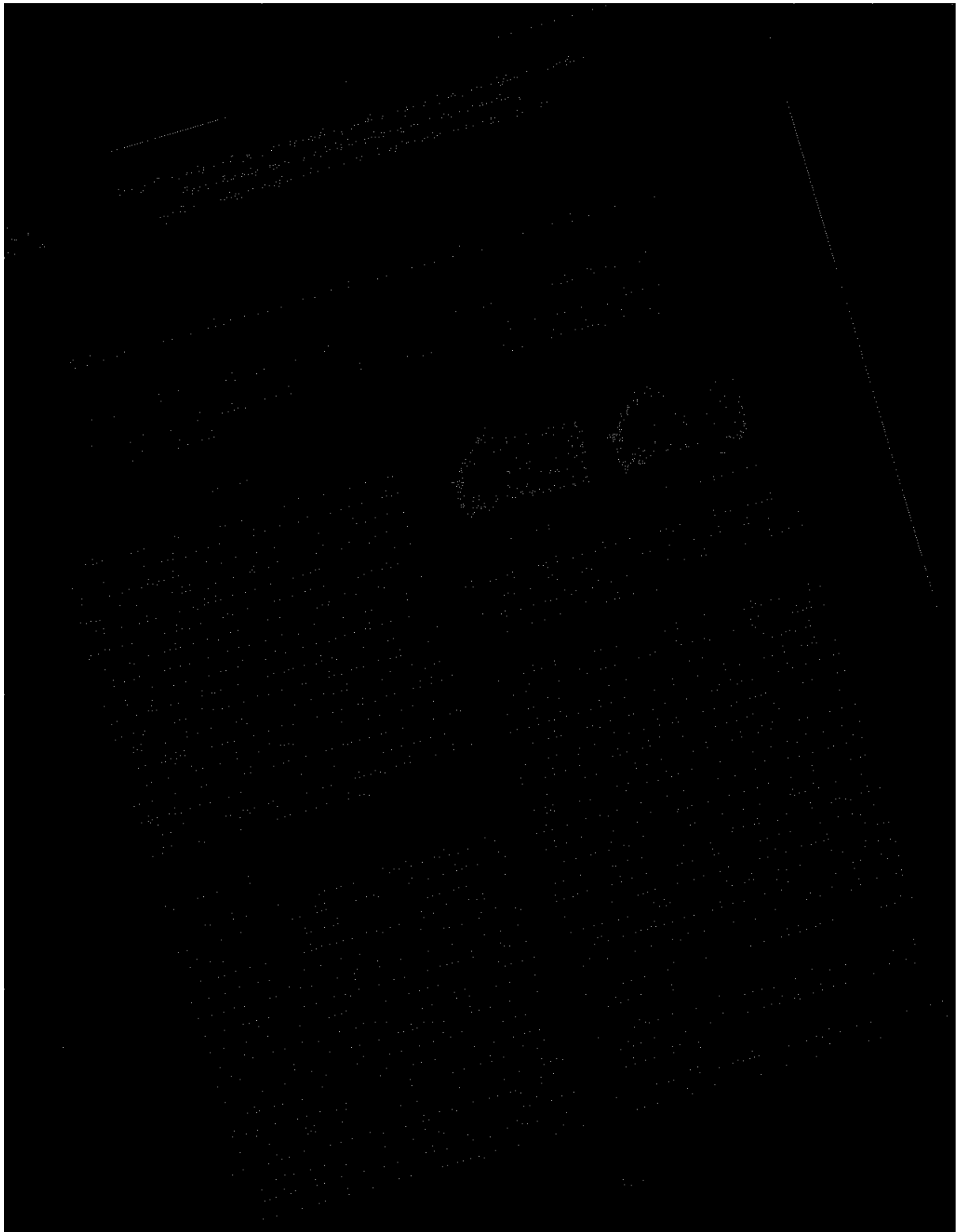
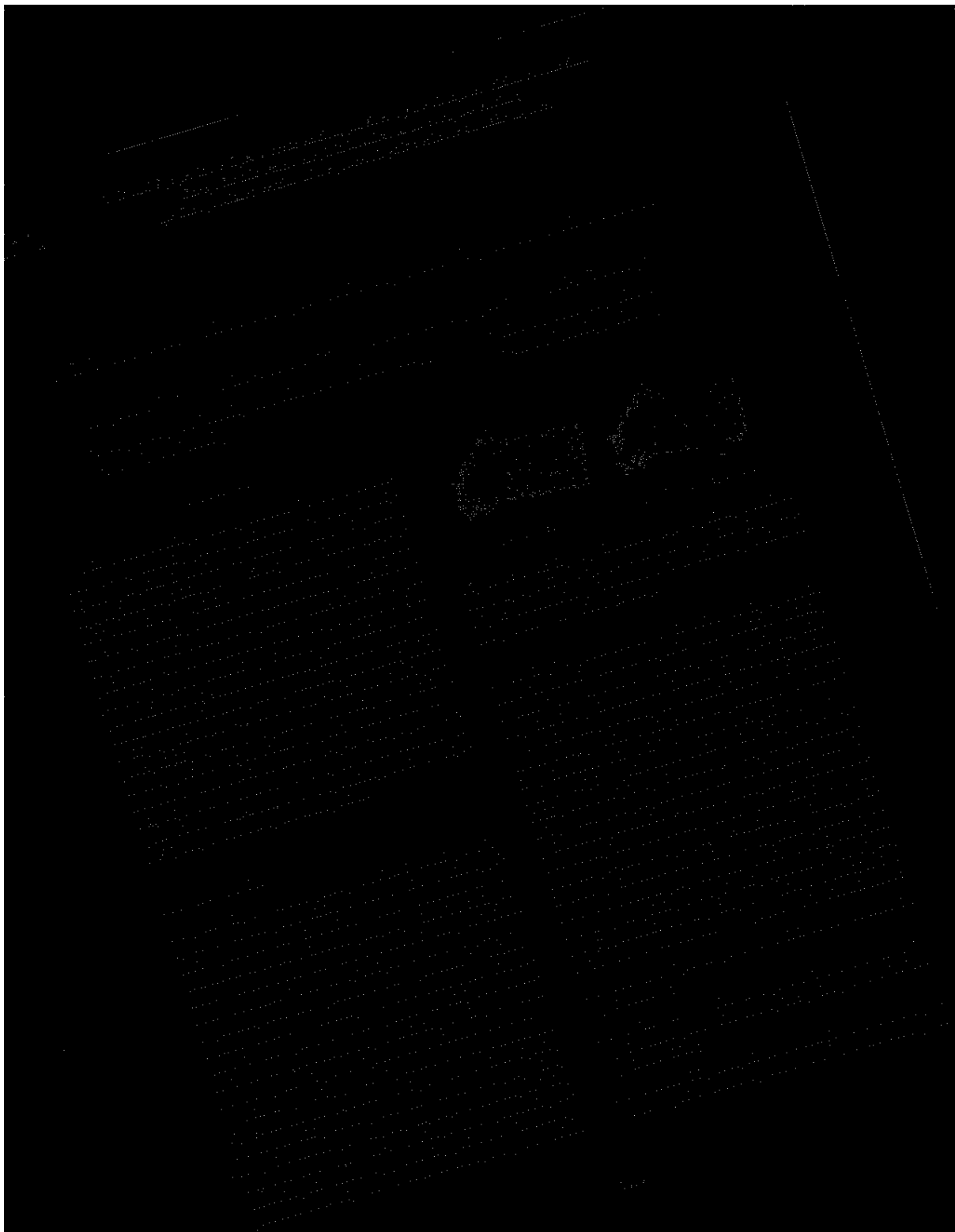


