



OPEN ACCESS

EDITED BY

Zhiqiang Huo,
King's College London, United Kingdom

REVIEWED BY

Yanchao Mao,
Zhengzhou University, China
Jesya Janet Kumari J.,
The Oxford College of Engineering, India
Chonglun Guo,
Third Affiliated Hospital of Nanchang
University, China
Roi Yozevitch,
Ariel University, Israel
Prithwijit Mukherjee,
University of Calcutta, India

*CORRESPONDENCE

Zhizhong Xing
✉ xingzhizhong@kmmu.edu.cn
Gengfeng Zheng
✉ hcomet@163.com
Huidong Wu
✉ wuhuidong@kmmu.edu.cn

RECEIVED 11 December 2024

ACCEPTED 08 April 2025

PUBLISHED 02 May 2025

CITATION

Xing Z, Meng Z, Zheng G, Ma G, Yang L,
Guo X, Tan L, Jiang Y and Wu H (2025)
Intelligent rehabilitation in an aging
population: empowering human-machine
interaction for hand function rehabilitation
through 3D deep learning and point cloud.
Front. Comput. Neurosci. 19:1543643.
doi: 10.3389/fncom.2025.1543643

COPYRIGHT

© 2025 Xing, Meng, Zheng, Ma, Yang, Guo,
Tan, Jiang and Wu. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License](#)
(CC BY). The use, distribution or reproduction
in other forums is permitted, provided the
original author(s) and the copyright owner(s)
are credited and that the original publication
in this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Intelligent rehabilitation in an aging population: empowering human-machine interaction for hand function rehabilitation through 3D deep learning and point cloud

Zhizhong Xing^{1,2,3*}, Zhijun Meng², Gengfeng Zheng^{3*},
Guolan Ma⁴, Lin Yang^{5,6}, Xiaojun Guo⁷, Li Tan¹, Yuanqiu Jiang¹
and Huidong Wu^{1*}

¹School of Rehabilitation, Kunming Medical University, Kunming, China, ²Medical Imaging Key Laboratory of Sichuan Province, North Sichuan Medical College, Nanchong, China, ³Fujian Key Laboratory of Special Intelligent Equipment Safety Measurement and Control, Fujian Special Equipment Inspection and Research Institute, Fuzhou, China, ⁴School of Pharmaceutical Science and Yunnan Key Laboratory of Pharmacology for Natural Products, Kunming Medical University, Kunming, China, ⁵School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an, China, ⁶Xi'an Aerospace Automation Co., Ltd, Xi'an, China, ⁷School of Mechanical and Automotive Engineering, South China University of Technology, Guangzhou, China

Human-machine interaction and computational neuroscience have brought unprecedented application prospects to the field of medical rehabilitation, especially for the elderly population, where the decline and recovery of hand function have become a significant concern. Responding to the special needs under the context of normalized epidemic prevention and control and the aging trend of the population, this research proposes a method based on a 3D deep learning model to process laser sensor point cloud data, aiming to achieve non-contact gesture surface feature analysis for application in the field of intelligent rehabilitation of human-machine interaction hand functions. By integrating key technologies such as the collection of hand surface point clouds, local feature extraction, and abstraction and enhancement of dimensional information, this research has constructed an accurate gesture surface feature analysis system. In terms of experimental results, this research validated the superior performance of the proposed model in recognizing hand surface point clouds, with an average accuracy of 88.72%. The research findings are of significant importance for promoting the development of non-contact intelligent rehabilitation technology for hand functions and enhancing the safe and comfortable interaction methods for the elderly and rehabilitation patients.

KEYWORDS

3D perception, neural network, human-machine interaction, deep learning, non-contact rehabilitation

1 Introduction

In the context of the ongoing normalization of global epidemic prevention and control measures and the accelerating aging of the population, the dual challenge of safeguarding public health and enhancing the well-being of the elderly and rehabilitation patients has emerged as a pressing issue in contemporary society (Huang et al., 2023; Baraković et al., 2024; Lopes et al., 2023; Harada et al., 2021; Li L. et al., 2023). Especially in the field of intelligent rehabilitation of hand functions in human-machine interaction, traditional rehabilitation methods are limited by contact based operations and infection risks, making it difficult to meet current needs (An et al., 2022; Gao et al., 2023; Jiang et al., 2021; Zhang et al., 2023; Samhan et al., 2020; Tang et al., 2025). Therefore, developing a non-contact, efficient, and accurate gesture recognition system is of great significance for promoting the development of intelligent rehabilitation technology for human-machine interaction hand function.

There are many limitations in the field of hand function rehabilitation (Gu et al., 2022; Edger-Lacoursière et al., 2023; Zestas and Tselikas, 2023; Ergen et al., 2024; Bates and Sunderam, 2023). Traditional hand function rehabilitation methods are mostly contact based, such as therapists manually assisting patients with hand movement training. This approach relies on the manpower input of professional therapists, and in the face of the growing population of elderly rehabilitation patients, treatment resources are often difficult to meet the demand. In addition, existing rehabilitation assessment methods largely rely on the subjective judgment of therapists and lack objective and accurate data support. When evaluating the accuracy and flexibility of hand function, there may be differences in the evaluation criteria of different therapists, resulting in lower reliability of the evaluation results. The non-contact gesture recognition method proposed in this research avoids direct contact between patients and treatment equipment or therapists. At the same time, it can achieve intelligent gesture analysis and recognition, reduce reliance on professional therapists' manpower, improve the efficiency of rehabilitation treatment, and better meet the needs of elderly rehabilitation patients. By analyzing the hand surface point cloud, our proposed method can extract rich hand features and provide accurate data support for rehabilitation assessment.

Different sensors can be used to obtain hand surface information. LEAP device plays an important role in fields such as human-machine interaction and hand motion capture (Vaitkevičius et al., 2019; Ameer et al., 2020; Najafinejad and Korayem, 2023; Li et al., 2019; Galván-Ruiz et al., 2020). The LEAP device mainly utilizes optical imaging principles to obtain 3D coordinate information of the hand, with performance advantages of high frame rate and high accuracy, and can track subtle hand movements in real time. In the application of virtual reality, game development, and scientific research, LEAP equipment played a crucial role. LEAP device performs well in indoor environments with stable lighting and close range operations. What's more, MediaPipe, as a powerful and widely used open-source framework, has achieved remarkable results in the field of pose detection (Suwabe et al., 2024; Cao et al., 2024; Samaan et al., 2022; Mariappan et al., 2024). With excellent performance, it can accurately recognize hand movements and postures, and has outstanding application performance in many scenarios. MediaPipe combines various advanced algorithms, greatly promoting the

development of computer vision and human-machine interaction fields, and providing important support for related research and applications. In addition, laser sensors, as a high-precision, non-contact measurement tool, have been widely used in the fields of material surface morphology analysis, defect detection, and so on (Zheng et al., 2023; Ye et al., 2023; Rufe et al., 2022; Sadaoui et al., 2022; Theodose et al., 2021). The point cloud data collected by laser sensors can obtain 3D structural information of material surfaces, providing important basis for material performance analysis and interface engineering. Meanwhile, laser sensor point clouds, as a non-contact gesture acquisition technology, can obtain 3D information of the hand with high precision and high spatiotemporal resolution, providing a rich data source for gesture analysis. However, in the field of human-machine interaction and intelligent rehabilitation of hand functions, the application of laser sensor point cloud data is still relatively limited. Therefore, this research aims to apply laser sensor point cloud data to gesture recognition, process point cloud data through a 3D deep learning model, and achieve non-contact gesture surface feature analysis.

There are various methods in the field of point cloud processing, among which the point cloud processing method based on multi view method is widely used (Zhang et al., 2018; Tong et al., 2020). This method projects three-dimensional point cloud data onto multiple two-dimensional views, and then uses traditional two-dimensional image processing techniques for analysis and processing. Its advantage is that it can use mature two-dimensional image processing algorithms to reduce processing difficulty. However, this method inevitably loses some three-dimensional spatial information during the projection process, resulting in damage to the integrity of point cloud data and limiting subsequent analysis. The point cloud processing based on voxel method divides the three-dimensional space into regular voxel grids (Li J. et al., 2021; Li Y. et al., 2021; Aljumaily et al., 2023), which fills the point cloud data into corresponding voxels, and then uses a three-dimensional convolutional neural network for processing. This method can effectively preserve the spatial structure information of point cloud data. However, high-resolution voxels can lead to an exponential increase in computational complexity, while low resolution voxels cannot accurately express the detailed features of point clouds. PointNet and PointNet++ break through the limitations of point cloud processing methods (Qi et al., 2017a; Qi et al., 2017b). As deep learning models capable of directly processing point cloud data, they break the traditional convolutional neural network's dependence on structured data. PointNet and PointNet++ can effectively extract global features of point clouds by introducing symmetric functions, demonstrating good performance in tasks such as point cloud classification and object detection.

The continuous emergence of efficient solutions for upper limb motor performance evaluation, hand rehabilitation, human-computer interaction, and other excellent achievements (Guo et al., 2023; Yozevitch et al., 2023; Zhu et al., 2025; Lu et al., 2024). We have achieved a lot of research results based on graph neural network in the early stage (Xing et al., 2023; Xing et al., 2025; Xing et al., 2022), but non-contact gesture surface feature analysis for intelligent rehabilitation of hand function in human-machine interaction is still blank. Moreover, due to the complexity and diversity of gesture surface features, how to effectively extract and analyze gesture surface features from point cloud data has become a challenging problem.

In summary, this research combines laser sensor point cloud data with 3D deep learning model to construct an efficient and accurate gesture recognition system in the context of normalized epidemic prevention and control and aging population trends. Through this system, we can achieve precise analysis of the surface features of gestures, providing new solutions for intelligent rehabilitation of hand functions in human-machine interaction. This research achievement can provide a non-contact gesture analysis tool for rehabilitation medical institutions and elderly care institutions, reduce the risk of infection transmission, improve the efficiency and accuracy of rehabilitation treatment, and promote the cross integration of interface science and deep learning technology, opening up new paths for research in related fields. In addition, this method can also be applied in fields such as augmented reality and gesture control interfaces, providing users with a more natural and intuitive way of human-machine interaction.

