

Article

Recognition of 3D Images by Fusing Fractional-Order Chebyshev Moments and Deep Neural Networks

Lin Gao ^{1,2}, Xuyang Zhang ^{3,*}, Mingrui Zhao ³ and Jinyi Zhang ^{1,4} ¹ School of Information Science and Engineering, Shenyang Ligong University, Shenyang 110159, China; zhangjinyi@sylu.edu.cn (J.Z.)² School of Computer Science and Engineering, Northeastern University, Shenyang 110169, China³ School of Mechanical Engineering, Shenyang Ligong University, Shenyang 110159, China⁴ Faculty of Engineering, Gifu University, Gifu 501-1193, Japan

* Correspondence: zhangxy_0403@126.com; Tel.: +86-187-42543541

Abstract: In order to achieve efficient recognition of 3D images and reduce the complexity of network parameters, we proposed a novel 3D image recognition method combining deep neural networks with fractional-order Chebyshev moments. Firstly, the fractional-order Chebyshev moment (FrCM) unit, consisting of Chebyshev moments and the three-term recurrence relation method, is calculated separately using successive integrals. Next, moment invariants based on fractional order and Chebyshev moments are utilized to achieve invariants for image scaling, rotation, and translation. This design aims to enhance computational efficiency. Finally, the fused network embedding the FrCM unit (FrCMs-DNNs) extracts depth features to analyze the effectiveness from the aspects of parameter quantity, computing resources, and identification capability. Meanwhile, the Princeton Shape Benchmark dataset and medical images dataset are used for experimental validation. Compared with other deep neural networks, FrCMs-DNNs has the highest accuracy in image recognition and classification. We used two evaluation indices, mean square error (MSE) and peak signal-to-noise ratio (PSNR), to measure the reconstruction quality of FrCMs after 3D image reconstruction. The accuracy of the FrCMs-DNNs model in 3D object recognition was assessed through an ablation experiment, considering the four evaluation indices of accuracy, precision, recall rate, and F1-score.



Citation: Gao, L.; Zhang, X.; Zhao, M.; Zhang, J. Recognition of 3D Images by Fusing Fractional-Order Chebyshev Moments and Deep Neural Networks. *Sensors* **2024**, *24*, 2352. <https://doi.org/10.3390/s24072352>

Academic Editors: Chih-Yang Lin and Ir. Kahlil Muchtar

Received: 26 February 2024

Revised: 23 March 2024

Accepted: 26 March 2024

Published: 7 April 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: fractional order; deep neural network; Chebyshev moments; image recognition

1. Introduction

Image recognition is an important field of artificial intelligence research, and different forms of moments are key descriptors for extracting relevant information in 3D images. The study of image moments has aroused strong interest among researchers. Moments are widely applied in image reconstruction [1–3], image analysis, image indexing [4–7], digital image research [8–10], spectral image super-resolution mapping [11], hyperspectral target detection [12], radar target recognition [13,14], SAR target recognition [15], sound classification [16], and other fields. Li et al. [17] employed an innovative face recognition method that integrated the Gabor wavelet representation of face images with an enhanced discriminator, the Complete Kernel Fisher Discriminant (CKFD), and fractional power polynomial (FPP) models to improve recognition performance and discrimination ability. The continuous functions in the orthogonal moments are employed as kernel functions, as they are not affected by rotation, scaling, or translation. The orthogonal moments include Legendre moments [18,19], Zernike moments [20], Fourier–Mellin moments [21], Chebyshev–Fourier moments [22], and so on. Due to the low efficiency of traditional image recognition methods, scholars have studied the application of fractional moments. Zhang et al. [23] adopted fractional-order orthogonal Fourier–Mellin moments, which can improve the calculation performance of image moments by removing the factorial term in

orthogonal polynomials. El Ogri et al. [24] used fractional generalized Laguerre moment invariants (FrGLMIs) to realize pattern recognition. Kaur et al. [25] used a support vector machine and fractional-order Zernike moments (FrZMs). Hosny et al. [26] created a set of fractional-order shifted Gegenbauer moments (FrSGMs) for image understanding and recognition. Horlando et al. [27] adopted fractional-order circular moments to solve some problems in image analysis. Guo et al. [28] introduced the Fractional-Order Fish Migration Optimization algorithm, which provides an optimal solution that can easily skip the whole order speed by using a new position generation strategy based on a global optimal solution. Zhang et al. [29] used fractional-order differentiation and closed image matting to perform multifocus image fusion.

Three-dimensional images find extensive applications in various fields including medicine, industry, and the military. For many applications in these areas, efficient identification and accurate analysis are essential. However, due to the large amount of 3D image data, its high complexity, and the need to capture both local and global features, traditional methods usually face a series of challenges.