Fig 1.1: Negative image with candidate points where foreground pixels are candidates. [Mode 1]



**Fig 1.2:** Negative Image where candidate points are the mean of all coordinates in a component.  
[Mode 2]



**Fig 1.3:** Negative Image where point with max y coordinate value in a component is the candidate point. [Mode 3]

## Task 2:

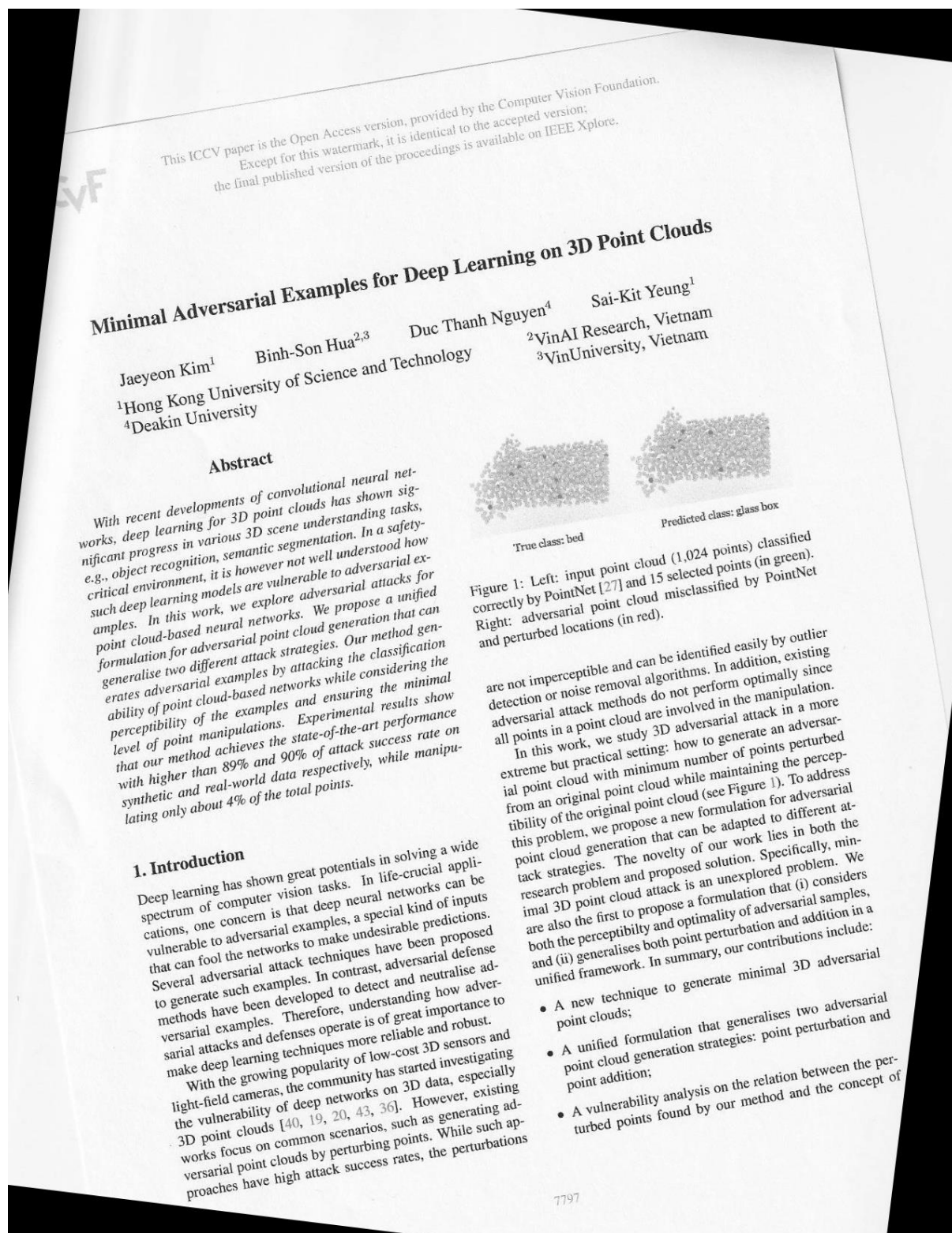


Fig 2.1: De-skewed doc with strategy 1 [All foreground pixels]