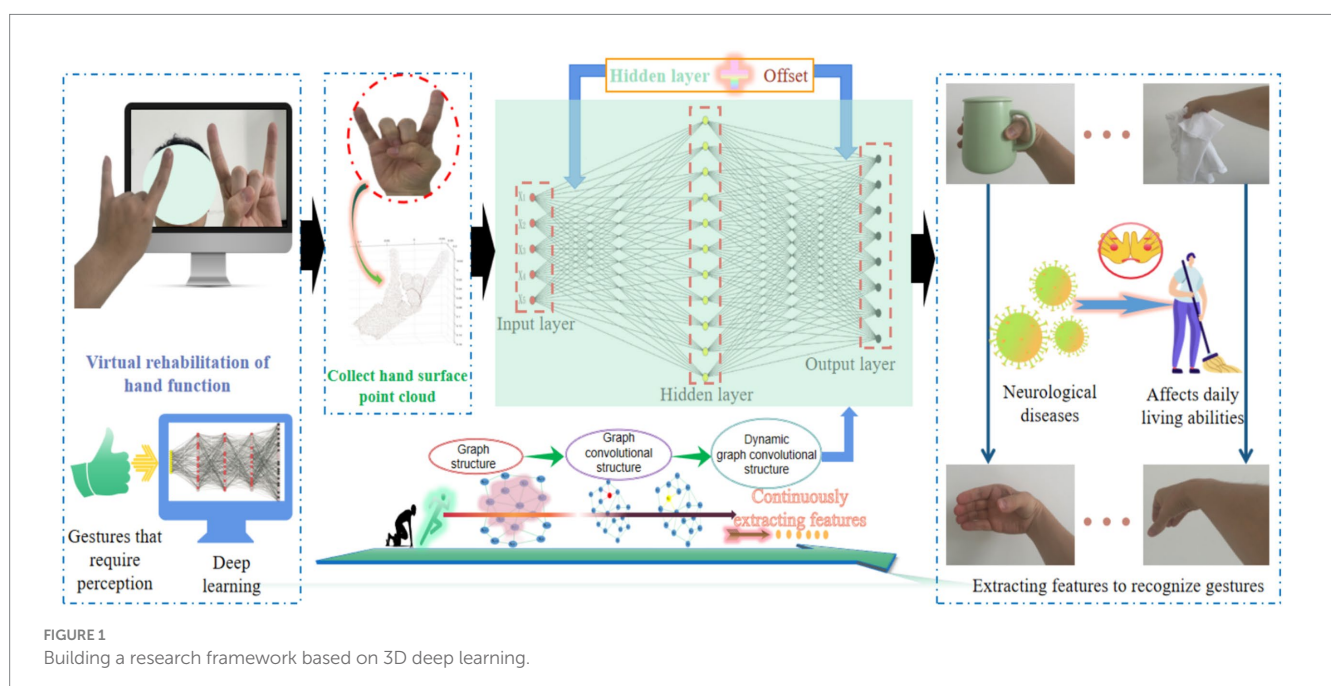
2 Materials and methods

Gesture recognition is a crucial step in achieving natural human-machine interaction in hand function virtual rehabilitation systems (Li W. et al., 2023; Liu et al., 2022; Wang et al., 2023; Zhu et al., 2021). Due to the high degree of freedom of human hand joints, achieving precise and robust breakthroughs in gesture recognition technology is currently a challenging task. A point cloud refers to a collection of points on the surface characteristics of a target object, which have 3D coordinates and simple expression methods (Tan et al., 2022; Xiao et al., 2023; Akhtar et al., 2022), making it easy to express spatial features of different gestures in a digital form. More importantly, point cloud data is generally obtained through non-contact measurement devices such as laser sensors. This research builds a gesture surface feature analysis network (GSFAN) based on DGCNN to revolutionize the research approach of gesture recognition in 3D space (Wang et al., 2019), as shown in Figure 1.

2.1 Hand 3D data

The collection of hand surface point clouds is the primary step in this research. By collecting point cloud representations of hand surface data, it provides a foundation for subsequent analysis of hand surface features and functional rehabilitation. In order to obtain 3D information of the hands, we used light laser detection and ranging as a data acquisition device to obtain point cloud data of the hands. In the process of collecting point clouds on the surface of the hand, choose an environment with sufficient light and no obstructions, while ensuring that the background of the hand collection area forms a clear contrast with the hand, in order to better extract the features of the hand. In this research, we collected point clouds on the surface of the hands of different individuals to ensure data diversity. When collecting hand surface point clouds, different individuals displayed different postures in order according to regulations, as shown in Figure 2.

The total number of participants is 4. The distance between the sensor and the hand is not fixed, which facilitates us to quickly and conveniently collect hand surface point cloud. Although variable distances can increase noise, we can use CloudCompare to remove it. In the actual data collection process, the technology used by laser radar is Direct Time of Flight (DToF). The main reasons for choosing laser radar are as follows: Firstly, laser radar has the advantage of non-contact, which avoids the risk of cross infection that may be caused by contact devices. Secondly, laser radar has high-precision characteristics and can capture subtle structures of the hand. Finally, the anti-interference ability of laser radar enables stable acquisition of depth information even under complex lighting conditions. Therefore, laser radar plays a crucial role in the entire process of collecting hand surface point cloud. It can accurately capture hand surface features and generate high-quality point cloud data, which provides a solid data foundation for our research. This research sets each finger (thumb, index finger, middle finger, ring finger, little finger) as an independent variable, and each finger is in two states of



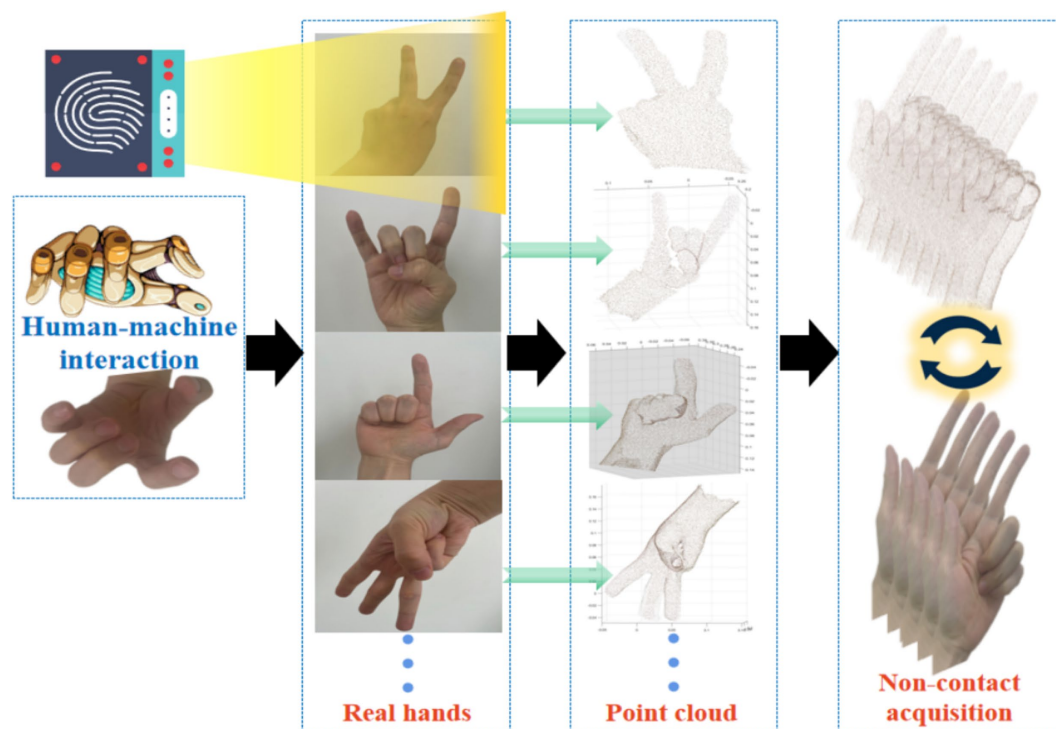


FIGURE 2
Collecting 3D data of hand surface for non-contact human-machine interaction.

extension and bending. A total of 32 gestures were calculated through combination.

After the data collection was completed, we adopted a series of steps to process the raw point cloud data to remove irrelevant information. Firstly, we will import the collected point cloud data into the professional point cloud processing software CloudCompare, which has powerful visualization and processing capabilities, allowing us to intuitively view the quality and distribution of point cloud data. Secondly, after importing the data, we checked if there is any interference from other objects around the hand surface point cloud. If there are non-hand surface point cloud, we used CloudCompare's segmentation tool to process them. Specifically, we combined the software visualization interface to accurately select the hand surface point cloud data to be preprocessed in the point cloud data list, ensuring accurate operation. Next, we used segmentation tools to draw a closed shape that fully includes the non-hand surface point cloud that needs to be deleted. We further utilized the segmentation tool sub option to separate non hand surface point cloud from the original point cloud data, and used the deletion function of the point cloud data list to completely delete the separated point clouds. Finally, after completing the above processing operations, we stored the processed hand surface point cloud data.

2.2 Hand surface point cloud edge convolution for human-machine interaction

In GSFAN, edge convolution is a key technique that can effectively capture local structures and geometric features in hand surface point

cloud data. The point cloud on the surface of the hand can be viewed as an unordered set of N points, thus it can be represented as an undirected graph, where each point is a node in the graph, and the edges between nodes represent their proximity (Figure 3a). In edge convolution, it is necessary to construct a graph structure to capture local proximity information of the hand surface point cloud. The commonly used methods for constructing graphs include different algorithms such as radius neighbors, which can determine the connection relationship based on the distance between each center point and its nearest neighbor point. Once the graph structure is constructed, edge convolution operations can be performed. The basic idea of edge convolution is to update the node representation by aggregating the features of neighboring nodes of each node. In the point cloud of the hand surface, each node can be represented as a vector containing position and other features. The edge convolution operation obtains a new representation of each node by aggregating and updating the features of its neighboring nodes (Figure 3b). GSFAN extracts higher-level feature representations by stacking multiple layers of edge convolution, gradually learning richer hand surface point cloud features. However, whether the more edge convolution layers, the better the effect, will be analyzed in detail and depth in this research.

