



Title	PointNet + + Based Concealed Object Classification Utilizing an FMCW Millimeter-Wave Radar
Author(s)	Wang, Yaheng; Su, Jie; Murakami, Hironaru et al.
Citation	Journal of Infrared, Millimeter, and Terahertz Waves. 2024, 45(11-12), p. 1040-1057
Version Type	VoR
URL	<a href="https://hdl.handle.net/11094/98397">https://hdl.handle.net/11094/98397</a>
rights	This article is licensed under a Creative Commons Attribution 4.0 International License.
Note	

*The University of Osaka Institutional Knowledge Archive : OUKA*

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka



# PointNet++ Based Concealed Object Classification Utilizing an FMCW Millimeter-Wave Radar

Yaheng Wang<sup>1</sup> · Jie Su<sup>1</sup> · Hironaru Murakami<sup>1</sup> · Masayoshi Tonouchi<sup>1</sup>

Received: 9 July 2024 / Accepted: 30 September 2024  
© The Author(s) 2024

## Abstract

In the field of millimeter-wave (MMW) imaging, the integration of artificial intelligence (AI) has emerged as a crucial solution for addressing automation challenges. In this study, concealed object classification was successfully achieved on point cloud data from MMW radar high-precision imaging using the PointNet++ deep learning method. The utilized dataset comprises point cloud data generated through the transformation of 3D models and reconstruction of physical objects with an accuracy of less than 1 mm via MMW radar scanning. Classification accuracy was significantly improved by introducing data enhancement techniques, including the generation of homologous data and optimization of sampling points. After several evaluations, 300 epochs of training were conducted using 8192 sampling points, the results showed an accuracy of 0.998 for the training dataset and 0.996 for the test dataset. Moreover, evaluations of samples not included in the original dataset as well as multi-surface scans of concealed objects within the cardboard both resulted in correct predictions, which further validates the effectiveness and reliability of the study and demonstrates the potential of AI applied to MMW imaging.

**Keywords** Millimeter wave imaging · Millimeter wave radar · Object classification · Deep learning

## 1 Introduction

Millimeter waves (MMWs) refer to electromagnetic waves with frequencies ranging from 30 to 300 GHz and wavelengths between 10 and 1 mm. There are various unique features that make them highly attractive. One of the distinctive features is high transparency to various non-metallic objects, including cardboard and clothing, allowing imaging through obstacles that impede visibility with traditional measurement techniques. Moreover, unlike other radiation-based methods, MMW

---

✉ Masayoshi Tonouchi  
tonouchi.masayoshi.ile@osaka-u.ac.jp

<sup>1</sup> Institute of Laser Engineering, Osaka University, Osaka, Japan

imaging techniques do not pose exposure risks, thereby enhancing safety for both operators and subjects under inspection. Additionally, MMWs exhibit exceptional weather resistance, maintaining functionality even under adverse weather conditions such as rain and fog. These inherent advantages have spurred the development of various applications utilizing MMWs, including vehicle-mounted radar for driving assistance systems and autonomous driving [1–3], biometric monitoring technologies for non-contact measurement of vital signs [4–7], infrastructure inspection for detecting cracks beneath paint coatings [8–11], non-destructive testing for object defect detection [12–15], and security technologies for detecting hazardous objects [16–18]. Furthermore, with the rapid advancement of artificial intelligence (AI) technology, integration with AI has emerged as a significant breakthrough in these fields [19]. For instance, through deep learning techniques, applications such as autonomous vehicles can accurately detect and identify various objects in the environment using LiDAR point cloud data [20].

In the field of MMW imaging, most of the AI integration research has been focused on 2D images due to the maturity of the 2D imaging processing techniques. In passive imaging, machine learning has successfully identified the location of concealed objects on the human body with an accuracy of approximately 94% [21]. In non-passive imaging, a 27 GHz with 5 GHz bandwidth MMW radar is used to scan the human body, and deep learning techniques applied to the acquired 2D images successfully segment body parts and estimate poses with an accuracy of approximately 98.4% [22]. However, 2D imaging techniques suffer from an inherent limitation: the inability to identify objects at different depths. Currently, AI integration research on object classification using MMW 3D point clouds is relatively limited. This limitation stems from the low resolution of the 3D images obtained by the MMW imaging system, which is due to the lack of high-precision image processing techniques to accurately reconstruct object contours [23]. For example, in an AI integration study [24], a 79 GHz with 1.6 GHz bandwidth MMW radar was used to classify large objects such as cars and buses; however, due to the lack of high-precision image processing techniques, the resolution was approximately 60 cm, allowing only simple contours of large objects to be identified. Nevertheless, our previous research has successfully established a multiple-input multiple-output synthetic aperture radar (MIMO-SAR) frequency-modulated continuous wave (FMCW) MMW radar imaging system and developed a technique for precise 2D and 3D reconstruction using observed point cloud data [25]. By effectively reducing noise in smooth surface 3D point cloud data, a resolution of less than 1 mm was achieved. With the successful acquisition of high-precision data using the MMW radar, key technical support has been provided for the further development of high-precision MMW 3D imaging and AI integration, which makes it feasible to perform precise object classification of MMW imaging through deep learning.

Therefore, this study aims to combine MMW 3D imaging and deep learning to achieve the automatic object classification of concealed objects by observing high-precision point cloud features. This work fills a gap in the application of AI integration in high-precision MMW 3D imaging and contributes to addressing the

challenges of automation and productivity enhancement in MMW imaging applications, thereby promoting the development and application of relevant technologies.

## 2 PointNet++ Architecture

In the field of deep learning for point cloud data, the processing and analysis of such data pose challenging tasks due to their high dimensionality, non-structural, and irregular nature. However, with the continuous advancement of deep learning technology, a series of deep learning methods for point cloud data have emerged, including point-based methods such as PointNet, PointNet++, Dynamic Graph CNN, and PointCNN [26–29], voxel-based methods such as VoxNet, OctNe, and GridNet [30–32], and image-based methods such as Frustum PointNet and MVCNN [33, 34].

Point-based methods exhibit several advantages over voxel-based and image-based methods: the raw point cloud data can be directly handled without requiring additional preprocessing or transformation; more flexibility in adapting to different shapes and structures of point cloud data, enabling better capture of local and global information; and the ability to process point cloud data of arbitrary shapes and structures, independently of the mesh structure.

In point-based methods, PointNet++, which was proposed by Qi et al. from Stanford University in 2017, introduces a hierarchical point set aggregation mechanism, progressively aggregating feature information of point cloud data from local to global levels [27]. This multi-level feature learning mechanism enables PointNet++ to better capture local and global information of point cloud data, exhibiting strong flexibility and adaptability. Therefore, PointNet++ holds significant advantages in the field of deep learning for point cloud data, leading this study to select PointNet++ as the learning model.