A novel 3D image recognition approach, integrating fractional Chebyshev moments with deep neural networks, offers a promising solution to the aforementioned challenges. In this method, fractional Chebyshev moments are combined with deep neural networks. The concept of fractional calculus is used to extract multiscale, nonuniform, and nonlocal information from 3D images. The deep neural networks are combined with global spatial information for feature fusion and classification recognition. This method can effectively streamline network parameter selection, enhancing both the accuracy and speed of 3D image recognition. The benefits of this method include the following:

Improved recognition accuracy: The traditional 3D image recognition model may be inaccurate due to unreasonable network design or insufficient extraction of data features. The fractional Chebyshev moment and deep neural network combined method can capture local and global features in 3D images more comprehensively and accurately and can improve recognition accuracy.

Reduced complexity: For large-scale 3D image data processing, the implementation of traditional methods requires significant human effort, material resources, and time costs. In this method, a fractional Chebyshev moment algorithm is introduced for multidimensional feature extraction, and a deep convolutional neural network (DCNN) is used to quickly and accurately classify the processed data, thus ensuring the accuracy and reducing the complexity of network parameter selection.

Significant practical value: The efficient 3D image recognition method combining fractional Chebyshev moments and deep neural networks has been widely used and has achieved good performance in image classification, face recognition, mapping and modeling, object recognition, and medical image recognition, especially when dealing with noise interference and other complex cases. It significantly contributes to improving image recognition performance.

Therefore, the efficient 3D image recognition method combining fractional Chebyshev moments and deep neural networks is of great significance, and it has a very wide application prospect in solving practical problems.

2. 3D Object Recognition Based on FrCMs and DNNs

2.1. Fractional-Order Chebyshev Moments

The FrCMs adopt successive integrals to compute Chebyshev moments and three recursive relation methods, which effectively achieve the invariants of rotation, translation, and scaling.

2.1.1. Fractional-Order Chebyshev Moments

About the given function $f(x, y, z)$, the FrCMs of the order $\alpha(n + m + p)$ are defined within the region $[0, 1] \times [0, 1] \times [0, 1]$, and it is possible to calculate this by continuous integration.

$$FrCM_{nmp}^{\alpha} = \int_0^1 \int_0^1 \int_0^1 f(x, y, z) \tilde{F}T_n^{\alpha_x}(x) \tilde{F}T_m^{\alpha_y}(y) \tilde{F}T_p^{\alpha_z}(z) dx dy dz, \quad (1)$$

where α_x , α_y , and $\alpha_z > 0$, and $\tilde{F}T_n^{\alpha_x}(x)$, $\tilde{F}T_m^{\alpha_y}(y)$, and $\tilde{F}T_p^{\alpha_z}(z)$ are representative fractional-order Chebyshev polynomials.

For a digital image intensity function $f(i, j, k)$ of size $N \times M \times K$, the $FrCM_{nmp}^{\alpha}$ is expressed as

$$FrCM_{nmp}^{\alpha} = \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \sum_{k=0}^{K-1} f(i, j, k) \tilde{F}T_n^{\alpha_x}(x_i) \tilde{F}T_m^{\alpha_y}(y_j) \tilde{F}T_p^{\alpha_z}(z_k) \Delta_x \Delta_y \Delta_z \quad (2)$$

$\Delta_x = \frac{1}{N}$, $\Delta_y = \frac{1}{M}$, and $\Delta_z = \frac{1}{K}$, and the image coordinate of the mapping is expressed as

$$x_i = \frac{i}{N} + \frac{\Delta_x}{2}, y_j = \frac{j}{M} + \frac{\Delta_y}{2}, z_k = \frac{k}{K} + \frac{\Delta_z}{2}, \quad (3)$$

where $i = 1, 2, \dots, N$, $j = 1, 2, \dots, M$, and $k = 1, 2, \dots, K$.

The original image is approximated as

$$f(i, j, k) = \sum_{i=0}^{n_{max}} \sum_{j=0}^{m_{max}} \sum_{k=0}^{p_{max}} FrCM_{nmp}^{\alpha} \tilde{F}T_n^{\alpha_x}(x_i) \tilde{F}T_m^{\alpha_y}(y_j) \tilde{F}T_p^{\alpha_z}(z_k). \quad (4)$$