The node set V contains all the points in the hand surface point cloud. Each point v_i corresponds to a coordinate point in three-dimensional space, and the set of node V can be expressed as Equation (1).

$$V = \{v_1, v_2, \dots, v_N\} \quad (1)$$

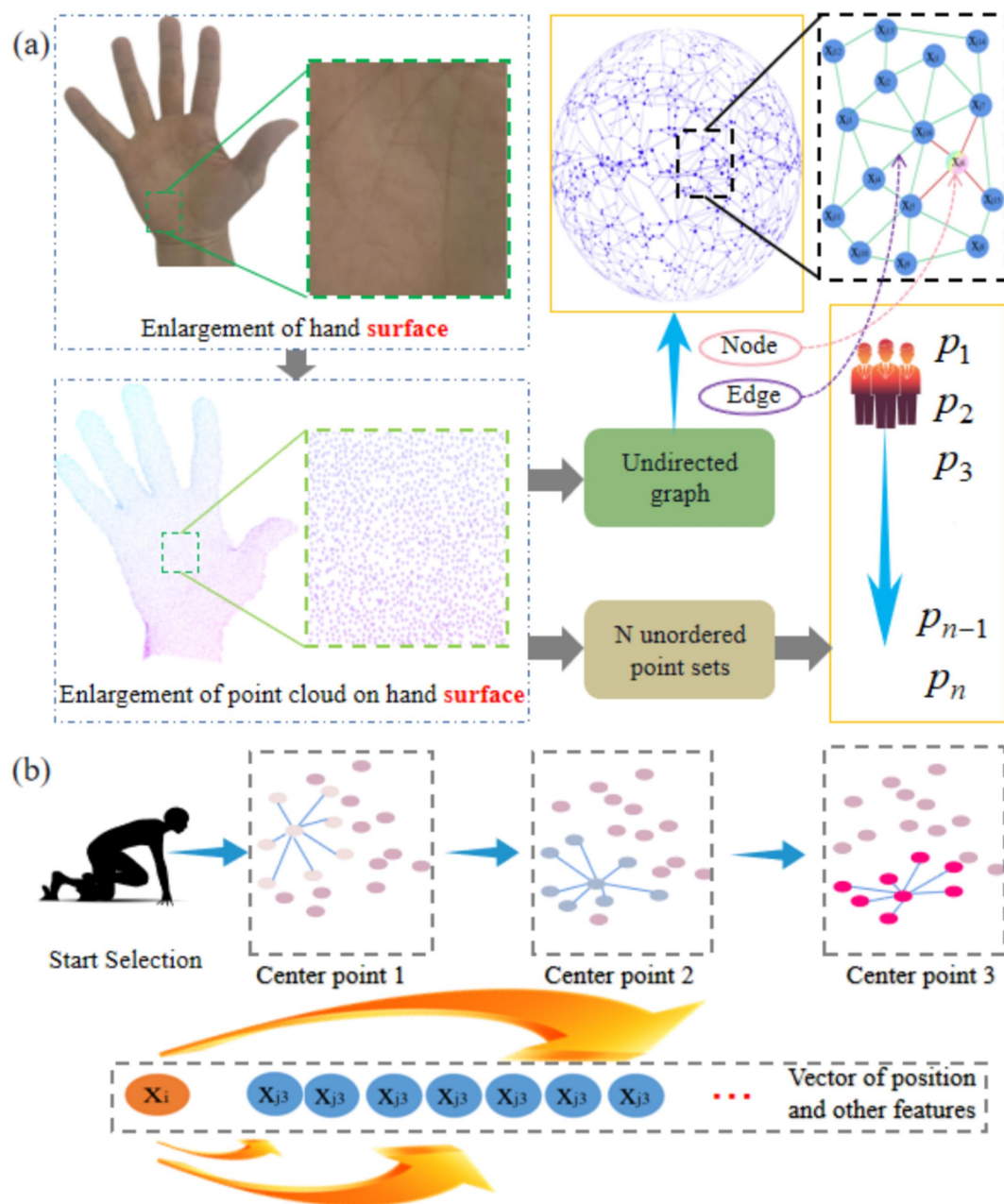


FIGURE 3

(a) Graph structure representation of real hand surface point clouds. (b) Each point in the hand surface point clouds is traversed as the center point.

Where N is the number of points in the point cloud on the surface of the hand. Each node v_i can be represented by a feature vector x_i , which contains the coordinates and other relevant information of the point, and can be expressed as Equation (2).

$$x_i = \{x_i^{coord}, x_i^{other}\} \quad (2)$$

Where x_i^{coord} is the three-dimensional coordinates of the point (for example, $x_i = [x, y, z]$), and x_i^{other} is other features such as color and normal.

The edge set E defines the connection relationship between nodes in the graph, and for hand surface point clouds, threshold distance is used to define edges. For any two nodes v_i and v_j , if the spatial distance between them is less than a certain threshold d_{thresh} , it can be expressed as Equation (3).

$$E = \{e_{ij} | dist(v_i, v_j) < d_{thresh}\} \quad (3)$$

Where $dist(v_i, v_j)$ is a function for calculating the distance between nodes v_i and v_j . In order to more conveniently represent the

graph structure, this research uses the adjacency matrix A , which can be expressed as Equation (4).

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1N} \\ A_{21} & A_{22} & \cdots & A_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ A_{N1} & A_{N2} & \cdots & A_{NN} \end{bmatrix} \quad (4)$$

If there is an edge between nodes v_i and v_j , $A_{ij} = 1$, otherwise $A_{ij} = 0$. In this research, we assume that there is a hand surface point cloud dataset of P , where each point p_i has a 3D coordinate (x_i, y_i, z_i) . For point p_i , its local density can be expressed as Equation (5).

$$p_i = \frac{1}{|N_k(p_i)|} \sum_{p_j \in N_k(p_i)} d(p_i, p_j) \quad (5)$$

Where $N_k(p_i)$ is the set of K -nearest neighbors of point p_i , and $d(p_i, p_j)$ is the Euclidean distance between points p_i and p_j . The neighborhood $N(p_i)$ of point p_i is defined as the set of all points up to p_i that do not exceed r_i , which can be expressed as Equation (6).

$$N(p_i) = \{p_j \in P \mid d(p_i, p_j) \leq r_i\} \quad (6)$$

Where r_i is the neighborhood radius of each point. We use the weight matrix W and the nonlinear activation function σ to obtain the aggregated feature h_i , which can be expressed as Equation (7).

$$h_i = \sigma \left(W \cdot \text{aggregate} \left(\{x_j \mid j \in N(p_i) \cup \{i\}\} \right) \right) \quad (7)$$

In order to enable each node to consider its own characteristics, we add self-loops and normalize them during the aggregation process, which can be expressed as Equation (8).

$$h_i = \sigma \left(W \cdot \frac{1}{|N(p_i) \cup \{i\}|} \sum_{j \in N(p_i) \cup \{i\}} x_j \right) \quad (8)$$

Where $|N(p_i)|$ represents the number of nodes in the neighborhood $N(p_i)$.

Traditional point cloud processing methods, such as simple voxelization or direct feature extraction, which often struggles to accurately capture local geometric structure information in point cloud data. Taking voxelization as an example, it divides the three-dimensional space into regular voxel grids and discretizes point cloud data into these grids. Although this approach reduces the complexity of the data to a certain extent, it also loses a large amount of local detail information. Edge convolution can fully consider the relationship between each point and its neighboring points when processing hand surface point cloud. Through edge convolution operation, the model can learn the local geometric structure in point cloud data based on the relative position between the center point and neighboring points.

In addition, traditional point cloud processing methods typically use fixed feature extraction methods, which may lack adaptability for

different types of point cloud data. Edge convolution is different, it is a dynamic feature extraction method. In edge convolution, the features of each point are obtained by aggregating the features of its neighboring points. This feature extraction method enables edge convolution to learn more representative features based on the local structure and feature distribution of point cloud data. When processing hand surface point cloud, different gestures may have different local feature distributions. Edge convolution can adjust the feature extraction method based on these differences, thereby improving the expression ability of features and better distinguishing different gestures.

2.3 Enhancement of hand surface point cloud dimension abstraction in human-machine interaction

In hand surface point cloud recognition methods, improving the dimensional information of the point cloud helps to capture richer hand surface features and enhance the recognition ability of the model. Multilayer perceptron, as a feedforward neural network, can effectively handle nonlinear problems and abstract the intrinsic features of data (Yu et al., 2023; Lyu et al., 2022; Yang and Fang, 2024; Khan et al., 2022). Therefore, after edge convolution, we introduce a multilayer perceptron to further process hand surface point cloud data. After each fully connected layer, a multilayer perceptron follows an activation function to introduce nonlinear transformations. By stacking multiple such layers, multilayer perceptron can learn complex mappings from low to high dimensions and extract advanced features of hand surface point clouds.

Before inputting gesture point cloud data into a multilayer perceptron, we first perform preliminary processing on the gesture point cloud using edge convolution and graph algorithms to obtain local features for each point. Then, we use these local features as inputs to the multilayer perceptron, and gradually abstract the global features of the gesture point cloud through layer by layer transmission and transformation. An important feature of multilayer perceptron is its ability to enhance the dimensionality of data. By increasing the number of fully connected layers and neurons, we can map the input low dimensional features to a high dimensional space, thereby extracting more complex feature representations. The above process is shown in Figure 4. In hand surface point cloud recognition, this dimension enhancement helps to capture subtle changes and diversity in gestures, improving the recognition accuracy of GSFAN.

After edge convolution and KNN graph algorithm processing, the local feature representation of each hand surface point cloud can be obtained. Set $F = \{f_1, f_2, \dots, f_N\}$ to represent the local feature set of the hand surface point cloud, where $f_i \in R^{D_0}$, N is the number of points, and D_0 is the initial feature dimension.