The deep learning process of PointNet++ is as follows: input point cloud data; extract local features and combine them to obtain local and global features; aggregate features at different levels through hierarchical aggregation; input the processed features into network layers to accomplish classification.






## 3 Composition of the Dataset

Accurate recognition by neural networks relies heavily on extensive data for continual feature learning. This necessitates the creation of datasets, critical for training and evaluating deep learning models. These datasets typically encompass training, test, and evaluation subsets. The training dataset enables the model to learn input–output relationships and adjust parameters to minimize prediction errors, while the test dataset fine-tunes the model hyper-parameters and evaluates performance after each round of training. Finally, the evaluation dataset evaluates the generalization and accuracy of the trained model on new data.

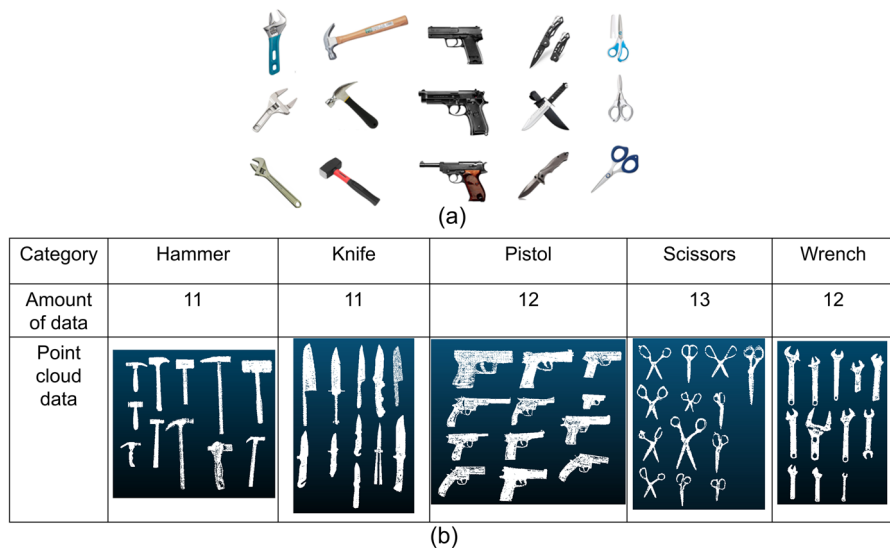
For the current commonly used 3D object recognition datasets, such as ShapeNet and ModelNet40 [35, 36], since the objects in them are mostly used for object recognition in rooms or large spaces, while the objects required in this

study are smaller and precise, therefore, we downloaded the 3D objects in GLB format from the Internet and converted all of them to point cloud objects in PLY format. For the initial training dataset, we downloaded five categories, namely hammer, knife, pistol, scissors, and wrench, as point cloud data for training, comprising 66 data entries for hammers, 70 for knives, 70 for pistols, 70 for scissors, and 68 for wrenches. Part of the training dataset of point cloud data is shown in Fig. 1.

For the testing set, we utilized an MMW imaging system to acquire point cloud data of objects. Considering the time and monetary costs associated with scanning and processing MMW imaging point cloud data and purchasing samples, we acquired a test dataset consisting of 11 data entries for hammers, 11 for knives, 12 for pistols, 13 for scissors, and 12 for wrenches. It is noteworthy that while actual guns are typically metallic, the samples we acquired were plastic replicas. To obtain a more comprehensive sample contour, we applied metallic exterior paint to the plastic and wooden parts of the samples. The low reflective intensity of plastic and wood compared to metal can cause uneven contours in the reconstructed point cloud. By applying a conductive coating, the reflective intensity of these regions is significantly increased, resulting in a more uniform point cloud contour of the object. However, this process is only used during dataset collection to improve the quality of training models. In practical application scenarios, we will use samples without metallic coatings for validation experiments to ensure the model's performance is more aligned with real-world conditions and to verify its effectiveness when handling unprocessed objects. Part of the actual sample of the test set is shown in Fig. 2a. Each object underwent scanning from four sides utilizing an IWR1443 module affixed to a two-axis mechanical stage. Analyzing IF signals from transmitted and received chirps produced Range FFT spectra, offering distance and reflection intensity information based on frequency. Subsequently, point cloud data reconstruction was accomplished through our previously developed precise 3D reconstruction algorithm, which includes monolayering, smoothing, and 3D stitching [25]. By effectively reducing noise in 3D point cloud data, a resolution of less than 1 mm was achieved. The results of the reconstruction process for all objects are visually depicted in Fig. 2b.

Category	Hammer	Knife	Pistol	Scissors	Wrench
Amount of data	66	70	70	70	68
Point cloud data					

**Fig. 1** Composition of the training dataset



**Fig. 2** **a** Actual sample of the test dataset. **b** Composition of the test dataset

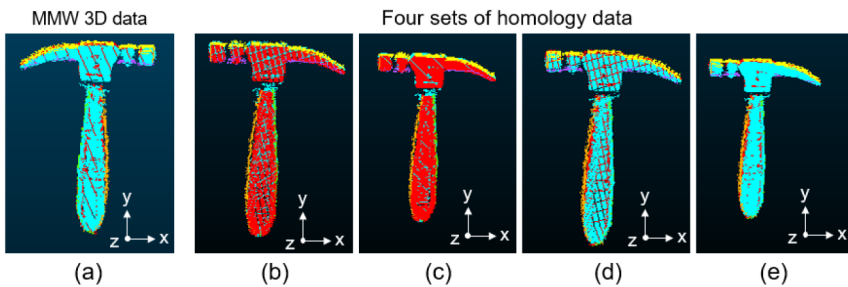
#### 4 Impact of Homologous Dataset

A dataset consisting of five objects was created in the previous section, where the training and test datasets showed significant differences in point cloud contours. By comparing the training dataset (shown in Fig. 1) with the test dataset (shown in Fig. 2b), we can observe the lack of noise and smooth edges of object contours in the 3D models that were downloaded from the Internet. In contrast, for the dataset obtained from the MMW imaging system, the vibration of the scanner or the interference phenomenon of the reflected signals from the MMW resulted in significant noise in the captured data and non-uniform data distribution. It is difficult to achieve high accuracy only on the above training and test sets, as specific features of the MMW point cloud cannot be learned efficiently by using a training dataset that does not contain noise. In order to increase the diversity of data obtained from the MMW system in the training dataset, we performed transformation operations such as panning, scaling, zooming, and rotating on the test set to generate homologous data, which were added to the initial training dataset. It is important to note that panning operations on a 3D point cloud can change its position, scaling and zooming can change its size, and rotation can change its orientation. With these transformations, we can generate new data samples that are similar but not identical to the original data samples, thus enriching the training dataset. This data augmentation technique is widely used in deep learning to help increase the diversity of training data and improve the generalization ability and robustness of the model [37–40]. It is also worth noting that real-world scenarios often involve objects appearing at different scales, orientations, and positions. Data augmentation techniques can effectively simulate these situations, thus making the model more robust and adaptable to a variety of real-world environments.