2.1.2. Fractional-Order 3D Moment Invariants

Given the image function $f(i, j, k)$, 3D fractional-order moment invariants (FrGMs) of order $(\alpha_x p + \alpha_y q + \alpha_z r)$ defined continuously on the region $N \times M \times K$ can be expressed as follows:

$$FrGM_{pqr}^{\alpha_x \alpha_y \alpha_z} = \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \sum_{k=0}^{K-1} f(i, j, k) m_{pqr}^{\alpha_x \alpha_y \alpha_z}(x_i, y_j, z_k), \quad (5)$$

where

$$m_{pqr}^{\alpha_x \alpha_y \alpha_z}(x_i, y_j, z_k) = \int_{x_i - \frac{\Delta x_i}{2}}^{x_i + \frac{\Delta x_i}{2}} \int_{y_j - \frac{\Delta y_j}{2}}^{y_j + \frac{\Delta y_j}{2}} \int_{z_k - \frac{\Delta z_k}{2}}^{z_k + \frac{\Delta z_k}{2}} x^{\alpha_x p} y^{\alpha_y q} z^{\alpha_z r} dx dy dz. \quad (6)$$

By dividing the function, it simplifies to

$$m_{pqr}^{\alpha_x \alpha_y \alpha_z}(x_i, y_j, z_k) = I_{Xp}^{\alpha_x}(x_i) I_{Yq}^{\alpha_y}(y_j) I_{Zr}^{\alpha_z}(z_k), \quad (7)$$

where

$$\begin{aligned} I_{Xp}^{\alpha_x}(x_i) &= \int_{x_i - \frac{\Delta x_i}{2}}^{x_i + \frac{\Delta x_i}{2}} x^{\alpha_x p} dx = \frac{1}{\alpha_x p + 1} \left[u_{i+1}^{\alpha_x p + 1} - u_i^{\alpha_x p + 1} \right] \\ I_{Yq}^{\alpha_y}(y_j) &= \int_{y_j - \frac{\Delta y_j}{2}}^{y_j + \frac{\Delta y_j}{2}} y^{\alpha_y q} dy = \frac{1}{\alpha_y q + 1} \left[v_{j+1}^{\alpha_y q + 1} - v_j^{\alpha_y q + 1} \right] \\ I_{Zr}^{\alpha_z}(z_k) &= \int_{z_k - \frac{\Delta z_k}{2}}^{z_k + \frac{\Delta z_k}{2}} z^{\alpha_z r} dz = \frac{1}{\alpha_z r + 1} \left[w_{k+1}^{\alpha_z r + 1} - w_k^{\alpha_z r + 1} \right]. \end{aligned} \quad (8)$$

After merging, the expression is expressed as

$$m_{pqr}^{\alpha_x \alpha_y \alpha_z} = \frac{1}{(\alpha_x p + 1)(\alpha_y q + 1)(\alpha_z r + 1)} \left[u_{i+1}^{\alpha_x p + 1} - u_i^{\alpha_x p + 1} \right] \left[v_{j+1}^{\alpha_y q + 1} - v_j^{\alpha_y q + 1} \right] \left[w_{k+1}^{\alpha_z r + 1} - w_k^{\alpha_z r + 1} \right], \quad (9)$$

where

$$\begin{aligned} u_i &= (i - 0.5)\Delta x_i; v_j = (j - 0.5)\Delta y_j; w_k = (k - 0.5)\Delta z_k \\ u_{i+1} &= (i + 0.5)\Delta x_i; v_{j+1} = (j + 0.5)\Delta y_j; w_{k+1} = (k + 0.5)\Delta z_k \end{aligned} \quad (10)$$

The expression for the center point is denoted as

$$X = \frac{FrGM_{100}^{\alpha_x \alpha_y \alpha_z}}{FrGM_{000}^{\alpha_x \alpha_y \alpha_z}}, Y = \frac{FrGM_{010}^{\alpha_x \alpha_y \alpha_z}}{FrGM_{000}^{\alpha_x \alpha_y \alpha_z}}, Z = \frac{FrGM_{001}^{\alpha_x \alpha_y \alpha_z}}{FrGM_{000}^{\alpha_x \alpha_y \alpha_z}}. \quad (11)$$

The expression for the fractional-order translational invariant center distance is expressed as

$$\eta_{pqr}^{\alpha_x \alpha_y \alpha_z} = \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \sum_{k=0}^{K-1} f(i, j, k) T_{pqr}^{\alpha_x \alpha_y \alpha_z}(x_i, y_j, z_k), \quad (12)$$

where

$$T_{pqr}^{\alpha_x \alpha_y \alpha_z} = \int_{x_i - \frac{\Delta x_i}{2}}^{x_i + \frac{\Delta x_i}{2}} \int_{y_j - \frac{\Delta y_j}{2}}^{y_j + \frac{\Delta y_j}{2}} \int_{z_k - \frac{\Delta z_k}{2}}^{z_k + \frac{\Delta z_k}{2}} (x - \hat{x})^{\alpha_x p} (y - \hat{y})^{\alpha_y q} (z - \hat{z})^{\alpha_z r} dx dy dz. \quad (13)$$