## Minimal Adversarial Examples for Deep Learning on 3D Point Clouds

Jaeyeon Kim<sup>1</sup> Binh-Son Hua<sup>2,3</sup> Duc Thanh Nguyen<sup>4</sup> Sai-Kit Yeung<sup>1</sup>

<sup>1</sup>Hong Kong University of Science and Technology

<sup>2</sup>VinAI Research, Vietnam

<sup>4</sup>Deakin University

<sup>3</sup>VinUniversity, Vietnam

### Abstract

With recent developments of convolutional neural networks, deep learning for 3D point clouds has shown significant progress in various 3D scene understanding tasks, e.g., object recognition, semantic segmentation. In a safety-critical environment, it is however not well understood how such deep learning models are vulnerable to adversarial examples. In this work, we explore adversarial attacks for point cloud-based neural networks. We propose a unified formulation for adversarial point cloud generation that can generalise two different attack strategies. Our method generates adversarial examples by attacking the classification ability of point cloud-based networks while considering the perceptibility of the examples and ensuring the minimal level of point manipulations. Experimental results show that our method achieves the state-of-the-art performance with higher than 89% and 90% of attack success rate on synthetic and real-world data respectively, while manipulating only about 4% of the total points.

### 1. Introduction

Deep learning has shown great potentials in solving a wide spectrum of computer vision tasks. In life-crucial applications, one concern is that deep neural networks can be vulnerable to adversarial examples, a special kind of inputs that can fool the networks to make undesirable predictions. Several adversarial attack techniques have been proposed to generate such examples. In contrast, adversarial defense methods have been developed to detect and neutralise adversarial examples. Therefore, understanding how adversarial attacks and defenses operate is of great importance to make deep learning techniques more reliable and robust.

With the growing popularity of low-cost 3D sensors and light-field cameras, the community has started investigating the vulnerability of deep networks on 3D data, especially 3D point clouds [40, 19, 20, 43, 36]. However, existing works focus on common scenarios, such as generating adversarial point clouds by perturbing points. While such approaches have high attack success rates, the perturbations

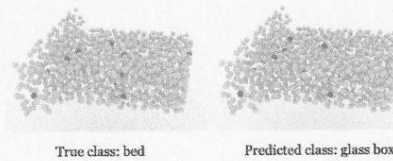


Figure 1: Left: input point cloud (1,024 points) classified correctly by PointNet [27] and 15 selected points (in green). Right: adversarial point cloud misclassified by PointNet and perturbed locations (in red).

are not imperceptible and can be identified easily by outlier detection or noise removal algorithms. In addition, existing adversarial attack methods do not perform optimally since all points in a point cloud are involved in the manipulation.

In this work, we study 3D adversarial attack in a more extreme but practical setting: how to generate an adversarial point cloud with minimum number of points perturbed from an original point cloud while maintaining the perceptibility of the original point cloud (see Figure 1). To address this problem, we propose a new formulation for adversarial point cloud generation that can be adapted to different attack strategies. The novelty of our work lies in both the research problem and proposed solution. Specifically, minimal 3D point cloud attack is an unexplored problem. We are also the first to propose a formulation that (i) considers both the perceptibility and optimality of adversarial samples, and (ii) generalises both point perturbation and addition in a unified framework. In summary, our contributions include:

- A new technique to generate minimal 3D adversarial point clouds;
- A unified formulation that generalises two adversarial point cloud generation strategies: point perturbation and point addition;
- A vulnerability analysis on the relation between the perturbed points found by our method and the concept of

Fig 2.2: De-skewed doc with strategy 2 [mean of all coordinates]

## Minimal Adversarial Examples for Deep Learning on 3D Point Clouds

Jaeyeon Kim<sup>1</sup> Binh-Son Hua<sup>2,3</sup> Duc Thanh Nguyen<sup>4</sup> Sai-Kit Yeung<sup>1</sup>

<sup>1</sup>Hong Kong University of Science and Technology

<sup>2</sup>VinAI Research, Vietnam

<sup>4</sup>Deakin University

<sup>3</sup>VinUniversity, Vietnam

### Abstract

With recent developments of convolutional neural networks, deep learning for 3D point clouds has shown significant progress in various 3D scene understanding tasks, e.g., object recognition, semantic segmentation. In a safety-critical environment, it is however not well understood how such deep learning models are vulnerable to adversarial examples. In this work, we explore adversarial attacks for point cloud-based neural networks. We propose a unified formulation for adversarial point cloud generation that can generalise two different attack strategies. Our method generates adversarial examples by attacking the classification ability of point cloud-based networks while considering the perceptibility of the examples and ensuring the minimal level of point manipulations. Experimental results show that our method achieves the state-of-the-art performance with higher than 89% and 90% of attack success rate on synthetic and real-world data respectively, while manipulating only about 4% of the total points.

### 1. Introduction

Deep learning has shown great potentials in solving a wide spectrum of computer vision tasks. In life-crucial applications, one concern is that deep neural networks can be vulnerable to adversarial examples, a special kind of inputs that can fool the networks to make undesirable predictions. Several adversarial attack techniques have been proposed to generate such examples. In contrast, adversarial defense methods have been developed to detect and neutralise adversarial examples. Therefore, understanding how adversarial attacks and defenses operate is of great importance to make deep learning techniques more reliable and robust.