We performed four different homologous data generation operations on the test set: scaling up by  $1.1\times$  and rotating by 180 degrees; scaling down by  $0.95\times$  and rotating by 180 degrees; scaling up by  $1.1\times$  and mirroring; and scaling down by  $0.9\times$  and mirroring. It is worth noting that in order to minimize the difference in point distance, the point cloud was scaled between 0.9 and 1.1 times. These transformations produce four sets of homogeneous data for enriching the initial training dataset. Taking the hammer as an example, the results of these homologous data are shown in Fig. 3. Figure 3a shows the point cloud data of the hammer after reconstruction by the MMW imaging system, while Fig. 3 b, c, d, and e show the homologous data after scaling up by a factor of 1.1 and rotating it by 180 degrees around Y-axis, scaling down by a factor of 0.95 and rotating it by 180 degrees around Y-axis, scaling up by a factor of 1.1 and mirroring it, and scaling down by a factor of 0.9 and mirroring it, respectively. It is clear from this example that the homologous data is not identical to the test set data in terms of position, size, and orientation, and is substantially a new sample. Adding these homologous data to the training set can effectively increase the diversity of the training data, and this technique helps to train the model to recognize and understand the features of objects after MMW imaging processing, making the model more robust and adaptable.

The objective of adding homologous data is to allow the neural network to learn these features even if the homologous data contains features specific to the MMW point cloud, such as missing, hollow, noise, and overlapping features. However, if the amount of homologous data is too large, defective features in the homologous data may dominate the learning process, resulting in the network not being able to efficiently learn the shape features of the object, which will reduce the correctness rate. In order to determine the appropriate proportion of homologous data, we added homologous data with factors ranging from 0 to 4 in the initial training dataset, respectively. The specific homologous data added include:

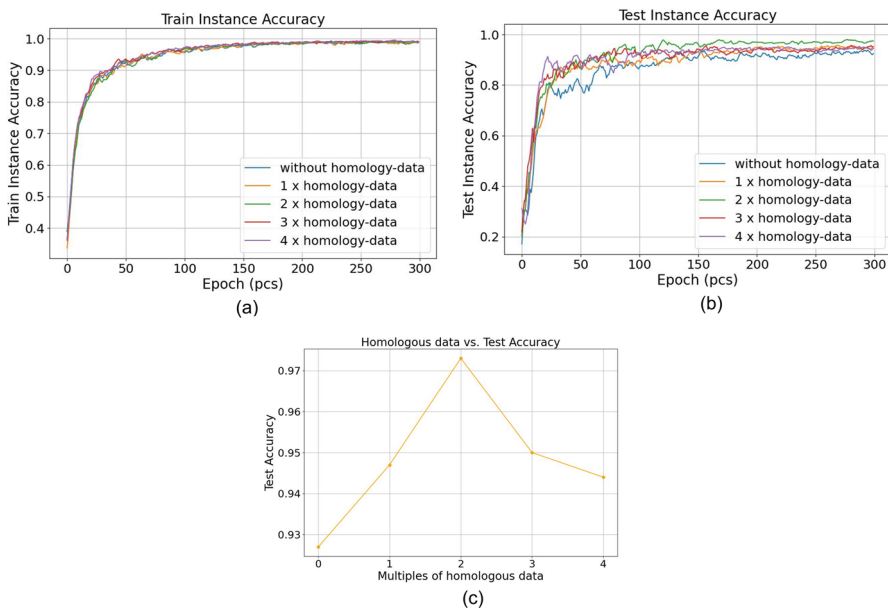
- Increase homologous data by a factor of 1: Scale up  $1.1\times$  and rotate by 180 degrees.
- Increase homologous data by a factor of 2: Scale up  $1.1\times$  and rotate by 180 degrees, scale down  $0.95\times$  and rotate 180 degrees.



**Fig. 3** **a** Point cloud data after MMW imaging reconstruction. **b** Scaling up by  $1.1\times$  and rotating by 180 degrees around Y-axis. **c** Scaling down by  $0.95\times$  and rotating by 180 degrees around Y-axis. **d** Scaling up by  $1.1\times$  and mirroring. **e** Scaling down by  $0.9\times$  and mirroring

- Increase homologous data by a factor of 3: Scale up  $1.1\times$  and rotate 180 degrees, scale down  $0.95\times$  and rotate 180 degrees, scale up  $1.1\times$  and mirror.
- Increase homologous data by a factor of 4: Scale up  $1.1\times$  and rotate 180 degrees, scale down  $0.95\times$  and rotate 180 degrees, scale up  $1.1\times$  and mirror, scale down  $0.9\times$  and mirror.

We employed 2048 sampling points and performed 300 epochs of training. The accuracy rate of the training dataset is shown in Fig. 4a and the accuracy rate of the test dataset is shown in Fig. 4b. From the results of the training set, it can be seen that adding homologous data has almost no effect, which is because the complexity of the model and the learning parameters have not been changed. From the results of the test dataset, it can be seen that with the increase of homologous data, the accuracy rate reaches the peak when it is increased to the factor of 2, and then gradually decreases. This is because the neural network has learned the specific features of the MMW point cloud data by increasing the homologous data up to a factor of 2. However, when the homologous data is increased to a factor of 3–4, the neural network learns an excessive number of defective features, leading to a decrease in the accuracy rate. Figure 4c shows this result more clearly, that is the highest accuracy rate, about 0.973, is achieved when the homologous data is increased to the factor of 2.