Based on the divisibility of the functional moments, Equation (10) simplifies to

$$T_{pqr}^{\alpha_x \alpha_y \alpha_z}(x_i, y_j, z_k) = IT_{Xp}^{\alpha_x}(x_i) IT_{Yq}^{\alpha_y}(y_j) IT_{Zr}^{\alpha_z}(z_k), \quad (14)$$

where

$$\begin{aligned} IT_{Xp}^{\alpha_x}(x_i) &= \frac{1}{\alpha_x p + 1} \left[(u_{i+1} - \hat{x})^{\alpha_x p + 1} - (u_i - \hat{x})^{\alpha_x p + 1} \right] \\ IT_{Yq}^{\alpha_y}(y_j) &= \frac{1}{\alpha_y q + 1} \left[(v_{j+1} - \hat{y})^{\alpha_y q + 1} - (v_j - \hat{y})^{\alpha_y q + 1} \right] \\ IT_{Zr}^{\alpha_z}(z_k) &= \frac{1}{\alpha_z r + 1} \left[(w_{k+1} - \hat{z})^{\alpha_z r + 1} - (w_k - \hat{z})^{\alpha_z r + 1} \right] \end{aligned} \quad (15)$$

The merged expressions are organized as follows:

$$\begin{aligned} T_{pqr}^{\alpha_x \alpha_y \alpha_z} &= \frac{1}{(\alpha_x p + 1)(\alpha_y q + 1)(\alpha_z r + 1)} \left[(u_{i+1} - \hat{x})^{\alpha_x p + 1} - (u_i - \hat{x})^{\alpha_x p + 1} \right] \\ &\times \left[(v_{j+1} - \hat{y})^{\alpha_y q + 1} - (v_j - \hat{y})^{\alpha_y q + 1} \right] \left[(w_{k+1} - \hat{z})^{\alpha_z r + 1} - (w_k - \hat{z})^{\alpha_z r + 1} \right] \end{aligned} \quad (16)$$

The 3D fractional-order moment invariants have rotational invariants, and the rotation of the matrix can be obtained as

$$R_{xyz}(\theta, \varphi, \psi) = \begin{pmatrix} \cos \varphi \cos \psi & \cos \varphi \sin \psi & -\sin \varphi \\ \sin \theta \sin \varphi \cos \psi - \cos \theta \sin \psi & \sin \theta \sin \varphi \sin \psi + \cos \theta \cos \psi & \cos \varphi \sin \theta \\ \cos \theta \sin \varphi \cos \psi + \sin \theta \sin \psi & \cos \theta \sin \varphi \sin \psi - \sin \theta \cos \psi & \cos \theta \cos \varphi \end{pmatrix}. \quad (17)$$

The rotation matrix is usually used for the linear transformation of the object coordinates, as shown below:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = R_{xyz}(\theta, \varphi, \psi) \begin{pmatrix} x - \hat{x} \\ y - \hat{y} \\ z - \hat{z} \end{pmatrix} = \begin{pmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{pmatrix} \begin{pmatrix} x - \hat{x} \\ y - \hat{y} \\ z - \hat{z} \end{pmatrix}, \quad (18)$$

where (R_{ij}) $1 \leq i \leq m$ $1 \leq j \leq n$ is an element of the matrix $R_{xyz}(\theta, \varphi, \psi)$.

The 3D moment invariants of fractional order are simply written as FrGMIs:

$$FrGMI_{pqr}^{\alpha_x \alpha_y \alpha_z} = \lambda^{-\gamma} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \sum_{k=0}^{K-1} f(i, j, k) \mu_{pqr}^{\alpha_x \alpha_y \alpha_z}(x_i, y_j, z_k), \quad (19)$$

where

$$\mu_{pqr}^{\alpha_x \alpha_y \alpha_z} = \int_{x_i - \frac{\Delta x_i}{2}}^{x_i + \frac{\Delta x_i}{2}} \int_{y_j - \frac{\Delta y_j}{2}}^{y_j + \frac{\Delta y_j}{2}} \int_{z_k - \frac{\Delta z_k}{2}}^{z_k + \frac{\Delta z_k}{2}} \begin{bmatrix} (R_{11}(x - \hat{x}) + R_{12}(y - \hat{y}) + R_{13}(z - \hat{z}))^{\alpha_x p} \\ (R_{21}(x - \hat{x}) + R_{22}(y - \hat{y}) + R_{23}(z - \hat{z}))^{\alpha_y q} \\ (R_{31}(x - \hat{x}) + R_{32}(y - \hat{y}) + R_{33}(z - \hat{z}))^{\alpha_z r} \end{bmatrix} dx dy dz. \quad (20)$$