With the growing popularity of low-cost 3D sensors and light-field cameras, the community has started investigating the vulnerability of deep networks on 3D data, especially 3D point clouds [40, 19, 20, 43, 36]. However, existing works focus on common scenarios, such as generating adversarial point clouds by perturbing points. While such approaches have high attack success rates, the perturbations

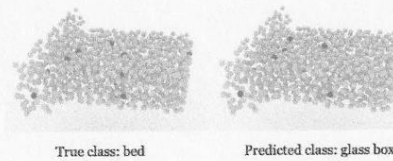


Figure 1: Left: input point cloud (1,024 points) classified correctly by PointNet [27] and 15 selected points (in green). Right: adversarial point cloud misclassified by PointNet and perturbed locations (in red).

are not imperceptible and can be identified easily by outlier detection or noise removal algorithms. In addition, existing adversarial attack methods do not perform optimally since all points in a point cloud are involved in the manipulation.

In this work, we study 3D adversarial attack in a more extreme but practical setting: how to generate an adversarial point cloud with minimum number of points perturbed from an original point cloud while maintaining the perceptibility of the original point cloud (see Figure 1). To address this problem, we propose a new formulation for adversarial point cloud generation that can be adapted to different attack strategies. The novelty of our work lies in both the research problem and proposed solution. Specifically, minimal 3D point cloud attack is an unexplored problem. We are also the first to propose a formulation that (i) considers both the perceptibility and optimality of adversarial samples, and (ii) generalises both point perturbation and addition in a unified framework. In summary, our contributions include:

- A new technique to generate minimal 3D adversarial point clouds;
- A unified formulation that generalises two adversarial point cloud generation strategies: point perturbation and point addition;
- A vulnerability analysis on the relation between the perturbed points found by our method and the concept of

Fig 2.3: De-skewed doc with strategy 3 (Point with max y coordinate)

### Task 3:

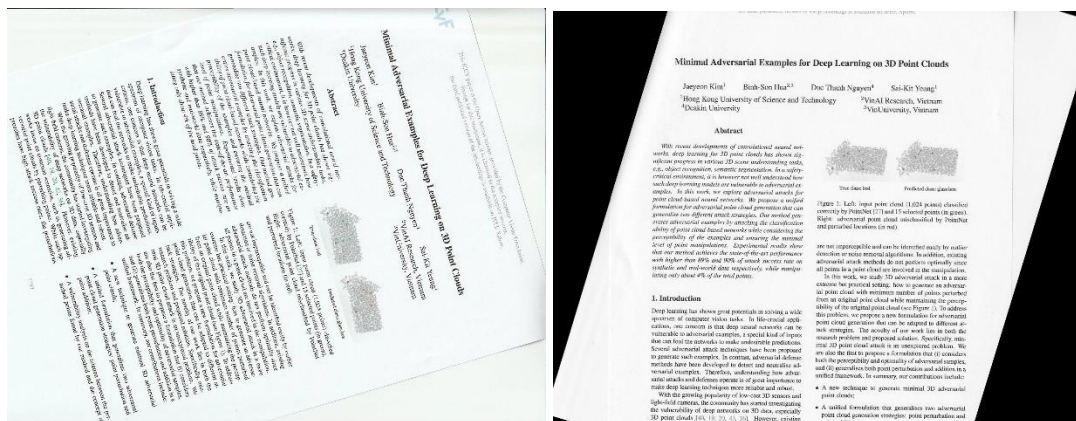


Fig 3.1: (a) doc\_1 (b) doc\_1 de-skewed

This shows the strength of the algorithm on too skewed images, we can see that the algorithm works fine on images with large skew values with same set of parameters. This shows Hough lines can be used with proper candidate selection strategy to de-skew documents.

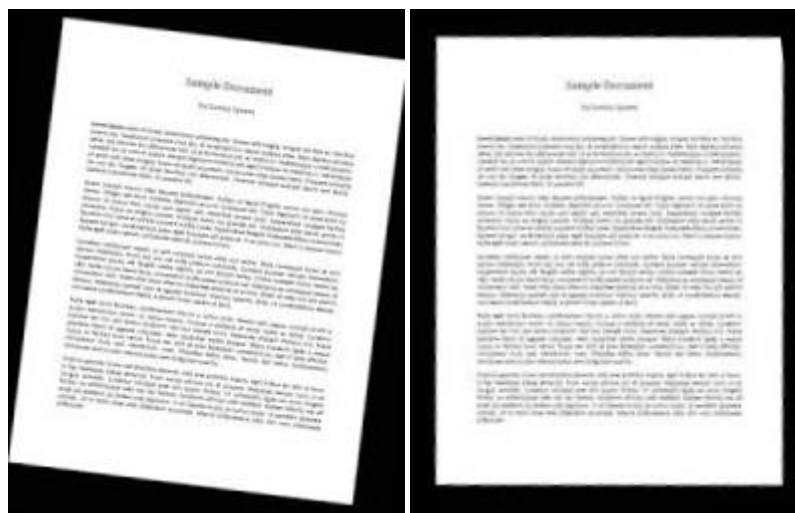


Fig 3.2: (a) doc\_2 (b) doc\_2 de-skewed

We can see that the algorithm works fine on this image where the skew angle is not large and Hough lines can be used to estimate the mean angle of the document. The algorithm works properly but because of low resolution (198 x 255) pytesseract is not able to parse pdf from this image.





Fig 3.3: doc\_3

This image was added to the dataset to include high cropping scenario, the algorithm fails on this image. The algorithm is not able to calculate angle on this image.