**Fig. 4** **a** The accuracy rate of the training dataset. **b** The accuracy rate of the test dataset. **c** The relationship between the multiples of homologous data and accuracy rate



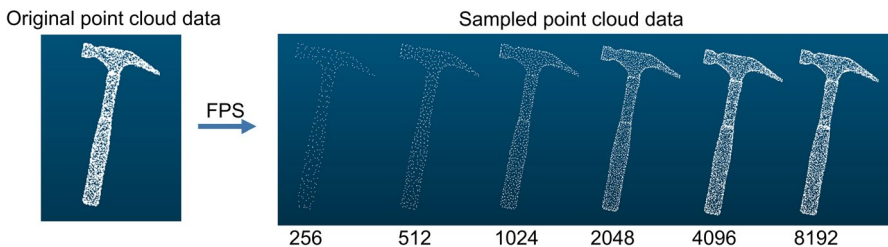
## 5 Impact of Sampling Points

The PointNet++ methodology employs the farthest point sampling (FPS) method to simplify massive point cloud data [27]. A higher number of sampling points captures more features, but also significantly increases the size of the dataset. This indicates that increasing the number of sampling points improves accuracy, but excessive sampling points can dramatically increase the training time of the neural network, making it difficult to build a practical system. To determine the optimal number of sampling points that balances high accuracy with acceptable learning time, we conducted a series of training experiments that gradually changed the number of sampling points without changing the dataset.

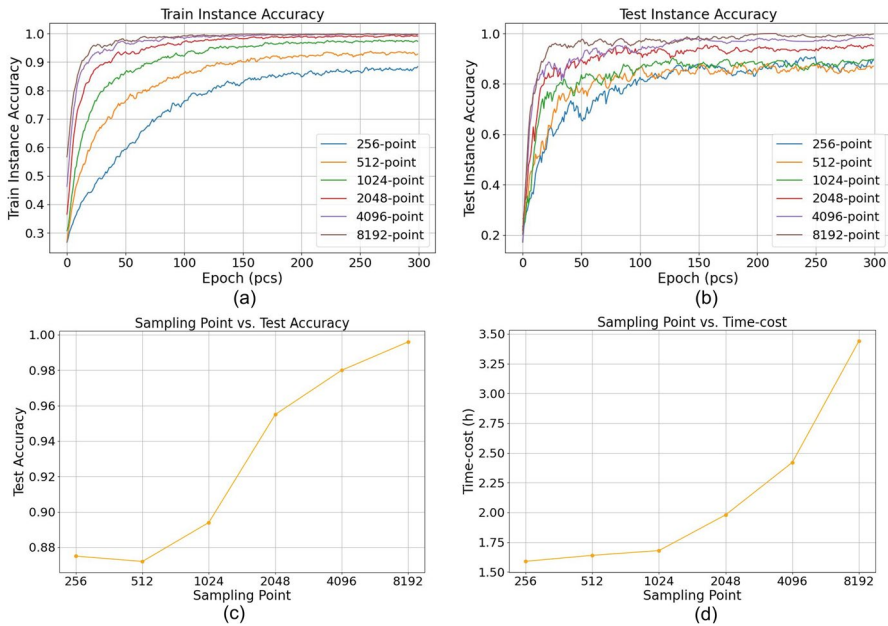
We added homologous data by a factor of 2 to the initial training dataset for training and used the point cloud data obtained from the MMW imaging system as the test dataset to train for 300 epochs. The number of sampling points started from 256 points and multiplied exponentially until it reached 8192 points (the limit of the device), as shown in Fig. 5.

Figure 6a illustrates the accuracy rate of the training dataset as the number of sampling points increases. As the number of sampling points increases, more shape features can be extracted, which improves the accuracy rate of the training set. When the number of sampling points is 8192, the training set has the highest accuracy rate, which is about 0.998. Figure 6b shows how the accuracy rate of the test dataset varies with the number of sampling points. The difference between 256 and 512 points is not significant, but the accuracy rate of the test dataset gradually increases from 1024 points onwards. Figure 6c shows more clearly the result, with the highest accuracy rate of the test set, about 0.996, at a sampling point count of 8192.

An increase in the number of sampling points represents an increase in computation and training time. Figure 6d shows the learning time required to complete 300 epochs with the number of sampling points ranging from 256 to 8192 points. As the number of sampling points increases, the required training time also increases. For 8192 sample points, the required learning time is about 3.44 h, which is close to the limit of our computer equipment. In order to obtain the best accuracy, we chose to use 8192 sampling points to train the optimal model. In addition, another important influence on the training time is the hardware of the computer used. The more the number and performance of GPUs, the shorter the training time. In this study, the GPU we used was the NVIDIA GeForce RTX 2070 SUPER. In future



**Fig. 5** Results of different numbers of sampling points using the FPS method

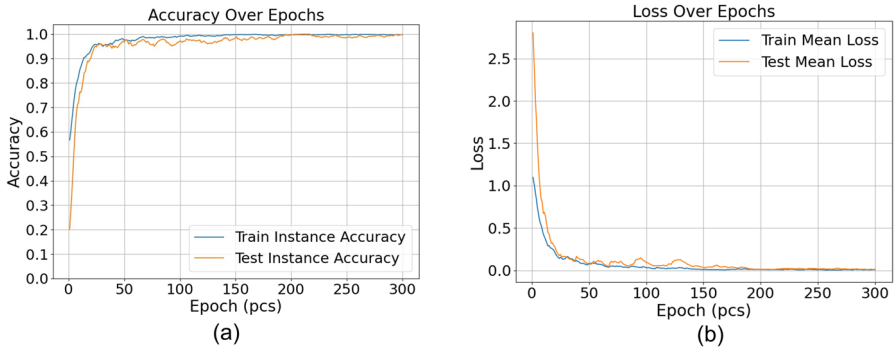


**Fig. 6** **a** The accuracy rate of the training dataset. **b** The accuracy rate of the test dataset. **c** The relationship between the sampling point and accuracy rate. **d** The relationship between the sampling point and training time–cost

studies, we will further explore the relationship between processing time and performance and consider using more and higher-performance GPUs to accelerate the training process. In different applications, processing time is closely tied to hardware requirements. For real-time applications like autonomous driving, low latency and high inference efficiency are crucial. These scenarios require multiple high-end GPUs to ensure sufficient computational speed. In contrast, offline applications like medical image analysis focus more on processing large datasets, relying on high-capacity memory and fast storage devices. While offline tasks don't demand real-time response, reducing training time remains key to improving efficiency.

## 6 Object Classification Results

Based on the previous evaluation of the homologous data and the sampling points of the FPS method, we added the factor of two homologous data to the initial training dataset as the training dataset and used the point cloud obtained from the MMW system as the test dataset, with the number of sampling points set to 8192, and performed the training for 300 epochs. The training results are shown in Fig. 7a. The blue data in the figure represents the accuracy rate of the training dataset, which reaches 0.998 after 300 epochs. The orange data represents the accuracy rate of the test dataset, which reaches 0.996 after 300 epochs. These training results show that we have achieved a very high accuracy rate in model training.