For the i output feature of the l layer in a multilayer perceptron, it can be expressed as Equation (9).

$$h_i^{(l)} = \sigma \left(\sum_{j=1}^{D_{l-1}} \omega_{ij}^{(l)} h_j^{(l-1)} + b_i^{(l)} \right) \quad (9)$$

Where $h_i^{(l)}$ is the feature vector of the i row of the output matrix $H^{(l)}$ in the l layer, $\omega_{ij}^{(l)}$ is the element of the i row and j column of the

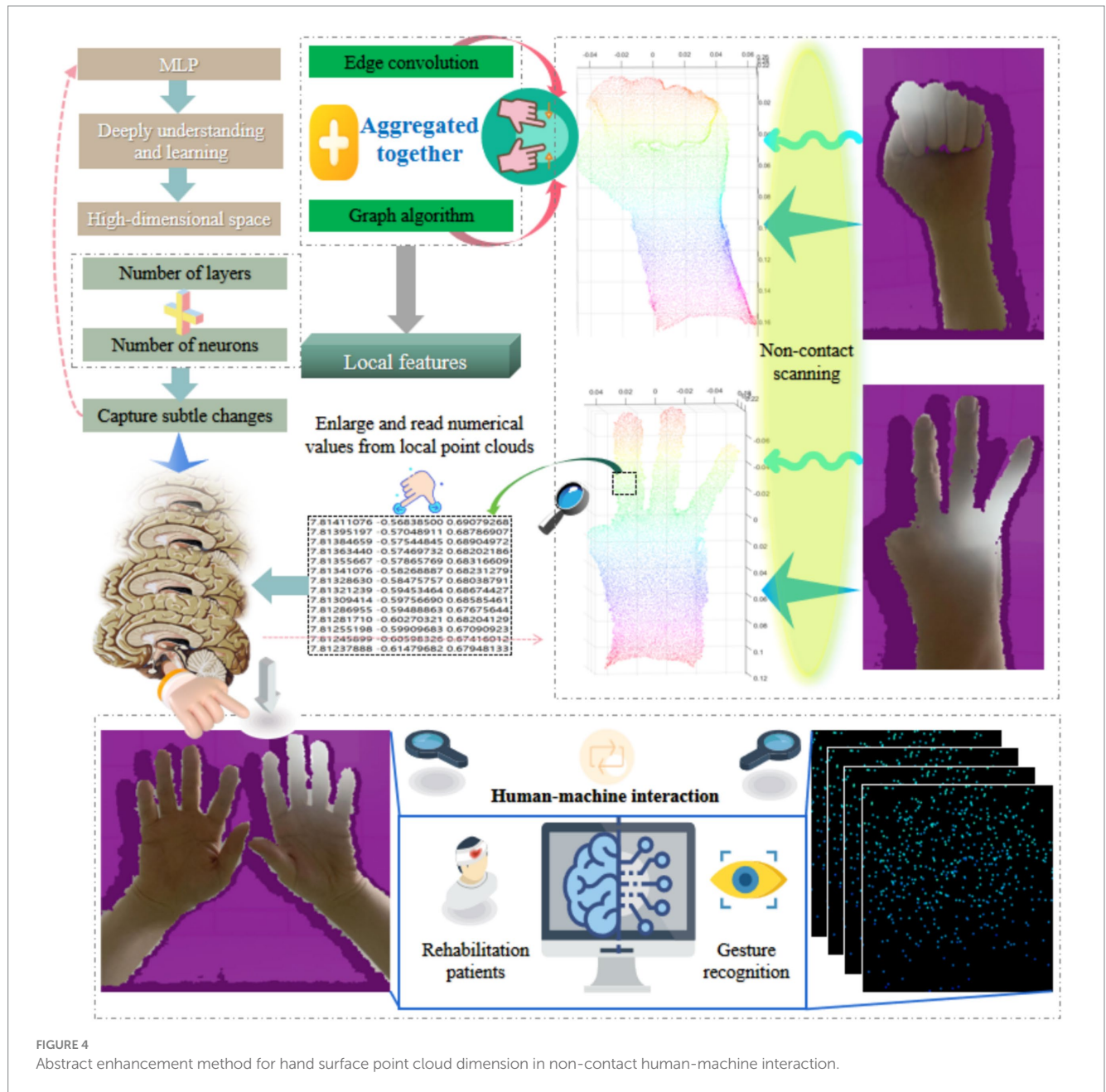


FIGURE 4

Abstract enhancement method for hand surface point cloud dimension in non-contact human-machine interaction.

weight matrix $W^{(l)}$, $h_j^{(l-1)}$ is the vector composed of the features of the j column of the output matrix $H^{(l-1)}$ in the $l-1$ layer. $b_i^{(l)}$ is the i -th element of the bias vector $b^{(l)}$. In order to obtain the output matrix $H^{(l)}$ of the entire l layer, we need to perform the above calculation on each $i \in \{1, 2, \dots, D_l\}$ and stack the results in matrix form, which can be expressed as Equation (10).

$$H^{(l)} = \sigma(W^{(l)}H^{(l-1)} + 1_N \otimes b^{(l)}) \quad (10)$$

Where 1_N is a full 1-column vector of $N \times 1$, $\otimes$ represents Kronecker product or outer product, used to extend the bias vector

$b^{(l)}$ to the same shape as $H^{(l-1)}$ for element wise addition. The complete matrix $H^{(0)}$ can be expressed as Equation (11).

$$H^{(0)} = \begin{bmatrix} \sum_{j=1}^N \omega_{1j} \cdot M(j \cdot 1) & \sum_{j=1}^N \omega_{1j} \cdot M(j \cdot 2) & \dots & \sum_{j=1}^N \omega_{1j} \cdot M(j \cdot D_0) \\ \sum_{j=1}^N \omega_{2j} \cdot M(j \cdot 1) & \sum_{j=1}^N \omega_{2j} \cdot M(j \cdot 2) & \dots & \sum_{j=1}^N \omega_{2j} \cdot M(j \cdot D_0) \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{j=1}^N \omega_{Nj} \cdot M(j \cdot 1) & \sum_{j=1}^N \omega_{Nj} \cdot M(j \cdot 2) & \dots & \sum_{j=1}^N \omega_{Nj} \cdot M(j \cdot D_0) \end{bmatrix} \quad (11)$$

In the formula, M is the mapping function. We introduce a combination of additional nonlinear functions, matrix multiplication, and element wise multiplication, which can be expressed as Equation (12).

$$\tilde{z} = V \cdot [\sigma(z) \odot \tanh(W_2 \cdot z + \beta) + \gamma] \quad (12)$$

Where V and W_2 are weight matrices, β and γ are additional learnable parameters, $\sigma(z)$ is the result of applying non-linear functions to z , $\tanh(W_2 \cdot z + \beta)$ is the result of non-linear transformation, $\odot$ represents element wise multiplication.

3 Results and discussions

3.1 Experimental details

We divided the collected hand surface point cloud dataset into training and testing sets for GSFAN training and evaluation. In order to ensure a certain degree of randomness and representativeness in the training and evaluation data during the experimental process, and to maintain a balance between different gesture categories in the dataset, in order to avoid GSFAN bias caused by class imbalance.

In this research, we used accuracy to evaluate the model. Choosing accuracy for evaluation is because it has significant advantages. In the gesture recognition scenario of hand function rehabilitation for the elderly population that we focus on, the accuracy can quickly demonstrate the overall reliability of the model for gesture recognition in practical applications, providing a strong basis for judging whether the model can meet basic needs. In addition, we also use tools such as confusion matrices to visualize and analyze the classification results of the model, in order to further research the performance of the model.

We not only conducted multiple repeated experiments, but also, due to different patients undergoing hand function rehabilitation at home, GSFAN needs to face diverse hand surface point clouds. Therefore, we collected the hand surface point clouds of different experimenters to test GSFAN and fully demonstrate its generalization. The configuration for conducting relevant experiments in this section is as follows: the processor is the 13th Gen Intel Core i9-13900 K, with a reference frequency of 3.00GHz and a maximum acceleration frequency of 5.80GHz. The graphics card is an NVIDIA GeForce RTX 4080 independent graphics card, with a core frequency of 2,205 MHz, a graphics memory frequency of 1,400 MHz, and a boost frequency of 2,505 MHz. The memory configuration consists of two 32GB DDR5 memory modules. The operating system is Windows 10 64 bit. In addition, different cutting-edge deep neural networks were constructed using Python 3.7.9 and TensorFlow 2.3.0 deep learning frameworks. The key hyperparameter settings are shown in Table 1.

Although the ReLU function solves the problem of vanishing gradients, when $x < 0$, the gradient becomes 0, making the neuron invalid and not updated in the subsequent training process. Leaky ReLU does not use the method of all zeros in non-positive parts, but assigns a non-zero slope in non-positive parts (Wang and Liu, 2024;

Gezawa et al., 2023). Therefore, Leaky ReLU is selected as the activation function.

3.2 Multi perspective analysis of edge convolution results

The number of edge convolution layers plays a crucial role in capturing local geometric features and learning discriminative representations of input point cloud data. In order to research the effect of different edge convolution layers, we conducted a series of experiments, and the number of edge convolution layers gradually increased from one layer. In all experiments, the remaining parts of the network and parameter settings were kept consistent to ensure the fairness of the experiment.

The more edge convolution layers attached within GSFAN, the larger the receptive field range of the hand surface point cloud, as shown in Figure 5a. We analyzed the results obtained by four experimenters under different numbers of edge convolution layers. In order to ensure the comprehensiveness and objectivity of the experiment, the results of each experiment were the average of multiple repeated experiments, as shown in Figure 5b. The experimental results of each experimenter under different numbers of edge convolution layers are shown in Figure 5c.