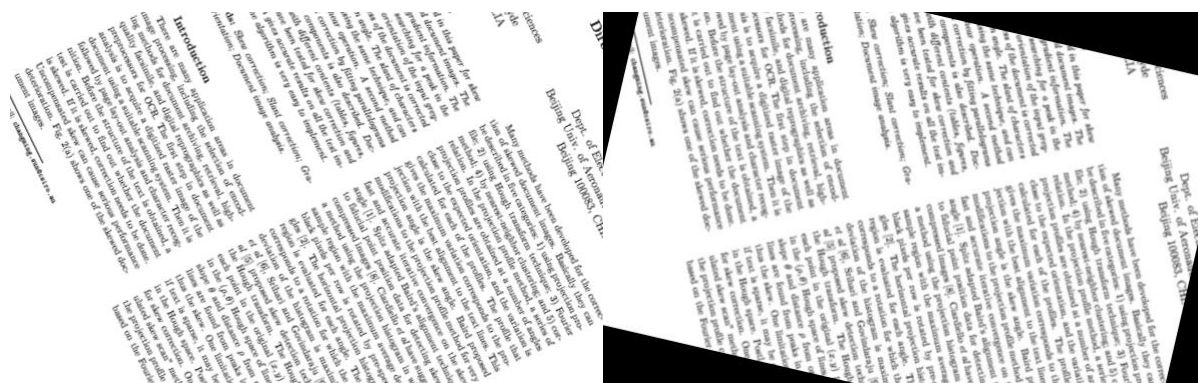


Fig 3.4: (a) doc\_4

(b) doc\_4 de-skewed

This document was included for another difficult scenario for this algorithm where we can see that the same parameters which work for slightly skewed document, will not work for documents that have a larger skew value. Although the document is horizontally aligned





Fig 3.5: doc\_5

The algorithm is not able to calculate lines angle for this document. This shows that the algorithm fails with the same set of parameters on such images.

The above experiments show that while the algorithm can work well with the same set of parameters on variety of images (doc, doc\_1, doc 2, and doc 4) but it will fail on images where candidate selection strategy fails, and no lines can be plotted. e.g.: (doc\_3 and doc\_5).

Although it is important to highlight that doc 4 was correctly aligned horizontally and doc 2 was correctly aligned but the text could not be parse due to low resolution.

Thus, we can conclude that Hough Transform is a strong method to de-skew a document provided the image is a **good quality image** and **not too low resolution**. Also, the image should have visible characters so that the **angles can be calculated**.

#### Task 4:

yacye™ es  
sions got  
ape Senne  
  
me gent

Fig 4.1: Result of print text from skewed document

Jaeyeon Kim! Binh-Son Hua?  
Hong Kong Unive  
'Deakin University

## Abstract

With recent developments of convolutional neural networks, deep learning for 3D point clouds has shown significant progress in various 3D scene understanding tasks, object recognition, semantic segmentation. In a safety-critical environment, it is however not well understood how

such deep learning models are vulnerable to adversarial examples for point cloud-based neural networks. We propose a unified formulation for adversarial point cloud generation that can generalise two different attack strategies. Our method generates adversarial examples by attacking the classification

Fig 4.2: Result of print text from de-skewed document. This highlights the importance of de-skewing in the domain of Document Recognition.

***Please Note: The ablation study for different thresholds has been shown in the notebook with visualisations. Task 2.2 (Parameter Setting)***