**Fig. 7** Results of training with  $2\times$  homologous data and 8192 sampling points. **a** Accuracy over epochs. **b** Loss over epochs

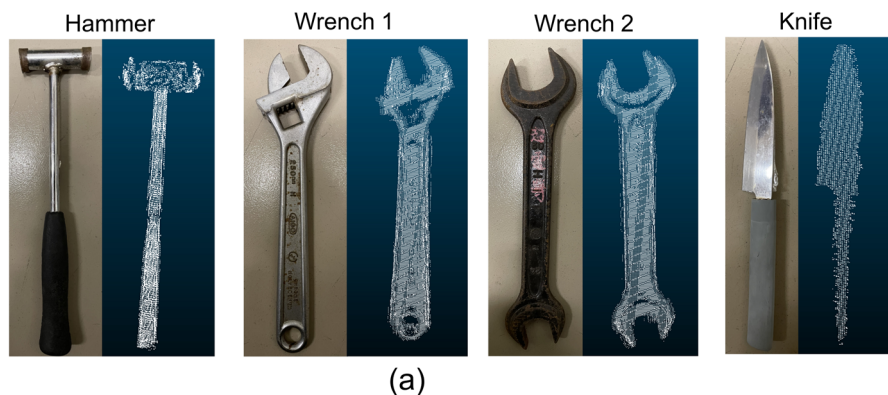
Additionally, the loss value is crucial in evaluating trained models, as it serves as the core metric for assessing predictive performance and guiding model optimization. In the PointNet++ classification model, cross-entropy is used as the loss function, represented as the average cross-entropy loss across all samples [27]. When the predicted and actual classes differ significantly, the loss value is high; conversely, the loss value is low when the predicted and actual classes are similar [41]. By calculating the loss function value, the model can quantify the difference between the current predictions and the true values. As shown in Fig. 7b, the blue line represents the train mean loss, while the orange line represents the test mean loss. The train mean loss rapidly decreases and converges within the first 30 epochs, ultimately approaching zero, indicating that the model is highly fitted to the training data. The test mean loss also decreases rapidly within the first 40 epochs, then fluctuates slightly, but eventually approaches zero, demonstrating good generalization performance on the test data. Moreover, the train loss and test loss run almost parallel, with the test loss slightly higher than the train loss, indicating consistent performance between training and testing data and no significant overfitting. In summary, the trained model exhibits ideal performance on both training and test data, showcasing strong learning capability and generalization ability.

In order to evaluate the trained model, each sample in Fig. 2b is transformed by scaling down by a factor of 0.95 and shifting to the right by 0.5 cm as an evaluation dataset input to the trained model. The data used for evaluation are point clouds of 11 hammers, 11 knives, 12 pistols, 13 scissors, and 12 wrenches. The prediction results are shown in Fig. 8. For the available samples, the trained model classified all objects correctly, which conforms with a high accuracy rate of 99.6%.

To ensure that the trained models have good generalization ability, further evaluation and validation of the trained models were performed. A sample not included in the dataset was used to evaluate the trained model. The samples are shown in Fig. 9a, with a hammer, two wrenches, and a knife. The right side of Fig. 9a shows the actual image of the samples, and the left side shows the point cloud obtained from the MMW imaging processing. The point clouds of the four objects are input

Predicted Actual	Hammer	Knife	Pistol	Scissors	Wrench
Hammer	11	0	0	0	0
Knife	0	11	0	0	0
Pistol	0	0	12	0	0
Scissors	0	0	0	13	0
Wrench	0	0	0	0	12

**Fig. 8** Prediction results with the evaluation dataset



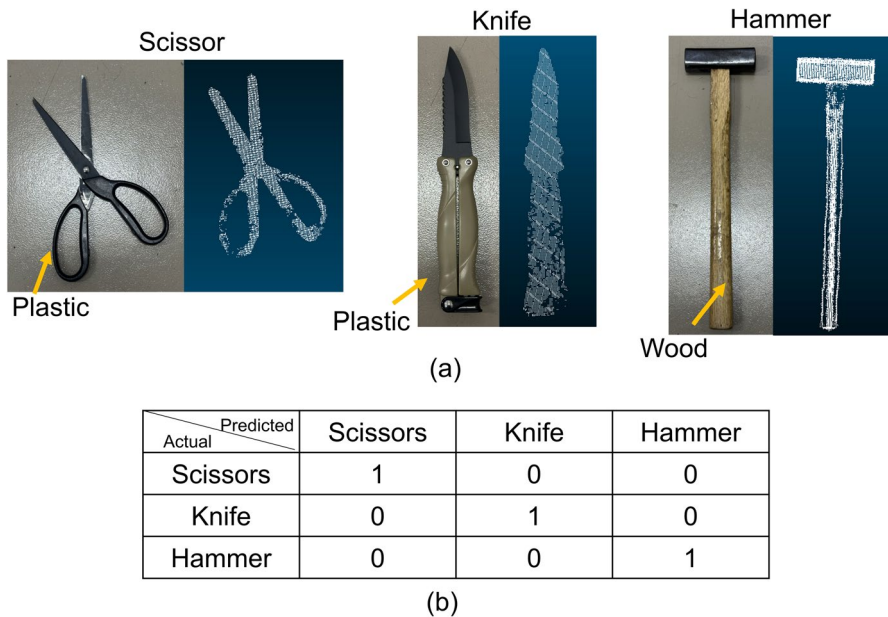
Predicted Actual	Hammer	Wrench 1	Wrench 2	Knife
Hammer	1	0	0	0
Wrench 1	0	1	0	0
Wrench 2	0	0	1	0
Knife	0	0	0	1

(b)

**Fig. 9** **a** Actual image and point cloud data of the samples. **b** Prediction results

into the trained model, and the results obtained are shown in Fig. 9b, where all the samples are classified correctly.

Considering the practical application that objects containing plastic or wood will not be sprayed with metallic paint, we evaluated objects that were not sprayed with metallic paint and also were not included in the dataset. The sample, shown in Fig. 10a, contains scissors with a plastic handle, a knife with a plastic handle, and a hammer with a wooden handle. The right side of Fig. 10a shows the actual image of the sample, and the left side shows the point cloud obtained after the MMW imaging reconstruction process. From the point cloud image, it can be seen



**Fig. 10** **a** Actual image and point cloud data of the samples. **b** Prediction results

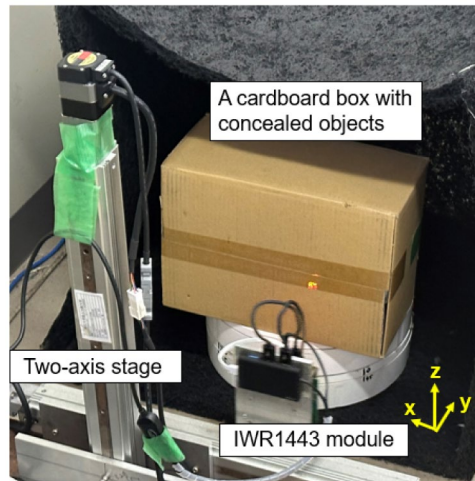
that the plastic or wood parts exhibit hollow areas, attributed to the high penetration of MMW to objects made of these materials, which are weak in reflecting signals. The point clouds of the three objects were input into the trained model, and the results obtained are shown in Fig. 10b. Even though the point cloud data have hollow defects in some parts, the samples are all correctly classified, which proves that the trained model has the potential for the classification of non-metallic objects in practical applications.