The performance of hand surface point cloud recognition methods is significantly affected by different numbers of edge convolution layers. We can see from the experimental results that as the number of edge convolution layers increases, the performance of hand surface point cloud recognition gradually improves. Starting from the number of edge convolution layers being 1, the accuracy gradually increases. Appropriately increasing the number of edge convolution layers can effectively improve the model's ability to extract hand surface point cloud features, thereby enhancing the accuracy of gesture recognition. The edge convolution operation is crucial in the model, as it aggregates and updates the features of each node's neighboring nodes to obtain a new representation of the node. When stacking multiple edge convolution layers, the model can capture deeper and farther distance hand surface point cloud features. Each layer of edge convolution further refines and integrates features based on the previous layer, enabling the model to gradually learn richer and more complex hand surface point cloud features. As the number of layers increases, the model can more accurately describe the entire gesture.

However, the number of edge convolution layers is not the more the better. When there are too many edge convolution layers, the performance of the model will actually decrease. This is because excessive edge convolution layers can make the features between neighboring information too smooth when aggregating the neighborhood information of the central point. In this process, originally discriminative features are excessively fused, resulting in a weakening of the uniqueness and differences of the features, making the features between different gestures similar and lacking sufficient discriminability. In this way, it is difficult for the model to accurately

TABLE 1 Hyperparameter settings.

Hyperparameters	Learning rate	Batch size	Momentum	Dropout rate	Loss function
Select	0.001	32	0.9	0.5	Cross-entropy

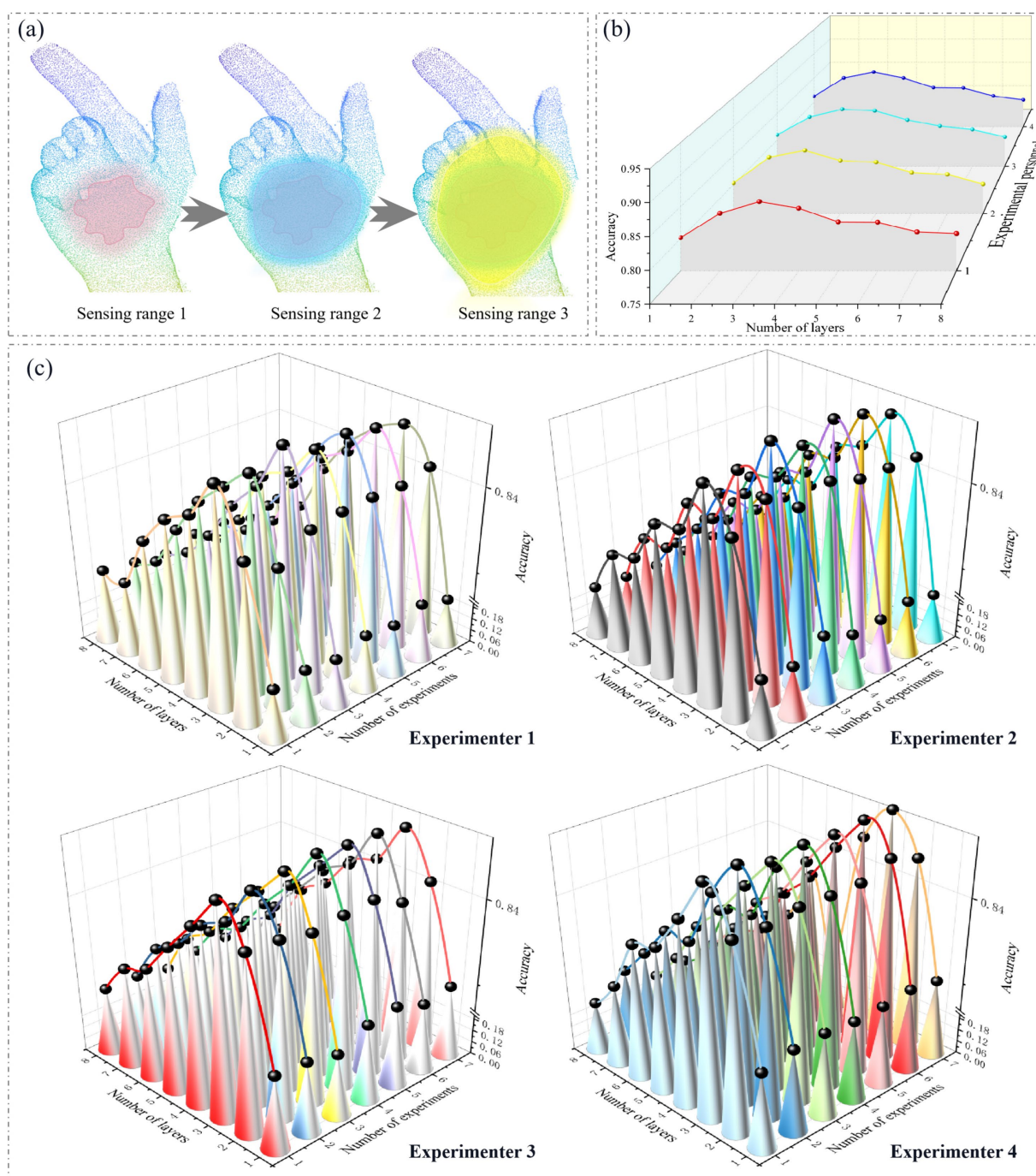


FIGURE 5

(a) Different receptive field ranges of hand surface point clouds. (b) Average values of experimenters in multiple experiments under different numbers of edge convolution layers. (c) Results of each experiment conducted by each experimenter under different numbers of edge convolution layers.

distinguish different gestures during gesture recognition, resulting in a decrease in recognition accuracy.

In addition, we can see that the accuracy changes obtained by different experimenters not only have consistent trends, but also the maximum accuracy values occur when the number of edge convolution layers is 3, achieving a relatively high level of performance. Furthermore, further increasing the number of edge convolution layers does not bring significant performance improvement. Increasing the number of edge convolution layers will increase the complexity of the model, as well as the number of parameters and

computational complexity. Therefore, it is reasonable to choose an appropriate number of edge convolution layers.

3.3 Multi perspective analysis of dimensional abstraction enhancement processing results

Abstract enhancement of the dimension of the hand surface point cloud can enable the 3D deep learning model to learn the features of the point cloud more accurately, as shown in Figure 6a. We designed

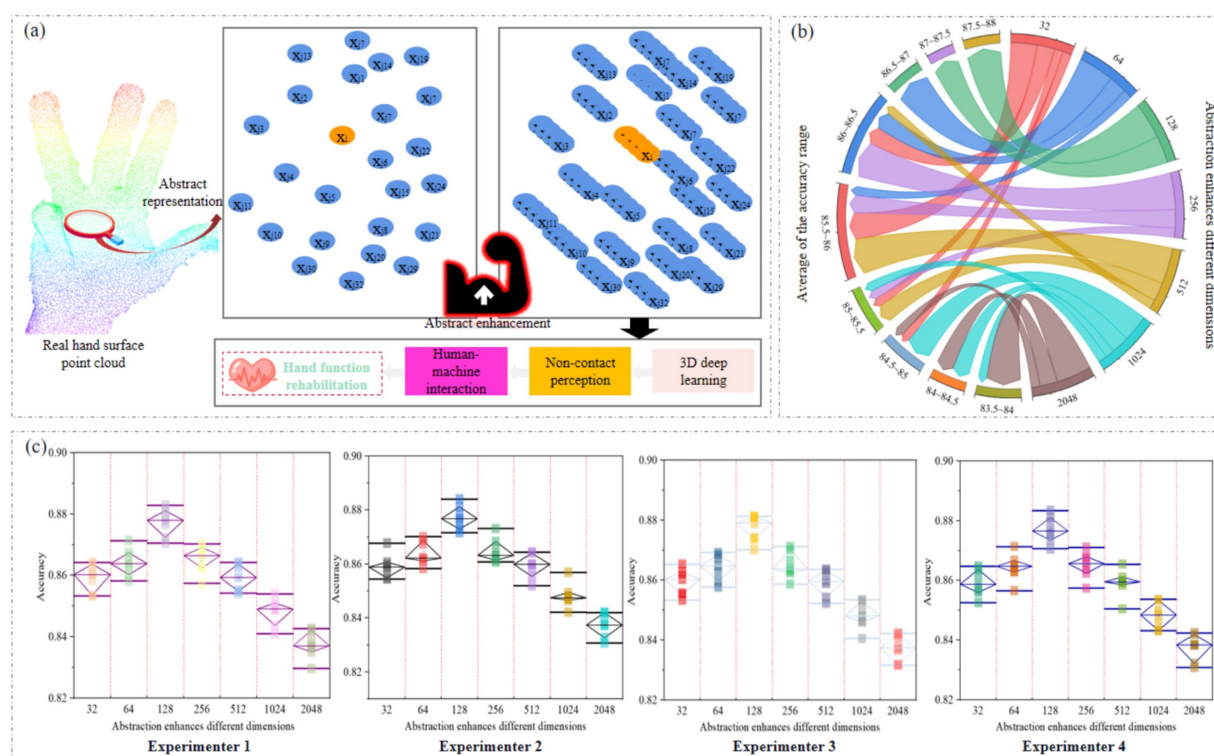


FIGURE 6

(a) Abstract enhancement of dimension of hand surface point cloud. (b) Chord diagrams of multiple repeated experiments conducted by all experimenters in different dimensions. (c) The experimental results of each experimenter in different dimensions.

multiple experiments to increase the dimensions of the hand surface point cloud to 32, 64, 128, 256, 512, 1,024, and 2,048. In each experiment, we used the same training and testing sets, and trained and evaluated them under the same network parameter configuration. At the same time, we analyzed the results obtained by four experimenters under different dimensions of abstraction enhancement. In order to ensure the comprehensiveness and objectivity of the experiment, each experimenter conducted multiple repeated experiments, as shown in Figure 6b. The experimental results of each experimenter in different dimensions are shown in Figure 6c.

The abstract enhancement of different dimensions has a significant impact on the performance of hand surface point cloud recognition methods. As the dimensions increase, the performance of the model shows a trend of first improving and then decreasing. Specifically, as the dimension gradually increases from 32, the accuracy of hand surface point cloud recognition gradually improves. This is because when the dimensions are small, the model's representation ability is limited, and it cannot fully capture and represent the advanced features of the hand surface point cloud. The hand surface point cloud contains rich information, and models in low dimensions find it difficult to effectively integrate and express this information, resulting in lower recognition accuracy. With the increase of dimensions, the representation ability of the model is enhanced, and it can handle more complex features, thereby improving the accuracy of gesture recognition. However, when the dimension reaches a certain threshold, the situation changes. Continuing to increase dimensions, the improvement in performance is no longer significant, and there may even be a decline. This is because when the dimensions are too

large, the model becomes increasingly complex, overfitting the training data during the training process, ultimately leading to a decrease in accuracy.