The normalized parameters are expressed as

$$\begin{aligned} \lambda &= FrGM_{000}^{\alpha_x \alpha_y \alpha_z}, \gamma = 1 + \frac{\alpha_x n + \alpha_y m + \alpha_z p}{3} \\ \theta &= \frac{1}{2} \tan^{-1} \left(\frac{2\eta_{011}}{\eta_{020} + \eta_{002}} \right), \phi = \frac{1}{2} \tan^{-1} \left(\frac{2\eta_{101}}{\eta_{200} + \eta_{002}} \right) \\ \psi &= \frac{1}{2} \tan^{-1} \left(\frac{2\eta_{110}}{\eta_{200} + \eta_{020}} \right) \end{aligned} \quad (21)$$

The exact calculation of fractional-order 3D moment invariants is not possible and can only be approximated, but due to the divisibility of the function moments, it is difficult to reduce the triple integral $\mu_{pqr}^{\alpha_x \alpha_y \alpha_z}$ to a simple integral using the numerical integration algorithm and the 3D Gaussian integration method.

Three-dimensional Gaussian integration is a way to integrate functions in three-dimensional space by dividing the three-dimensional space into infinitely small volume elements and accumulating the values of functions within the volume element multiplied by the volume of that volume element.

Gaussian integration is capable of solving various types of orthogonal polynomials, including Legendre, Chebyshev, Laguerre, Hermite, and others. However, there are still computational errors, and the integration interval should not be too large, otherwise the results will be inaccurate. In addition, the error of Gaussian integration is also related to the smoothness of the product function, and the worse the smoothness of the product function, the larger the error. Therefore, the Gaussian integral is the formula with the highest algebraic accuracy for a given number of nodes.

It can enhance the computational efficiency of fractional-order 3D moment invariants under the condition that the fractional parameters are satisfied, and it can accurately perform the calculation and achieve a desirable result for different types of 3D images, regardless of whether they are processed by rotation, scaling, translation transformation, noise processing, or filtering.

2.1.3. Fractional Chebyshev Moment Invariants

Fractional order can perform image recognition of 3D objects, and common moments can only achieve rotational transformation. Since Chebyshev moments add the two transformations of scaling and translation, fractional order is fused with Chebyshev moments to form fractional-order Chebyshev moments. By normalizing the transformed 3D objects and obtaining the rotational invariants of the parameters, the computing efficiency is improved under the premise of satisfying the parameter accuracy to achieve high efficiency in 3D image recognition.

The digital image strength function $f(i, j, k)$ is weighted to obtain the weighted image intensity function $\tilde{f}(i, j, k)$. FrCMs can be described as

$$\begin{aligned} FrCM_{nmp}^{\alpha} &= \int_0^1 \int_0^1 \int_0^1 \tilde{f}(i, j, k) \tilde{F}_n^{\alpha_x}(x_i) \tilde{F}_m^{\alpha_y}(y_j) \tilde{F}_p^{\alpha_z}(z_k) dx dy dz \\ &= \frac{1}{\sqrt{d_{n,\alpha_x}^2 d_{m,\alpha_y}^2 d_{p,\alpha_z}^2}} \int_0^1 \int_0^1 \int_0^1 f(i, j, k) \tilde{F}_n^{\alpha_x}(x_i) \tilde{F}_m^{\alpha_y}(y_j) \tilde{F}_p^{\alpha_z}(z_k) dx dy dz \end{aligned} \quad (22)$$

where

$$\tilde{f}(i, j, k) = [w_{\alpha_x}(x) \quad w_{\alpha_y}(y) \quad w_{\alpha_z}(z)]^{-1/2} f(i, j, k). \quad (23)$$

$$FrCM_{nmp}^{\alpha} = \frac{1}{\sqrt{d_{n,\alpha_x}^2 d_{m,\alpha_y}^2 d_{p,\alpha_z}^2}} \sum_{l=0}^n \sum_{s=0}^m \sum_{r=0}^p B_{n,l} B_{m,s} B_{p,r} FrGM_{pqr}^{\alpha_x \alpha_y \alpha_z}, \quad (24)$$

where $FrGM_{pqr}^{\alpha_x\alpha_y\alpha_z}$ is the fractional geometric moment. Chebyshev moment invariants can be denoted as

$$FrCMI_{nmp}^{\alpha} = \frac{1}{\sqrt{d_{n,\alpha_x}^2 d_{m,\alpha_y}^2 d_{p,\alpha_z}^2}} \sum_{l=0}^n \sum_{s=0}^m \sum_{r=0}^p B_{n,l} B_{m,s} B_{p,r} FrGM_{lps}^{\alpha_x\alpha_y\alpha_z}, \quad (25)$$

where the denominators of α_x , α_y , and α_z are odd.