### The summary of ablation study:

*Parameters tried: 5, 10, 15*

*Density\_threshold of 15 yields best accuracy visually.*

### The summary of candidate points:

Mode 1 time: 5.16 seconds

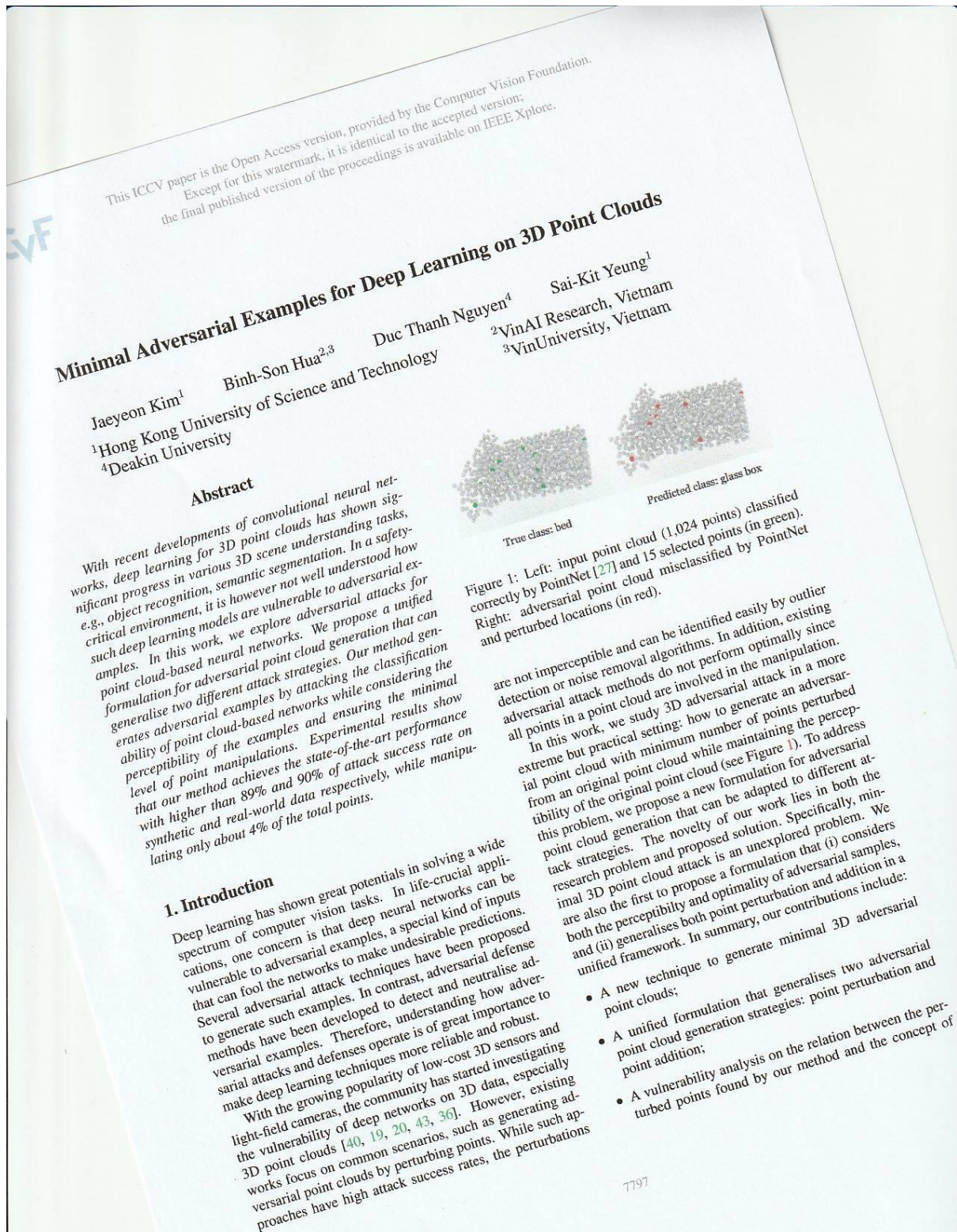
Mode 2 time: 0.62 seconds

Mode 3 time: 0.24 seconds

***Mode 3 was the fastest and most accurate, hence the optimal candidate points selection strategy.***

***This PDF contains all the images required for running the notebook in the appendix. The appendix also contains the pdf of doc.jpg and its de-skewed version.***

doc.jpg





The Journal of Supercomputing  
the final publication of record

**Minimal A Versarial Examples for Deep Learning on 3D Point Clouds**

Sun-Ki Yeung<sup>1</sup>      Duc Thanh Nguyen<sup>2</sup>      Binh-Son Hua<sup>2,3</sup>

<sup>1</sup>School of Information Technology, The Hong Kong Polytechnic University, Kowloon, Hong Kong

<sup>2</sup>VinAI Research, Vietnam

<sup>3</sup>VNU-University, Vietnam

© Springer 2020

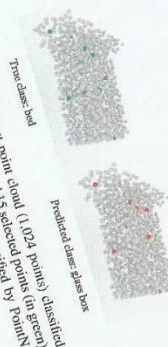
Jaeyeon Kim<sup>1</sup>  
<sup>1</sup>Hong Kong University  
<sup>4</sup>Deakin University

## Abstract

[illegible]

## 1. Introduction

Deep learning computer vision algorithms have been applied to a wide range of tasks, including image classification, object detection, and image segmentation. One of the main challenges in deep learning is the need for large amounts of labeled data. To address this, researchers have developed various techniques for data augmentation and transfer learning. In this paper, we propose a novel deep learning architecture for image classification that uses a combination of convolutional and recurrent neural networks. This architecture is designed to be robust to variations in the input data and to learn from a smaller amount of labeled data. We evaluate our architecture on a standard image classification dataset and show that it achieves state-of-the-art performance. Our results suggest that this architecture has the potential to be used in a wide range of other applications.



True class: bed

[illegible]

[illegible]

**Abstract**

A fast algorithm is presented in this paper for skew and slant correction in printed document images. The algorithm employs only the gradient information. The skew angle is obtained by searching for a peak in the histogram of the gradient orientation of a peak in the level image. The skewness of the document is corrected by a rotation at such an angle. The slant of characters can also be detected using the same technique, and can be corrected by a shear operation. A second method to the connected components is also described. Document images with different contents (tables, figures, and photos) have been tested for skew correction and slant correction. The algorithm gives accurate results on all the test images and the algorithm is very easy to implement.

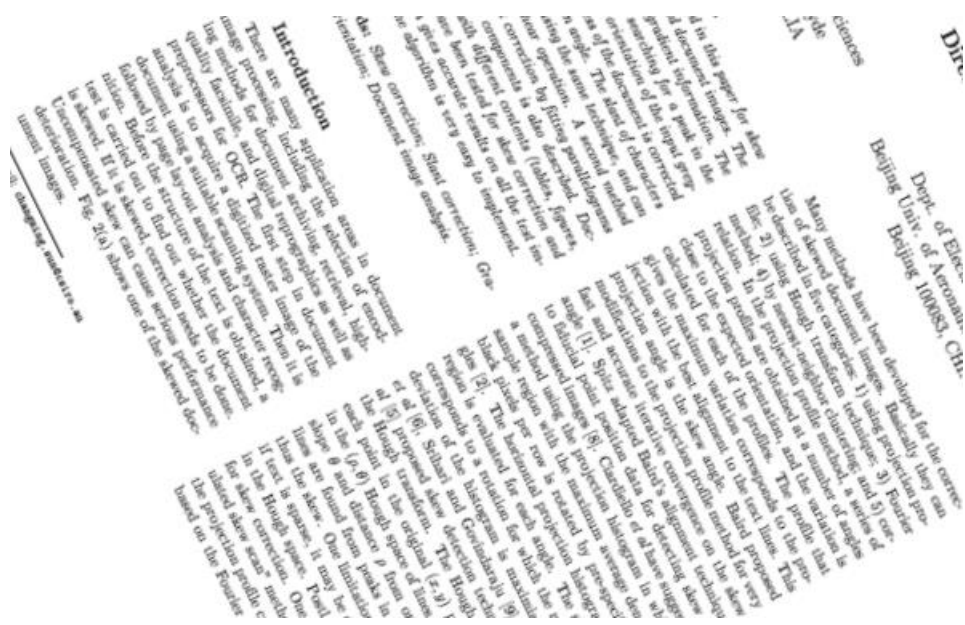
**Keywords:** Skew correction; Slant correction; Gradient orientation; Document image analysis.