Based on the experimental results, we found that when the dimension is around 128, the accuracy of hand surface point clouds is higher, and the model can fully capture the advanced features of hand surface point clouds. Besides, enhancing dimensions is like adding more attributes to data. The original hand surface point cloud only had simple information such as position, but after adding features, more abstract descriptions were added to the hand surface point cloud. This enables GSFAN to more accurately recognize different gestures and achieve non-contact gesture surface feature analysis in human-machine interaction.

3.4 Multi perspective analysis of different cutting-edge models

In order to verify the progressiveness and superiority of the model we built in this research, we will analyze the results obtained by GSFAN, PointNet, PointNet++ in processing point clouds on the hand surface of different experimenters (Figure 7a). For fair comparison, we trained and tested in the same hardware and software environment, using the same dataset and evaluation metrics. The overall accuracy of different cutting-edge models on the hand surface point clouds of different experimenters is shown in Figure 7b. The confusion matrix can observe the performance of the model on various categories. We analyze the confusion matrix of different cutting-edge models in

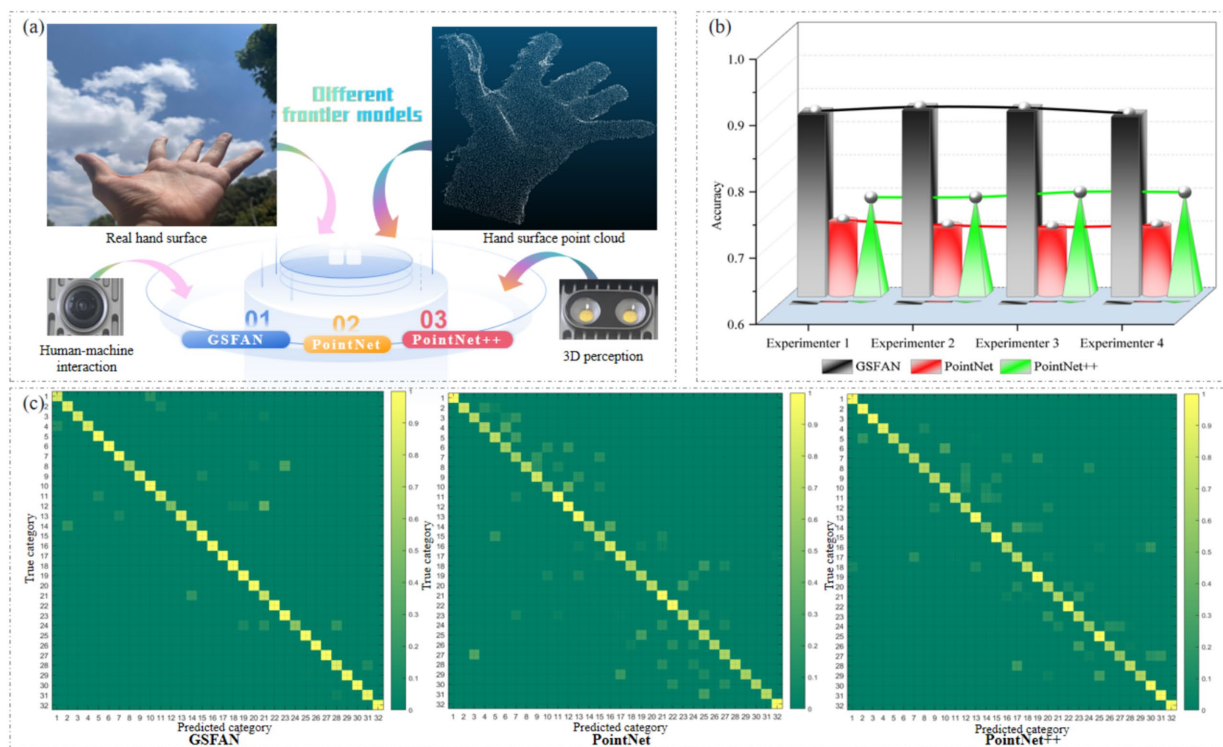


FIGURE 7

(a) Processing hand surface point clouds using different cutting-edge models. (b) Overall accuracy of different cutting-edge models on hand surface point clouds of different experimenters. (c) Confusion matrix of different cutting-edge models on second experimenter's hand surface point clouds.

the second experiment and evaluate the performance of different cutting-edge models on different gesture categories. Figure 7c shows the confusion matrix of 32 hand surface point clouds obtained from different cutting-edge models, where the horizontal axis represents the predicted results, the vertical axis represents the true results, and the diagonal matrix represents the recognition accuracy of each hand surface point cloud. The reason for choosing accuracy as the evaluation metric is that it can intuitively reflect the performance of the model in the overall gesture recognition task, and in actual hand function rehabilitation application scenarios, users are more concerned about whether the model can accurately recognize gestures, thereby ensuring the effectiveness of rehabilitation training.

In the comparative experiment, we observed that the GSFAN established in this research had an overall accuracy better than PointNet and PointNet++ for different experimenters, indicating that this model has higher accuracy in hand surface point cloud processing tasks. In addition, different cutting-edge models have certain sensitivity to the hand surface point clouds provided by different experimenters, and the results obtained may fluctuate and change. This is mainly because the hand surface point clouds of each experimenter cannot be completely consistent, but the accuracy obtained by GSFAN can still reach a high level, with an average accuracy of 88.72%. And the above experimental results also demonstrate that GSFAN can still achieve good results and has good robustness when facing different hand surface point clouds given by experimenters.

In order to conduct a more in-depth analysis of the performance of each model, we further plotted the confusion matrix. The confusion

matrix can clearly demonstrate the recognition ability of the model on various categories, especially on easily confused hand surface point cloud categories. It can be seen that for the 32 types of hand surface point clouds, GSFAN is able to distinguish each type of point cloud well and has good classification performance. PointNet and PointNet++ have a high false recognition rate on certain specific hand surface point cloud categories. Through comparative analysis, we found that the GSFAN established in this research exhibits superior performance in hand surface point cloud processing tasks. This is mainly due to the following aspects: the GSFAN constructed in this research obtains local information of the hand surface point cloud through edge convolution, which can more effectively capture the spatial structure and local detail features of gestures. Moreover, GSFAN increases the dimensional information of the gesture point cloud, which can further extract advanced feature representations of the hand surface point cloud, thereby improving the model's discriminative ability. In contrast, although PointNet and PointNet++ can handle unordered point cloud data, they have certain limitations when dealing with hand surface point cloud data. PointNet is unable to capture local detail features in point clouds. As an improved version of PointNet, although PointNet++ has improved the ability to extract local features to a certain extent, it still has not escaped the constraints of PointNet in essence, and still appears inadequate when dealing with complex hand surface point clouds.

Through the above experiments and result analysis, we can conclude that the model established in this research has good performance in hand surface point cloud processing tasks. This provides a new approach and method for non-contact analysis of

gesture surface features in human-machine interaction in the field of hand function rehabilitation.

In traditional hand function rehabilitation scenarios, rehabilitation methods often rely on physical contact and manual guidance. Patients need to perform various rehabilitation actions with the assistance of therapists, and the rehabilitation process is usually mechanically repeated according to established procedures, which makes the entire rehabilitation process appear dull and boring. Moreover, manual guidance has certain limitations, such as the difficulty for therapists to accurately monitor the quality and subtle changes of each patient's movements at all times. High accuracy gesture recognition enables patients to interact with rehabilitation equipment in a more natural and convenient way. Patients only need to make gestures, and rehabilitation equipment can accurately recognize and provide corresponding feedback without the need for tedious operations. The natural interaction greatly enhances the patient's rehabilitation experience and makes the rehabilitation process more enjoyable. Positive psychological experiences can greatly stimulate patients' rehabilitation enthusiasm, making them more willing to actively participate in rehabilitation training, thereby improving rehabilitation outcomes.

In addition, during the rehabilitation process, therapists need to accurately determine whether patients can complete specific rehabilitation training and adjust the training intensity and methods in a timely manner. High accuracy gesture recognition can provide reliable reference for therapists. Therapists can gain a clear understanding of patients' hand movements based on relevant feedback, which helps them develop more scientific and personalized rehabilitation plans, and improve the efficiency and quality of rehabilitation treatment.

In order to better adapt the model to the characteristics of each individual, we plan to introduce personalized training mechanisms in the future. In practical applications, users perform a series of simple gestures, and we use this data to fine tune the pre trained model. By fine-tuning, the model can adjust parameters based on the individual characteristics of new users, thereby better recognizing their gestures. This personalized training mechanism can improve the recognition accuracy of the model for each individual without affecting the overall performance of the model. In addition to the features of the hand surface point cloud itself, we will also consider integrating other relevant features, such as the user's basic information (age, gender, etc.) and the physiological characteristics of the hand (hand size, finger length, etc.). These additional features can provide the model with more information about individual differences, helping the model better understand and adapt to the characteristics of different individuals. Meanwhile, we will introduce an adaptive adjustment mechanism in the model to enable it to automatically adjust decision rules based on the characteristics of input data. For example, when the model detects that the input data is similar to an individual feature in the training data, it can automatically adjust the corresponding weights to improve recognition accuracy.