2.2. FrCMs-DNNs Model

DNNs, the multilayer unsupervised neural networks, systematically map features layer by layer to acquire an improved representation of the input. These networks incorporate a range of nonlinear mapping feature transformations for handling highly complex functions. Viewing the deep structure as a neuronal network, the core idea of a deep neural network can be succinctly described as follows:

- (1) Pre-training the network with unsupervised learning methods;
- (2) Layer-by-layer training using unsupervised learning;
- (3) Fine-tuning the network model with supervised learning.

DNNs are constructed upon the foundational perceptron model, which is a multiple-input single-output model. This model learns a linear relationship between inputs and outputs to generate the desired outputs.

$$z = \sum_{i=1}^m w_i x_i + b, \quad (26)$$

where w is the parameter of weight coefficient and b is the bias amount.

The output result is obtained following the neuronal activation function:

$$\text{sign}(z) = \begin{cases} -1, & z < 0 \\ 1, & z \geq 0 \end{cases}. \quad (27)$$

Although the structure of DNNs is very complex, its essence is still a perceptual system. The algorithm starts from any input layer, runs from left to right, and obtains a result at the last output layer. When the calculation results deviate greatly from the target value, the errors of each node are inverted from right to left, and the total weight of each node is modified. After reaching the input layer in reverse, the operation continues and repeats until each weight reaches an appropriate value. Compared with the traditional mathematical analysis, some parameters of this kind of differential equation adopt the mode of random selection and finally make it more accurate by modifying it.

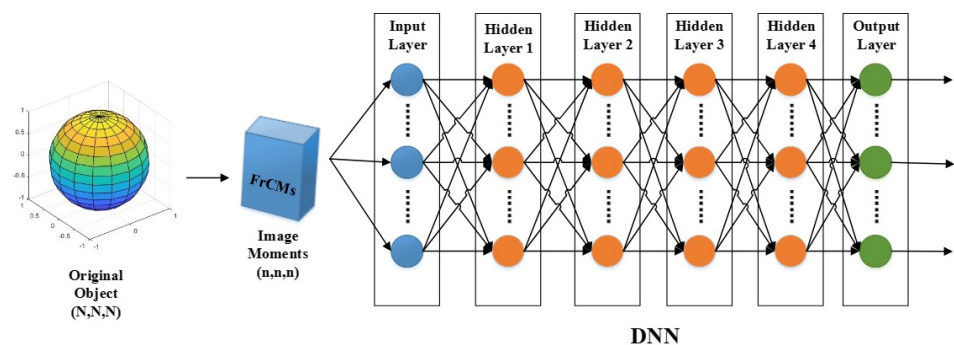
The algorithm can avoid the prior knowledge of each level of information and improve its performance. At the same time, its quantitative level enables the algorithm to deeply learn the distributed information and improve its effectiveness. Compared with the shallow model, the depth model can better describe the real information, with stronger details and better description ability, so that it can better identify the image effectively.

As the same with the architecture of DNNs, the structure of FrCMs-DNNs contains three layers, an input, hidden layers, and an output that are fully connected. By adjusting the weights and biases, FrCMs-DNNs achieves an output with the expected accuracy relative to the network input. Table 1 contains a detailed description of the FrCMs-DNNs system.

Table 1. The detailed description of the FrCMs-DNNs's structure.

Input Layers	Input Moment Vector	Activators $N \times N \times N$
1	Full Connection + BN + ELU + Dropout	100
2	Full Connection + BN + ReLU + Dropout	165
3	Full Connection + BN + ReLU + Dropout	245
4	Full Connection + BN + ReLU + Dropout	120
Output	Softmax	Quantity Subjects

As depicted in Figure 1, the FrCMs-DNNs model is computationally efficient and has a small memory requirement, and it is suitable for use in the ABC optimizer algorithm [30] for problems with a large amount of data and parameters.

**Figure 1.** The model structure of FrCMs-DNNs for 3D object classification.

The FrCMs-DNNs model includes an input, hidden layers, and an output. The output from the softmax layer corresponds to the number of classification labels. The input is expressed in terms of 3D FrCMs, and the descriptor vector consists of order r of 3D FrCMs, with r set by the experiment. The input vector can be expressed as

$$V = [FrCM_{nmp}^{\alpha} | n \times m \times p \in [0, 1, \dots, r]]. \quad (28)$$

When the maximum order of $20 \times 20 \times 20$ is used, V denotes 8000 dimensions. The dataset is categorized into two parts, the training and test sets. The hidden layers are used in the model with four hidden layers, containing 100, 165, 240, and 120 neurons.

$$Y_i = \eta_i(b^{(i)} + W^{(i)}Y_i). \quad (29)$$

Y_i represents the output of the hidden layer i , $W^{(i)}$ denotes the weight coefficient matrix, $b^{(i)}$ denotes the bias vector, and η_i denotes the activation function.