**Introduction**

In many application areas in document processing, including the selection of encoded documents, digital archiving, retrieval, high-resolution scanning system of the document, OCR. The first step in document analysis and character recognition of the text is obtained, a decision on whether the document is skewed or not needs to be made. If it is skewed, a correction must be applied to cause serious performance degradation. One of the skewed doc-

ment images has been developed in five categories: 1) using Hough transform technique; 2) using nearest-neighbor clustering relation. In the projection profile method, the projection profiles are obtained at a number of orientations, and the maximum variation corresponds to the expected orientation, and the projection with the best alignment to the text line gives the maximum variation. The skew angle is the skew angle. The modifications to the projection profile method for fast and accurate iterative convergence on the compressed images [8]. Chiarillo et al. have suggested a method using the projection histogram for detecting black pixels with the maximum average density of sample regions with the projection histogram in which a single pixel region is rotated by pre-specified angles [2]. The horizontal projection histogram of the region is evaluated for each angle. The skew angle corresponds to a rotation for which the mean square deviation of the histogram is maximized. Nakano et al. [5] proposed skew detection technique based on the Hough transform. The Hough transform maps each point in the original  $(x, y)$  plane to all points in the Hough space of lines through  $(x, y)$  with slope  $\theta$  and distance  $\rho$  from origin. The dominant lines are found from peaks in the Hough space. If the text is sparse, one limitation of this method is that if text is sparse, it may be difficult to choose a peak for skew correction. Post [7] proposed a projection profile method. One of them is based on the Fourier transform method, and the other is based on the Fourier transform method.

doc\_4. png



doc\_5. png



**PDFs of doc and doc de-skewed:**



# Minimal Adversarial Examples for Deep Learning on 3D Point Clouds

Jaeyeon Kim<sup>1</sup>

<sup>1</sup>Hong Kong University of Science and Technology  
<sup>4</sup>Deakin University

Binh-Son Hua<sup>2,3</sup>

Duc Thanh Nguyen<sup>4</sup>

Sai-Kit Yeung<sup>1</sup>

<sup>2</sup>VinAI Research, Vietnam  
<sup>3</sup>VinUniversity, Vietnam

## Abstract

With recent developments of convolutional neural networks, deep learning for 3D point clouds has shown significant progress in various 3D scene understanding tasks, e.g., object recognition, semantic segmentation. In a safety-critical environment, it is however not well understood how such deep learning models are vulnerable to adversarial examples. In this work, we explore adversarial attacks for point cloud-based neural networks. We propose a unified formulation for adversarial point cloud generation that can generalise two different attack strategies. Our method generates adversarial examples by attacking the classification ability of point cloud-based networks while considering the perceptibility of the examples and ensuring the minimal level of point manipulations. Experimental results show that our method achieves the state-of-the-art performance with higher than 89% and 90% of attack success rate on synthetic and real-world data respectively, while manipulating only about 4% of the total points.

## 1. Introduction

Deep learning has shown great potentials in solving a wide spectrum of computer vision tasks. In life-crucial applications, one concern is that deep neural networks can be vulnerable to adversarial examples, a special kind of inputs that can fool the networks to make undesirable predictions. Several adversarial attack techniques have been proposed to generate such examples. In contrast, adversarial defense methods have been developed to detect and neutralise adversarial examples. Therefore, understanding how adversarial attacks and defenses operate is of great importance to make deep learning techniques more reliable and robust. With the growing popularity of low-cost 3D sensors and light-field cameras, the community has started investigating the vulnerability of deep networks on 3D data, especially 3D point clouds [40, 19, 20, 43, 36]. However, existing works focus on common scenarios, such as generating adversarial point clouds by perturbing points. While such approaches have high attack success rates, the perturbations

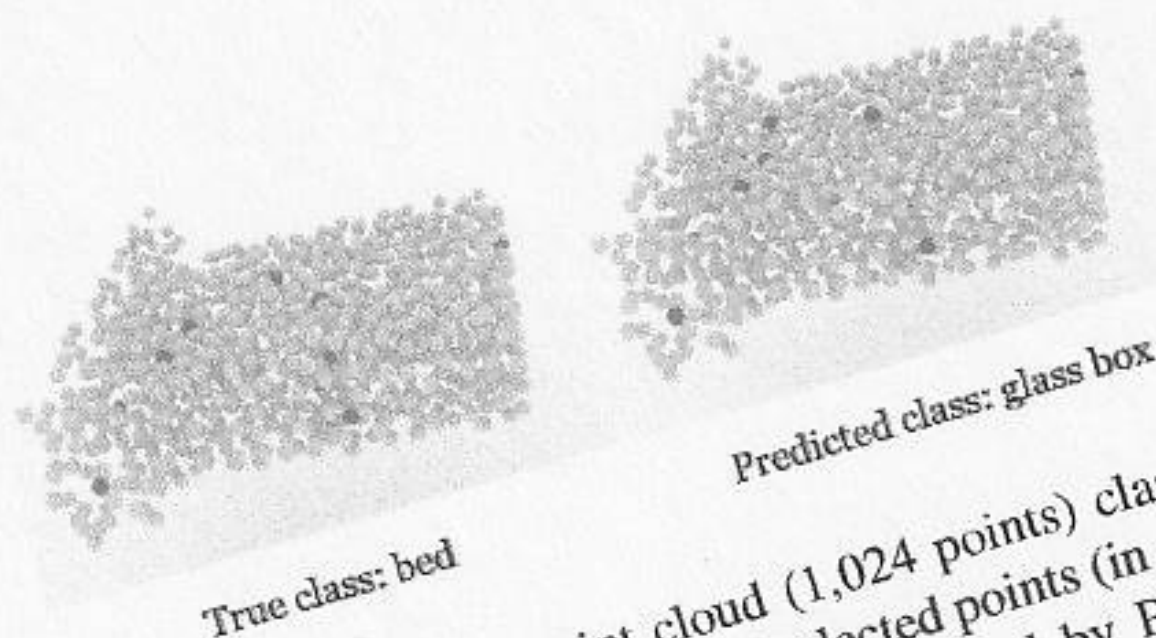


Figure 1: Left: input point cloud (1,024 points) classified correctly by PointNet [27] and 15 selected points (in green). Right: adversarial point cloud misclassified by PointNet and perturbed locations (in red).