Finally, in order to verify the feasibility of this study in practical applications, a cardboard box with concealed objects was used as a sample to be scanned by an MMW imaging system, the composition of which is shown in Fig. 11. The IWR1443 module was attached to a 2-axis mechanical stage for X–Z scanning purposes. The duration of the chirp signal in the 77–81 GHz range is 40  $\mu$ s. The x-axis scanning speed is 200mm/s and the time taken to scan an area of 500mm x 500mm is approximately 4 min.

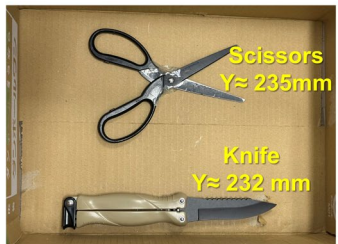
Due to the limitations of the scanning area, the five samples were divided into two sets and scanned in separate cardboard boxes. It is worth noting that of these 5 samples, the hammer, scissors, wrench, and knife are items not included in the dataset, and due to the limited number of samples of pistols, the pistol used in this concealed object identification was the sample included in the dataset.

The first set of concealed objects were a wooden-handled hammer, a plastic-handled knife, and plastic-handled scissors. The position of the hammer at the top layer is at approximately Y = 160 mm, as shown in Fig. 12a, and the knife and scissors at the bottom layer are located at approximately Y = 232 and 235 mm, respectively, as shown in Fig. 12b. Notably, the hammer and the two objects

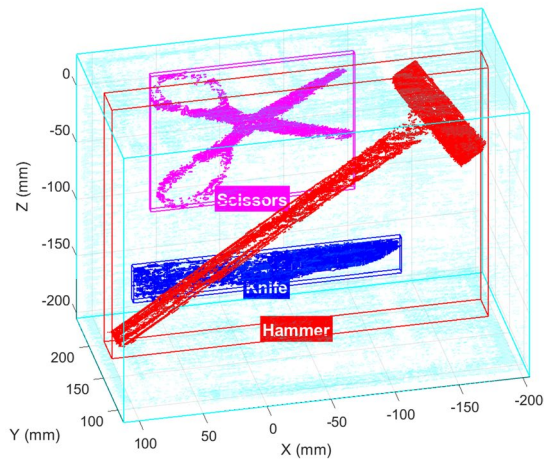
**Fig. 11** The composition of the MMW imaging system



(a)



(b)



(c)

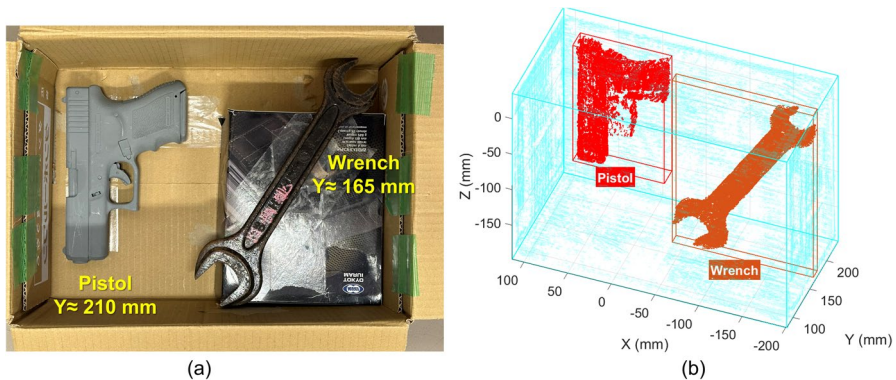
**Fig. 12** a Top layer of the cardboard box. b Bottom layer of the cardboard box. c Reconstruction and object classification results

at the bottom layer overlapped with each other. In Fig. 12c, the reconstruction results from four-sided scanning using our previously developed MMW point cloud reconstruction algorithm are illustrated. The light blue part is the perimeter of the cardboard box, the pink part is the scissors, the dark blue part is the knife, and the red part is the hammer. The results show that even though the objects overlap each other, the reconstruction results of four-sided scanning can still get the rough outline of each object, which is unachievable with 2D imaging. Due to



the fact that some of the canceled objects are made of non-metallic materials such as wood and plastic, the reconstructed contours of the objects appear uneven and hollow. Since the point cloud has been segmented by clustering in the first step of data processing during single-sided scanning, the box and the inner objects can be segmented directly in the result of four-sided scanning. Subsequently, the segmented objects were input into the trained model. It is worth noting that it took a total of 16 min to scan all four sides of the cardboard, and it only about 4 s to input the segmented point cloud into the training model to obtain the prediction results. Figure 12c presents the object classification results and the [supplementary](#) video showcases rotated 3D results, wherein the concealed objects are boxed, and the object classification results are displayed beneath the box. Notably, all classification outcomes are accurate.

The second set of concealed objects are the pistol and the wrench. The position of the wrench on the top layer is approximately at  $Y = 165$  mm, and the pistol on the bottom layer is located approximately at  $Y = 210$  mm, as shown in Fig. 13a. After scanning the six sides of the cardboard box, Fig. 13b shows the reconstruction results and the [supplementary](#) video showcases rotated 3D results. The light blue part is the perimeter of the cardboard box, the red part is the pistol, and the orange part is the wrench. The segmented objects inside the cardboard box are input into the trained model. Figure 13b showcases the results of object classification, where the concealed objects are boxed, and the object classification outcomes are displayed beneath the box, all of which are correctly classified. These outcomes demonstrate the robust performance of our trained model for object classification using MMW point cloud data, which often contain missing, hollow, noisy, and overlapping features. The efficacy of this method in practical applications has been successfully validated.