This research has limitations in some aspects. Firstly, we used laser sensors to collect hand surface point cloud, but this hardware has certain limitations. On the one hand, the accuracy of laser sensors is limited, which may result in inaccurate point cloud data collection and loss of some subtle features on the hand surface. On the other hand, the collection range of laser sensors is also limited. When the hand exceeds its effective range, some point cloud data will be missing,

which will affect subsequent feature extraction and model training. Secondly, in this research, the number of samples used for model training and testing was relatively small. This is mainly due to the fact that the data collection process requires a significant amount of time and labor costs. A small sample size may lead to insufficient generalization ability of the model, making it difficult to adapt well to gesture changes in different individuals and environments. For example, there may be differences in hand size among different individuals, and small sample data may not cover all of these changes, resulting in poor performance of the model when faced with new data. Finally, the current research mainly focuses on the accuracy of gesture recognition, but lacks validation of the real-time performance of the model. In practical human-machine interaction and rehabilitation application scenarios, real-time performance is a very important indicator. Patients need to receive timely feedback in order to adjust their actions. However, due to the high computational complexity of the model, it can cause delays in recognition results and fail to meet real-time requirements. In the future, we will research lightweight deep learning architectures to reduce the computational complexity of models.

4 Conclusion

In the current context of the interweaving trend of normalized epidemic prevention and control and aging population, this research focuses on the field of intelligent rehabilitation of hand function in human-machine interaction. A 3D deep learning model was constructed to process laser sensor point clouds and achieve non-contact gesture surface feature analysis. Through this innovative method, we are not only able to accurately recognize and analyze the surface features of gestures, but also have achieved significant breakthroughs in the security and convenience of human-machine interaction. Firstly, this research effectively captured the spatial structure and local detail features of gestures through edge convolution. Through experiments, it was found that the model performed best when the number of edge convolution layers in GSFAN was 3. Secondly, abstraction enhances the dimensional information of the hand surface point cloud, which can further extract advanced feature representations of the hand surface point cloud. Through experiments, it was found that the recognition accuracy of the hand surface point cloud is higher when the dimension is 128. Finally, this research verified the progressiveness and superiority of GSFAN through comparative experiments. Compared with the cutting-edge point cloud processing model, GSFAN showed higher accuracy and lower confusion rate when classifying 32 hand surface point clouds. The average accuracy of GSFAN is 88.72%. The results of this research indicate that our proposed model has stronger application potential and practical value in the field of intelligent rehabilitation of hand function in human-machine interaction.

In addition, in the subsequent research work, we will conduct extensive research on cutting-edge research achievements in the field of point cloud processing, select representative and influential advanced models, and carry out systematic comparative experiments. During the experiment, we will strictly control the experimental conditions to ensure the scientificity and reliability of the comparative results. By comparing the key indicators of different models in hand surface point cloud segmentation tasks, comprehensively evaluate the

performance and advantages of GSPAN. At the same time, we will conduct in-depth analysis and discussion of the experimental results, exploring the advantages and disadvantages of different models in processing hand surface point cloud.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding authors.

Author contributions

ZX: Formal analysis, Funding acquisition, Investigation, Methodology, Software, Supervision, Writing – original draft, Writing – review & editing. ZM: Conceptualization, Methodology, Writing – original draft. GZ: Investigation, Methodology, Software, Supervision, Writing – review & editing. GM: Data curation, Writing – original draft. LY: Methodology, Writing – original draft. XG: Software, Writing – review & editing. LT: Data curation, Writing – original draft. YJ: Software, Writing – original draft. HW: Funding acquisition, Investigation, Software, Supervision, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This research was supported by the Open Project Program of Fujian Key Laboratory of Special Intelligent Equipment Measurement and Control (grant number FJIES2023KF13), the Opening Project of Medical Imaging Key Laboratory of Sichuan Province (grant number MIKL202205), the

Yunnan Fundamental Research Projects (grant numbers 202401AU070071 and 202401AU070037), the Scientific Research Fund of the Yunnan Provincial Department of Education (grant numbers 2024J0337 and 2024J0147), the Kunming Medical University College Student Innovation Training Program (grant number 2024CYD001), the Kunming Medical University Talent Introduction Research Special Project (grant numbers 2404 and 2431), the Research Projects of the State Administration for Market Regulation (grant number 2023MK070), the Research Project of Fujian Provincial Market Supervision Bureau (grant number FJMS2023016), and the National Natural Science Foundation of China (grant number 82360365).

Conflict of interest

LY was employed by Xi'an Aerospace Automation Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Akhtar, A., Li, Z., Van der Auwera, G., Li, L., and Chen, J. (2022). Pu-dense: sparse tensor-based point cloud geometry upsampling. *IEEE Trans. Image Process.* 31, 4133–4148. doi: 10.1109/TIP.2022.3180904
- Aljumaily, H., Laefer, D. F., Cuadra, D., and Velasco, M. (2023). Point cloud voxel classification of aerial urban LiDAR using voxel attributes and random forest approach. *Int. J. Appl. Earth Obs. Geoinf.* 118:103208. doi: 10.1016/j.jag.2023.103208
- Ameur, S., Khalifa, A. B., and Bouhlel, M. S. (2020). A novel hybrid bidirectional unidirectional LSTM network for dynamic hand gesture recognition with leap motion. *Enter Comput* 35:100373. doi: 10.1016/j.entcom.2020.100373
- An, S., Zhu, H., Guo, C., Fu, B., Song, C., and Tao, P. (2022). Deng, T noncontact human-machine interaction based on hand-responsive infrared structural color. *Nat. Commun.* 13:1446. doi: 10.1038/s41467-022-29197-5
- Baraković, S., Akhtar, Z., and Baraković Husić, J. (2024). Quality of life improvement: smart approaches for the working and aging populations. *Front. Public Health* 12:1362019. doi: 10.3389/fpubh.2024.1362019
- Bates, M., and Sunderam, S. (2023). Hand-worn devices for assessment and rehabilitation of motor function and their potential use in BCI protocols: a review. *Front. Hum. Neurosci.* 17:1121481. doi: 10.3389/fnhum.2023.1121481
- Cao, W., Lu, P., and Cao, W. (2024). Multimodal gesture recognition with Spatio-temporal features fusion based on YOLOv5 and MediaPipe. *Int. J. Pattern Recognit. Artif. Intell.* 38:2455007. doi: 10.1142/S0218001424550073
- Edger-Lacoursière, Z., Deziel, E., and Nedelec, B. (2023). Rehabilitation interventions after hand burn injury in adults: a systematic review. *Burns* 49, 516–553. doi: 10.1016/j.burns.2022.05.005
- Ergen, H. İ., Keskinbiçki, M. V., and Öksüz, Ç. (2024). The effect of proprioceptive training on hand function and activity limitation after open carpal tunnel release surgery: a randomized controlled study. *Arch. Phys. Med. Rehabil.* 105, 664–672. doi: 10.1016/j.apmr.2023.12.004
- Galván-Ruiz, J., Travieso-González, C. M., Tejera-Fettmilch, A., Pinan-Roescher, A., Esteban-Hernández, L., and Domínguez-Quintana, L. (2020). Perspective and evolution of gesture recognition for sign language: a review. *Sensors* 20:3571. doi: 10.3390/s20123571
- Gao, N., Chen, P., and Liang, L. (2023). BCI–VR-based hand soft rehabilitation system with its applications in hand rehabilitation after stroke. *Int. J. Precis. Eng. Manuf.* 24, 1403–1424. doi: 10.1007/s12541-023-00835-2
- Gezawa, A. S., Liu, C., Jia, H., Nanekharan, Y. A., Almutairi, M. S., and Chiroma, H. (2023). An improved fused feature residual network for 3D point cloud data. *Front. Comput. Neurosci.* 17:1204445. doi: 10.3389/fncom.2023.1204445
- Gu, Y., Xu, Y., Shen, Y., Huang, H., Liu, T., Jin, L., et al. (2022). A review of hand function rehabilitation systems based on hand motion recognition devices and artificial intelligence. *Brain Sci.* 12:1079. doi: 10.3390/brainsci12081079
- Guo, K., Orban, M., Lu, J., Al-Quraishi, M. S., Yang, H., and Elsamanty, M. (2023). Empowering hand rehabilitation with ai-powered gesture recognition: a study of an semg-based system. *Bioengineering* 10:557. doi: 10.3390/bioengineering10050557
- Harada, A., Kawai, N., Ogawa, T., Hatakeyama, T., and Tamiya, T. (2021). Long-term multidisciplinary rehabilitation efficacy in older patients after traumatic brain injury: assessed by the functional independence measure. *Acta Med. Okayama* 75, 479–486. doi: 10.18926/AMO/62400
- Huang, Z., Xu, S., Liu, J., Wu, L., Qiu, J., Wang, N., et al. (2023). Effectiveness of inactivated COVID-19 vaccines among older adults in Shanghai: retrospective cohort study. *Nat. Commun.* 14:2009. doi: 10.1038/s41467-023-37673-9