are not imperceptible and can be identified easily by outlier detection or noise removal algorithms. In addition, existing adversarial attack methods do not perform optimally since all points in a point cloud are involved in the manipulation. In this work, we study 3D adversarial attack in a more extreme but practical setting: how to generate an adversarial point cloud with minimum number of points perturbed from an original point cloud while maintaining the perceptibility of the original point cloud (see Figure 1). To address this problem, we propose a new formulation for adversarial point cloud generation that can be adapted to different attack strategies. The novelty of our work lies in both the research problem and proposed solution. Specifically, minimal 3D point cloud attack is an unexplored problem. We are also the first to propose a formulation that (i) considers both the perceptibility and optimality of adversarial samples, and (ii) generalises both point perturbation and addition in a unified framework. In summary, our contributions include:

- A new technique to generate minimal 3D adversarial point clouds;
- A unified formulation that generalises two adversarial point cloud generation strategies: point perturbation and point addition;
- A vulnerability analysis on the relation between the perturbed points found by our method and the concept of



## Minimal Adversarial Examples for Deep Learning on 3D Point Clouds

Jaeyeon Kim<sup>1</sup>   Binh-Son Hua<sup>2,3</sup>   Duc Thanh Nguyen<sup>4</sup>   Sai-Kit Yeung<sup>1</sup>

<sup>1</sup>Hong Kong University of Science and Technology

<sup>2</sup>VinAI Research, Vietnam

<sup>4</sup>Deakin University

<sup>3</sup>VinUniversity, Vietnam

### Abstract

With recent developments of convolutional neural networks, deep learning for 3D point clouds has shown significant progress in various 3D scene understanding tasks, e.g., object recognition, semantic segmentation. In a safety-critical environment, it is however not well understood how such deep learning models are vulnerable to adversarial examples. In this work, we explore adversarial attacks for point cloud-based neural networks. We propose a unified formulation for adversarial point cloud generation that can generalise two different attack strategies. Our method generates adversarial examples by attacking the classification ability of point cloud-based networks while considering the perceptibility of the examples and ensuring the minimal level of point manipulations. Experimental results show that our method achieves the state-of-the-art performance with higher than 89% and 90% of attack success rate on synthetic and real-world data respectively, while manipulating only about 4% of the total points.

### 1. Introduction

Deep learning has shown great potentials in solving a wide spectrum of computer vision tasks. In life-crucial applications, one concern is that deep neural networks can be vulnerable to adversarial examples, a special kind of inputs that can fool the networks to make undesirable predictions. Several adversarial attack techniques have been proposed to generate such examples. In contrast, adversarial defense methods have been developed to detect and neutralise adversarial examples. Therefore, understanding how adversarial attacks and defenses operate is of great importance to make deep learning techniques more reliable and robust.

With the growing popularity of low-cost 3D sensors and light-field cameras, the community has started investigating the vulnerability of deep networks on 3D data, especially 3D point clouds [40, 19, 20, 43, 36]. However, existing works focus on common scenarios, such as generating adversarial point clouds by perturbing points. While such approaches have high attack success rates, the perturbations

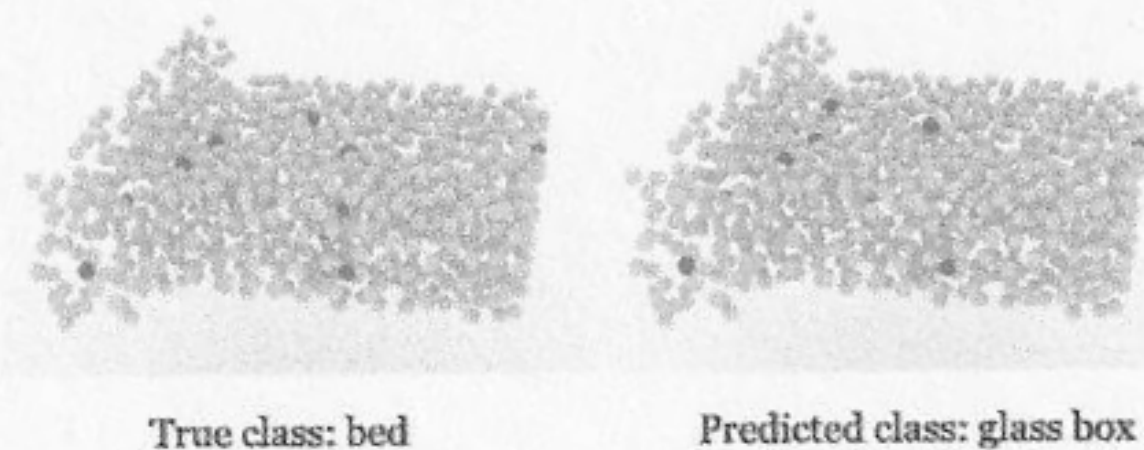


Figure 1: Left: input point cloud (1,024 points) classified correctly by PointNet [27] and 15 selected points (in green). Right: adversarial point cloud misclassified by PointNet and perturbed locations (in red).

are not imperceptible and can be identified easily by outlier detection or noise removal algorithms. In addition, existing adversarial attack methods do not perform optimally since all points in a point cloud are involved in the manipulation.

In this work, we study 3D adversarial attack in a more extreme but practical setting: how to generate an adversarial point cloud with minimum number of points perturbed from an original point cloud while maintaining the perceptibility of the original point cloud (see Figure 1). To address this problem, we propose a new formulation for adversarial point cloud generation that can be adapted to different attack strategies. The novelty of our work lies in both the research problem and proposed solution. Specifically, minimal 3D point cloud attack is an unexplored problem. We are also the first to propose a formulation that (i) considers both the perceptibility and optimality of adversarial samples, and (ii) generalises both point perturbation and addition in a unified framework. In summary, our contributions include:

- A new technique to generate minimal 3D adversarial point clouds;
- A unified formulation that generalises two adversarial point cloud generation strategies: point perturbation and point addition;
- A vulnerability analysis on the relation between the perturbed points found by our method and the concept of