**Fig. 13** **a** Internal view of the cardboard box. **b** Reconstruction and object classification results

## 7 Conclusion

This study successfully applied AI technology in the field of MMW imaging, effectively using the PointNet++ method to classify high-precision point cloud data obtained from MMW radar into five different object categories. The utilized dataset includes point cloud data generated through the transformation of 3D models and high-precision reconstruction of physical objects using MMW radar scanning. Through data augmentation techniques on the training and test datasets, including generating homologous data and optimizing the number of sampling points, we enhanced the accuracy in classifying five categories of objects. Considering evaluation results and hardware limitations, we augmented the initial training set with an addition in  $2\times$  homologous data for training, using point clouds obtained by the MMW imaging system as the test dataset, with 8192 sampled points, and trained for 300 epochs. The training dataset achieved an accuracy of 0.998, while the test dataset achieved an accuracy of 0.996.

Moreover, to further validate the potential of the trained model for practical applications, samples not included in the dataset were evaluated. The experiment results demonstrated exceptional performance in dealing with complex situations, including new samples and non-metallic materials, particularly in the classification of concealed objects inside the cardboard box for multi-surface scanning reconstruction. These findings highlight the effectiveness and feasibility of our method in classifying complex point cloud data from MMW imaging systems.

In conclusion, this study is significant in applying the AI technique to MMW imaging successfully and realizing the efficient object classification of complex point cloud data, which offers strong support for the practical application of MMW imaging technology. In future research, it is possible to enhance training speed by improving hardware or optimizing neural network architectures and to adapt to a wider range of complex applications by expanding the variety of datasets.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s10762-024-01017-5>.

**Authors Contributions** Conceptualization, Y.W., J.S., H.M, and M.T.; methodology, Y.W. and J.S.; software, Y.W. and J.S.; validation, Y.W., J.S., H.M, and M.T.; formal analysis, Y.W.; investigation, Y.W. and J.S.; experimental data acquisition, Y.W. and J.S.; data curation, Y.W.; writing—original draft preparation, Y.W.; writing—review and editing, Y.W., H.M, and M.T. All authors have read and agreed to the published version of the manuscript.

**Funding** Open Access funding provided by Osaka University. No funding was supported for this work.

**Data Availability** No datasets were generated or analysed during the current study.

## Declarations

**Ethical Approval** Not applicable.

**Consent to Participate** The authors consent to participate.

**Consent to Publish** The authors consent to publish.



**Competing Interests** The authors declare no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Hasch, J., Topak, E., Schnabel, R., Zwick, T., Weigel, R., Waldschmidt, C.: Millimeter-Wave Technology for Automotive Radar Sensors in the 77 GHz Frequency Band. *IEEE Transactions on Microwave Theory and Techniques*. 60, 845–860 (2012). <https://doi.org/10.1109/tmmt.2011.2178427>
2. Patole, S.M., Torlak, M., Wang, D., Ali, M.: Automotive radars: A review of signal processing techniques. *IEEE Signal Processing Magazine*. 34, 22–35 (2017). <https://doi.org/10.1109/msp.2016.2628914>
3. Zhou, T., Yang, M., Jiang, K., Wong, H., Yang, D.: MMW Radar-Based Technologies in Autonomous Driving: A Review. *Sensors*. 20, 7283 (2020). <https://doi.org/10.3390/s20247283>
4. Alizadeh, M., Shaker, G., Almeida, J.C.M.D., Morita, P.P., Safavi-Naeini, S.: Remote Monitoring of Human Vital Signs Using mm-Wave FMCW Radar. *IEEE Access*. 7, 54958–54968 (2019). <https://doi.org/10.1109/access.2019.2912956>
5. Yang, Z., Pathak, P.H., Zeng, Y., Xixi Liran, Mohapatra, P.: Monitoring vital signs using millimeter wave. *Proceedings of the 17th ACM International Symposium on Mobile Ad Hoc Networking and Computing*. (2016). <https://doi.org/10.1145/2942358.2942381>
6. Chao, L., Afsar, M.N., Korolev, K.A.: Millimeter wave dielectric spectroscopy and breast cancer imaging. 2012 7th European Microwave Integrated Circuit Conference. 572–575 (2012)
7. Gezimati, M., Singh, G.: Advances in terahertz technology for cancer detection applications. *Applied Physics B*. 55, (2022).
8. Nagatsuma, T.: Millimeter-wave Imaging and Its Application to Structure Diagnosis. *NTT Technical Journal*. 18, 25–28 (2006)
9. Oka, S., Togo, H., Kukutsu, N., Nagatsuma, T.: LATEST TRENDS IN MILLIMETER-WAVE IMAGING TECHNOLOGY. *Progress In Electromagnetics Research Letters*. 1, 197–204 (2008). <https://doi.org/10.2528/pier107120604>
10. Appleby, R., Anderton, R.N.: Millimeter-Wave and Submillimeter-Wave Imaging for Security and Surveillance. *Proceedings of the IEEE*. 95, 1683–1690 (2007). <https://doi.org/10.1109/jproc.2007.898832>
11. Hirata, A., Koji Suizu, Sudo, Y., Watanabe, I., Sekine, N., Akifumi Kasamatsu: Non-destructive Inspection of Concrete Surface Crack Using Near-Field Scattering. 2020 IEEE International Symposium on Radio-Frequency Integration Technology (RFIT). 244–246 (2020). <https://doi.org/10.1109/rfit49453.2020.9226171>
12. Kenneth, K.O., Choi, W., Han, R.: Perspective on active submillimeter electromagnetic wave imaging using CMOS integrated circuits technologies. *Journal of Applied Physics*. 133, (2023). <https://doi.org/10.1063/5.0143622>
13. Chopard, A., Guillet, J.-P., Gellie, P., Benoit RECUR, H Balacey, Mounaix, P.: Skeletonization and 3D Rendering with real time Terahertz tomography. *Optics Continuum*. 2, (2023). <https://doi.org/10.1364/optcon.486227>
14. Wang, Y., Yi, L., Tonouchi, M., Nagatsuma, T.: High-Speed 600 GHz-Band Terahertz Imaging Scanner System with Enhanced Focal Depth. *Photonics*. 9, 913 (2022). <https://doi.org/10.3390/photonics9120913>