The softmax function is an extension of the logic function:

$$\text{softmax}(y^j) = \left(\frac{e^{y^j}}{\sum_{i=1}^s e^{y^i}} \right), j = 0, 1, \dots, s. \quad (30)$$

The output of the model can be calculated as

$$f(V) = \text{softmax}(b^{(5)} + W^{(5)}Y_4). \quad (31)$$

BN: Batch normalization improves the learning rate, accelerates training, and avoids divergence and overfitting.

ELU: The exponential linear unit brings the average value of the activation function close to zero to speed up the learning. It enables avoidance of the problem of gradient disappearance.

ReLU: The rectified linear unit can be defined as $f = \max(0, x)$; it is insensitive to the gradient vanishing problem and improves the convergence speed.

Softmax function: The softmax function compresses a vector z of real numbers into another real vector $\sigma(z)$, ensuring that each element falls within the range $(0, 1)$ and the sum of the elements is 1.

Fractional-order Chebyshev moments and DNNs for 3D image recognition are shown in Table 2, including their main aspects, common datasets, characteristics, evaluation index, advantages, and limitations.

Table 2. Three-dimensional recognition using DNNs and fractional-order Chebyshev moments.

Model	Main Aspects	Datasets	Evaluation Index	Advantages	Limitations
3D-CNN [31]	3D object recognition	ModelNet40	Recognition accuracy	High identification accuracy	The model is sensitive to local feature learning, occlusion, and attitude changes
3D-encoder-predictor CNNs and shape synthesis [32]	DNN for 3D model classification	ModelNet40	Classification accuracy	High classification accuracy; high training complexity	The model is sensitive to small datasets and noise
FrCMs [33]	FrCMs combined with DNNs for 3D recognition	ShapeNet	Classification accuracy and robustness	FrCMs provide better characterization, which reduces risk of overfitting	Selection and adjustment of FrCM parameters are more complex
FrCMs [34]	3D shape analysis	ModelNet	Segmentation accuracy and computational efficiency	The fractional order can accommodate a wide range of data distributions; it enhances the robustness	The computational cost of FrCMs is relatively high
FrCMs combined with 3D-CNN [35]	3D shape analysis	ShapeNet	Segmentation accuracy robustness	FrCMs enhance the understanding of shape structure; 3D-CNN extracts higher-level features	The selection and adjustment of FrCM parameters is complicated
DNN–FrCMs joint optimization [36]	Joint optimization of DNNs and FrCMs	3DShapeNet	Overall performance indicators	Combines the powerful modeling capabilities of DNNs with the feature extraction advantages of FrCMs	Requires significant computational resources for training
3D-CNN in association with FrCMs [37]	Application of fractional-order features to 3D scene understanding	Sun RGBD	Semantic segmentation accuracy	Comprehensive use of deep learning and fractional features; good adaptability to complex scenes	It takes a lot of computing resources to train

3. Experiment

In this section, we designed three experiments from the perspectives of the effectiveness of FrCMs, 3D recognition ability, and practical application value, respectively. Firstly, fractional moments and 3D reconstruction play an important role in image feature tasks, and the relationship between them reflects the model's ability to retain and reconstruct image information. Therefore, image reconstruction experiments are designed to evaluate the reconstruction results through MSE and PSNR indicators and analyze the evaluation model's ability to extract image information effectively. Secondly, to prove that FrCMs-DNNs has certain advantages in multiscale feature extraction, global feature learning, robustness, and generalization ability in 3D recognition tasks, comparison and ablation experiments for FrCMs-DNNs are designed, and different evaluation indicators are used to

quantitatively analyze the effectiveness of the method. Finally, to prove that FrCMs-DNNs is good at extraction and feature representation in image recognition tasks and has strong adaptability and expression ability, the recognition experiment based on SAR image is designed to analyze and verify the universality and robustness of the method.

3.1. 3D Image Reconstruction

Experiment 1: Image reconstruction was carried out with different fractional-order moments.

Experiment 2: Different moment order parameters were used for feature extraction.

In this paper, MSE and PSNR, commonly used evaluation indexes, are used to compare the quality of images reconstructed with different fractional moments.

For the image whose original image $f(x, y, z)$ and reconstructed image $\hat{f}(x, y, z)$ size are both $N \times M \times K$, the mean square error is defined as:

$$MSE = \frac{1}{NMK} \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} \sum_{z=0}^{K-1} \|f(x, y, z) - \hat{f}(x, y, z)\|^2. \quad (32)$$

The MSE is the average of the differences between the two images, but as the quantization number increases, the MSE becomes larger. Hence, a smaller MSE value indicates better image reconstruction quality.

PSNR, defined by MSE, is frequently employed as a metric for signal reconstruction quality, particularly in fields like image compression. When the gray level of the image is set as L (8-bit gray level image L is 255), then

$$PSNR = 10 \cdot \lg \frac{L^2}{MSE}. \quad (33)$$

The value range of $PSNR$ is $(0, +\infty)$. A larger value of $PSNR$ means a better performance.