- Jiang, S., Kang, P., Song, X., Lo, B. P., and Shull, P. B. (2021). Emerging wearable interfaces and algorithms for hand gesture recognition: a survey. *IEEE Rev. Biomed. Eng.* 15, 85–102. doi: 10.1109/RBME.2021.3078190
- Khan, A., Khan, A., Ullah, M., Alam, M. M., Bangash, J. I., and Suud, M. M. (2022). A computational classification method of breast cancer images using the VGGNet model. *Front. Comput. Neurosci.* 16:1001803. doi: 10.3389/fncom.2022.1001803
- Li, Y., Luo, Y., Gu, X., Chen, D., Gao, F., and Shuang, F. (2021). Point cloud classification algorithm based on the fusion of the local binary pattern features and structural features of voxels. *Remote Sens.* 13:3156. doi: 10.3390/rs13163156
- Li, J., Sun, Y., Luo, S., Zhu, Z., Dai, H., Krylov, A. S., et al. (2021). P2V-RCNN: point to voxel feature learning for 3D object detection from point clouds. *IEEE Access* 9, 98249–98260. doi: 10.1109/ACCESS.2021.3094562
- Li, L., Taeihagh, A., and Tan, S. Y. (2023). A scoping review of the impacts of COVID-19 physical distancing measures on vulnerable population groups. *Nat. Commun.* 14:599. doi: 10.1038/s41467-023-36267-9
- Li, W., Wu, S., Li, S., Zhong, X., Zhang, X., Qiao, H., et al. (2023). Gesture recognition system using reduced graphene oxide-enhanced hydrogel strain sensors for rehabilitation training. *ACS Appl. Mater. Interfaces* 15, 45106–45115. doi: 10.1021/acsami.3c08709
- Li, H., Wu, L., Wang, H., Han, C., Quan, W., and Zhao, J. (2019). Hand gesture recognition enhancement based on spatial fuzzy matching in leap motion. *IEEE Trans. Inform. Inform.* 16, 1885–1894. doi: 10.1109/TII.2019.2931140
- Liu, L., Xu, W., Ni, Y., Xu, Z., Cui, B., Liu, J., et al. (2022). Stretchable neuromorphic transistor that combines multisensing and information processing for epidermal gesture recognition. *ACS Nano* 16, 2282–2291. doi: 10.1021/acsnano.1c08482
- Lopes, D. G., Mendonça, N., Henriques, A. R., Branco, J., Canhão, H., and Rodrigues, A. M. (2023). Trajectories and determinants of ageing in Portugal: insights from EpiDoC, a nationwide population-based cohort. *BMC Public Health* 23:1564. doi: 10.1186/s12889-023-16370-8
- Lu, L., Hu, G., Liu, J., and Yang, B. (2024). 5G nb-IoT system integrated with high-performance fiber sensor inspired by cirrus and spider structures. *Adv Sci* 11:e2309894. doi: 10.1002/advs.202309894
- Lyu, J., Shi, H., Zhang, J., and Norvilitis, J. (2022). Prediction model for suicide based on back propagation neural network and multilayer perceptron. *Front. Neuroinform.* 16:961588. doi: 10.3389/fninf.2022.961588
- Mariappan, S., Murugesan, P., and Selvan, H. M. (2024). Real-time interpreter for short sentences in Indian sign language using MediaPipe and deep learning. *Informat Technol Control* 53, 888–898. doi: 10.5755/j01.itc.53.3.33935
- Najafinejad, A., and Korayem, M. H. (2023). Detection and minimizing the error caused by hand tremors using a leap motion sensor in operating a surgeon robot. *Measurement* 221:113544. doi: 10.1016/j.measurement.2023.113544
- Qi, C. R., Su, H., Mo, K., and Guibas, L. J. PointNet: deep learning on point sets for 3D classification and segmentation. In: *IEEE Conference on Computer Vision and Pattern Recognition* (2017a), 77–85.
- Qi, C. R., Yi, L., Su, H., and Guibas, L. J. PointNet++: deep hierarchical feature learning on point sets in a metric space. In: *Advances in Neural Information Processing System*. (2017b), 5105–5114.
- Rufei, L., Jiben, Y., Hongwei, R., Bori, C., and Chenhao, C. (2022). Research on a pavement pothole extraction method based on vehicle-borne continuous laser scanning point cloud. *Meas. Sci. Technol.* 33:115204. doi: 10.1088/1361-6501/ac875c
- Sadaoui, S. E., Mehdi-Souzani, C., and Lartigue, C. (2022). Multisensor data processing in dimensional metrology for collaborative measurement of a laser plane sensor combined to a touch probe. *Measurement* 188:110395. doi: 10.1016/j.measurement.2021.110395
- Samaan, G. H., Wadie, A. R., Attia, A. K., Asaad, A. M., Kamel, A. E., Slim, S. O., et al. (2022). Mediapipe's landmarks with rnn for dynamic sign language recognition. *Electronics* 11:3228. doi: 10.3390/electronics11193228
- Samhan, A. F., Abdelhalim, N. M., and Elnaggar, R. K. (2020). Effects of interactive robot-enhanced hand rehabilitation in treatment of paediatric hand-burns: a randomized, controlled trial with 3-months follow-up. *Burns* 46, 1347–1355. doi: 10.1016/j.burns.2020.01.015
- Suwabe, R., Saito, T., and Hamaguchi, T. (2024). Verification of criterion-related validity for developing a Markerless hand tracking device. *Biomimetics* 9:400. doi: 10.3390/biomimetics9070400
- Tan, L., Lin, X., Niu, D., Wang, D., Yin, M., and Zhao, X. Projected generative adversarial network for point cloud completion. In: *IEEE Transactions on Circuits System Video Technology* (2022), 33, 771–781.
- Tang, Z., Hong, X., Lv, S., Cui, Z., Sun, H., and Shao, J. (2025). Hand rehabilitation exoskeleton system based on EEG spatiotemporal characteristics. *Expert Syst. Appl.* 270:126574. doi: 10.1016/j.eswa.2025.126574
- Theodose, R., Denis, D., Chateau, T., Frémont, V., and Checchin, P. (2021). A deep learning approach for LiDAR resolution-agnostic object detection. *IEEE Trans. Intell. Transp. Syst.* 23, 14582–14593. doi: 10.1109/TITS.2021.3130487
- Tong, G., Li, Y., Chen, D., Xia, S., Peethambaran, J., and Wang, Y. (2020). Multi-view features joint learning with label and local distribution consistency for point cloud classification. *Remote Sens.* 12:135. doi: 10.3390/rs12010135
- Vaitkevičius, A., Taroza, M., Blažauskas, T., Damaševičius, R., Maskeliūnas, R., and Woźniak, M. (2019). Recognition of American sign language gestures in a virtual reality using leap motion. *Appl. Sci.* 9:445. doi: 10.3390/app9030445
- Wang, X., and Liu, J. (2024). Spatially regularized leaky ReLU in dual space for CNN based image segmentation. *Inverse Probl. Imaging* 18, 1320–1342. doi: 10.3934/ipi.2024016
- Wang, X., Sun, Y., Liu, Z., Sarma, S. E., Bronstein, M. M., and Solomon, J. M. Dynamic graph CNN for learning on point clouds. In: *ACM Transactions on Graphics*. (2019), 38, 1–12.
- Wang, S., Wang, X., Wang, S., Yu, W., Yu, L., Hou, L., et al. (2023). Optical-nanofiber-enabled gesture-recognition wristband for human-machine interaction with the assistance of machine learning. *Adv. Intell. Syst.* 5:2200412. doi: 10.1002/aisy.202200412
- Xiao, A., Huang, J., Guan, D., Zhang, X., Lu, S., and Shao, L. (2023). Unsupervised point cloud representation learning with deep neural networks: a survey. *IEEE Trans. Pattern Anal. Mach. Intell.* 45, 11321–11339. doi: 10.1109/TPAMI.2023.3262786
- Xing, Z., Ma, G., Wang, L., Yang, L., Guo, X., and Chen, S. (2025). Towards visual interaction: hand segmentation by combining 3D graph deep learning and laser point cloud for intelligent rehabilitation. *IEEE Internet Things J.* 99:1. doi: 10.1109/JIOT.2025.3546874
- Xing, Z., Zhao, S., Guo, W., Guo, X., Wang, Y., Bai, Y., et al. (2022). Identifying balls feature in a large-scale laser point cloud of a coal mining environment by a multiscale dynamic graph convolution neural network. *ACS Omega* 7, 4892–4907. doi: 10.1021/acsomega.1c05473
- Xing, Z., Zhao, S., Guo, W., Meng, F., Guo, X., Wang, S., et al. (2023). Coal resources under carbon peak: segmentation of massive laser point clouds for coal mining in underground dusty environments using integrated graph deep learning model. *Energy* 285:128771. doi: 10.1016/j.energy.2023.128771
- Yang, Y., and Fang, Q. (2024). Prediction of glycopeptide fragment mass spectra by deep learning. *Nat. Commun.* 15:2448. doi: 10.1038/s41467-024-46771-1
- Ye, J., Stewart, E., Chen, Q., Roberts, C., Hajiyavand, A. M., and Lei, Y. (2023). Deep learning and laser-based 3D pixel-level rail surface defect detection method. *IEEE Trans. Instrum. Meas.* 72:2513612. doi: 10.1109/TIM.2023.3272033
- Yozevitch, R., Frenkel-Toledo, S., Elion, O., Levy, L., Ambaw, A., and Holdengreber, E. (2023). Cost-effective and efficient solutions for the assessment and practice of upper extremity motor performance. *IEEE Sensors J.* 23, 23494–23499. doi: 10.1109/JSEN.2023.3303892
- Yu, H., Liu, Y., Zhou, G., and Peng, M. (2023). Multilayer perceptron algorithm-assisted flexible Piezoresistive PDMS/chitosan/CMWCNT sponge pressure sensor for sedentary healthcare monitoring. *ACS Sensors* 8, 4391–4401. doi: 10.1021/acssensors.3c01885
- Zestas, O. N., and Tselikas, N. D. (2023). Sollerman hand function sub-test “write with a pen”: a computer-vision-based approach in rehabilitation assessment. *Sensors* 23:6449. doi: 10.3390/s23146449
- Zhang, H., Qu, H., Teng, L., and Tang, C. Y. (2023). LSTM-MSA: a novel deep learning model with dual-stage attention mechanisms forearm EMG-based hand gesture recognition. *IEEE Trans. Neural Syst. Rehabil. Eng.* 31, 4749–4759. doi: 10.1109/TNSRE.2023.3336865
- Zhang, L., Sun, J., and Zheng, Q. (2018). 3D point cloud recognition based on a multi-view convolutional neural network. *Sensors* 18:3681. doi: 10.3390/s18113681
- Zheng, Q., Zou, B., Chen, W., Wang, X., Quan, T., and Ma, X. (2023). Study on the 3D-printed surface defect detection based on multi-row cyclic scanning method. *Measurement* 223:113823. doi: 10.1016/j.measurement.2023.113823
- Zhu, K., Guo, W., Yang, G., Li, Z., and Wu, H. (2021). High-fidelity recording of EMG signals by multichannel on-skin electrode arrays from target muscles for effective human-machine interfaces. *ACS Appl. Electr. Mater.* 3, 1350–1358. doi: 10.1021/acsaelm.0c01129
- Zhu, P., Niu, M., Liang, S., Yang, W., Zhang, Y., Chen, K., et al. (2025). Non-hand-worn, load-free VR hand rehabilitation system assisted by deep learning based on ionic hydrogel. *Nano Res.* 18:94907301. doi: 10.26599/NR.2025.94907301