15. Yi, L., Li, Y., Tadao Nagatsuma: Photonic Radar for 3D Imaging: From Millimeter to Terahertz Waves. *IEEE Journal of Selected Topics in Quantum Electronics*. 29, 1–14 (2023). <https://doi.org/10.1109/jstqe.2023.3298933>
16. Arttu Luukanen, Appleby, R., Kemp, M., Salmon, N.A.: *Millimeter-Wave and Terahertz Imaging in Security Applications*. Springer Series in Optical Sciences. 171, 491–520 (2012). [https://doi.org/10.1007/978-3-642-29564-5\\_19](https://doi.org/10.1007/978-3-642-29564-5_19)
17. Zhuge, X., Yarovoy, A.G.: A Sparse Aperture MIMO-SAR-Based UWB Imaging System for Concealed Weapon Detection. *IEEE Transactions on Geoscience and Remote Sensing*. 49, 509–518 (2011). <https://doi.org/10.1109/tgrs.2010.2053038>
18. Sheen, D.M., McMakin, D.L., Hall, T.E.: Three-dimensional millimeter-wave imaging for concealed weapon detection. *IEEE Transactions on Microwave Theory and Techniques*. 49, 1581–1592 (2001). <https://doi.org/10.1109/22.942570>
19. Soumya, A., Krishna Mohan, C., Cenkeramaddi, L.R.: Recent Advances in mmWave-Radar-Based Sensing, Its Applications, and Machine Learning Techniques: A Review. *Sensors*. 23, 8901 (2023). <https://doi.org/10.3390/s23218901>
20. Guan, H., Yu, Y., Ji, Z., Li, J., Zhang, Q.: Deep learning-based tree classification using mobile LiDAR data. *Remote Sensing Letters*. 6, 864–873 (2015). <https://doi.org/10.1080/2150704x.2015.1088668>
21. López-Tapia, S., Molina, R., Pérez de la Blanca, N.: Using machine learning to detect and localize concealed objects in passive millimeter-wave images. *Engineering Applications of Artificial Intelligence*. 67, 81–90 (2018). <https://doi.org/10.1016/j.engappai.2017.09.005>
22. Meng, Z., Zhang, M., Wang, H.: CNN with Pose Segmentation for Suspicious Object Detection in MMW Security Images. *Sensors*. 20, 4974 (2020). <https://doi.org/10.3390/s20174974>
23. Yanik, M.E., Wang, D., Torlak, M.: Development and Demonstration of MIMO-SAR mmWave Imaging Testbeds. *IEEE Access*. 8, 126019–126038 (2020). <https://doi.org/10.1109/access.2020.3007877>
24. Bai, J., Zheng, L., Li, S., Tan, B., Chen, S., Huang, L.: Radar Transformer: An Object Classification Network Based on 4D MMW Imaging Radar. *Sensors*. 21, 3854 (2021). <https://doi.org/10.3390/s21113854>
25. Wang, Y., Su, J., Fukuda, T., Masayoshi Tonouchi, Murakami, H.: Precise 2D and 3D Fluoroscopic Imaging by Using an FMCW Millimeter-Wave Radar. *IEEE Access*. 11, 84027–84034 (2023). <https://doi.org/10.1109/access.2023.3302889>
26. Qi, C.R., Su, H., Mo, K., Guibas, Leonidas J: PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. <https://arxiv.org/abs/1612.00593>
27. Qi, C.R., Yi, L., Su, H., Guibas, Leonidas J: PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space. *Advances in neural information processing systems*. 30, (2017)
28. Phan, A.V., Nguyen, M.L., Nguyen, Y.L.H., Bui, L.T.: DGCNN: A convolutional neural network over large-scale labeled graphs. *Neural Networks*. 108, 533–543 (2018). <https://doi.org/10.1016/j.neunet.2018.09.001>
29. Li, Y., Bu, R., Sun, M., Wu, W., Di, X., Chen, B.: PointCNN: Convolution On  $\mathcal{X}$ -Transformed Points. *arXiv (Cornell University)*. (2018). <https://doi.org/10.48550/arxiv.1801.07791>
30. Maturana, D., Scherer, S.: VoxNet: A 3D Convolutional Neural Network for real-time object recognition. 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). (2015). <https://doi.org/10.1109/iros.2015.7353481>
31. Riegler, G., Ali Hakan Ulusoy, Geiger, A.: OctNet: Learning Deep 3D Representations at High Resolutions. *Computer Vision and Pattern Recognition*. (2017). <https://doi.org/10.1109/cvpr.2017.701>
32. Zotti, C., Luo, Z., Humbert, O., Lalande, A., Jodoin, P.-M.: GridNet with Automatic Shape Prior Registration for Automatic MRI Cardiac Segmentation. *Lecture notes in computer science*. 73–81 (2018). [https://doi.org/10.1007/978-3-319-75541-0\\_8](https://doi.org/10.1007/978-3-319-75541-0_8)
33. Qi, C.R., Liu, W., Wu, C., Su, H., Guibas, L.J.: Frustum PointNets for 3D Object Detection from RGB-D Data. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. (2018). <https://doi.org/10.1109/cvpr.2018.00102>
34. Jiang, J., Bao, Z., Chen, Z., Zhao, X., Gao, Y.: MLCNN: Multi-Loop-View Convolutional Neural Network for 3D Shape Retrieval. *Proceedings of the AAAI Conference on Artificial Intelligence*. 33, 8513–8520 (2019). <https://doi.org/10.1609/aaai.v33i01.33018513>
35. Chang, A.X., Funkhouser, T., Guibas, L., Hanrahan, P., Huang, Q., Li, Z., Savarese, S., Savva, M., Song, S., Su, H., Xiao, J., Yi, L., Yu, F.: ShapeNet: An Information-Rich 3D Model Repository. *ArXiv*. (2015)

36. Wu, Z., Song, S., Khosla, A., Fisher, Y., Zhang, L., Tang, X., Xiao, J.: 3D ShapeNets: A Deep Representation for Volumetric Shapes. *Proceedings of 28th IEEE Conference on Computer Vision and Pattern Recognition*. (2015)
37. Lian, H., Sun, P., Meng, Z., Li, S., Wang, P., Qu, Y.: LIDAR Point Cloud Augmentation for Dusty Weather Based on a Physical Simulation. *Mathematics*. 12, 141–141 (2023). <https://doi.org/10.3390/math12010141>
38. Chlap, P., Min, H., Vandenberg, N., Dowling, J., Holloway, L., Haworth, A.: A review of medical image data augmentation techniques for deep learning applications. *Journal of Medical Imaging and Radiation Oncology*. 65, 545–563 (2021). <https://doi.org/10.1111/1754-9485.13261>
39. Nalepa, J., Marcinkiewicz, M., Kawulok, M.: Data Augmentation for Brain-Tumor Segmentation: A Review. *Frontiers in Computational Neuroscience*. 13, (2019). <https://doi.org/10.3389/fncom.2019.00083>
40. Mikołajczyk, A., Grochowski, M.: Data augmentation for improving deep learning in image classification problem. *IEEE Xplore*. (2018). <https://ieeexplore.ieee.org/document/8388338>
41. Mao, A., Mohri, M., Zhong, Y.: Cross-Entropy Loss Functions: Theoretical Analysis and Applications. *Proceedings of Machine Learning Research*. 202, 23803–23828 (2023).

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.