According to Figure 2, when the order of the matrix is 40, most of the fractional-order moments have a good reconstruction effect in 3D ant images. And as shown in Figure 3, with the increase in order, the MSE values of FrCMs, FrOLMs, and FrGLMs gradually approach 0, and their PSNR values gradually increase. When the order reaches the highest order of 150, the PSNR values of FrOFMMs [38] and FrZMs are 0, the PSNR value of FrCMs is 44, the PSNR value of FrOLMs is 37, and the PSNR value of FrGLMs is 31.

	Moments order (n, m, k)							
	(10,10,10)	(20,20,20)	(30,30,30)	(40,40,40)	(50,50,50)	(80,80,80)	(100,100,100)	(150,150,150)
FrCMs $\alpha_x = 1.2$ $\alpha_y = 1.0$ $\alpha_z = 1.2$								
FrOFMMs $\alpha = 1.2$								
FrZMs $\alpha = 1.2$								
FrOLMs $\alpha = 1.2$								

Figure 2. Cont.

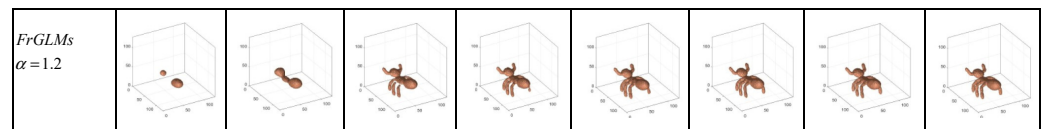


Figure 2. Comparison of various 3D image reconstruction results.

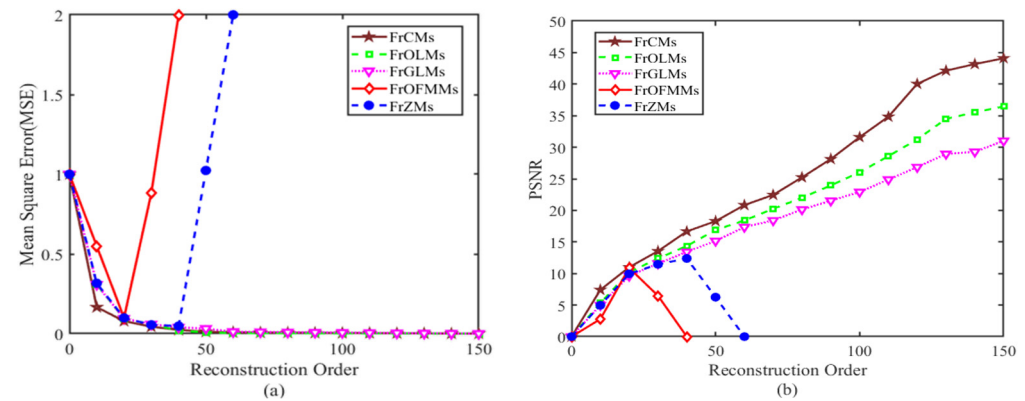


Figure 3. Comparison curves of different reconstructed methods from 3D image: (a) MSE; (b) PSNR.

Image reconstruction results are applied to evaluate the performance of the presented method. Figures 2 and 3 show the comparison of the image reconstruction results of FrCMs, FrOLMs, FrGLMs, FrOFMMs, FrZMs, and FrGLMs. With increasing fractional-order moments, the effect of the reconstructed image approaches that of the raw image. The reconstruction result of FrCMs is the best when the parameters are chosen as $\alpha_x = 1.2$, $\alpha_y = 1$, and $\alpha_z = 1.2$, which verifies that the fractional Chebyshev moments can effectively achieve visual reconstruction.

3.2. Feature Extraction

The image reconstruction was performed using the “bird” 3D image dataset, which is a combination of the original bird images. Multiple reconstruction experiments of bird images using different fractional parameters were conducted, which can verify the local image extraction ability of FrCMs.

Figure 4 shows the feature extraction results of 3D images by FrCMs with different α_x , α_y , α_z , and maximum orders. The FrCMs accurately represent the image information and show the ability of efficient local feature extraction. The approximate error of the FrCMs is smaller than the existing moments, and the advantages are obvious in 3D image reconstruction.

3.3. 3D Object Recognition

The PSB dataset [39] and the medical images dataset [40] are used to validate the effectiveness of 3D image recognition.

Experiment 1: As shown in Figure 5, 20 objects of different categories are selected in the PSB dataset, and as shown in Figure 6, 12 objects of different categories are selected in the medical images dataset, which can verify the recognition capability of FrCMs (fractional-order Chebyshev moment invariants), respectively.

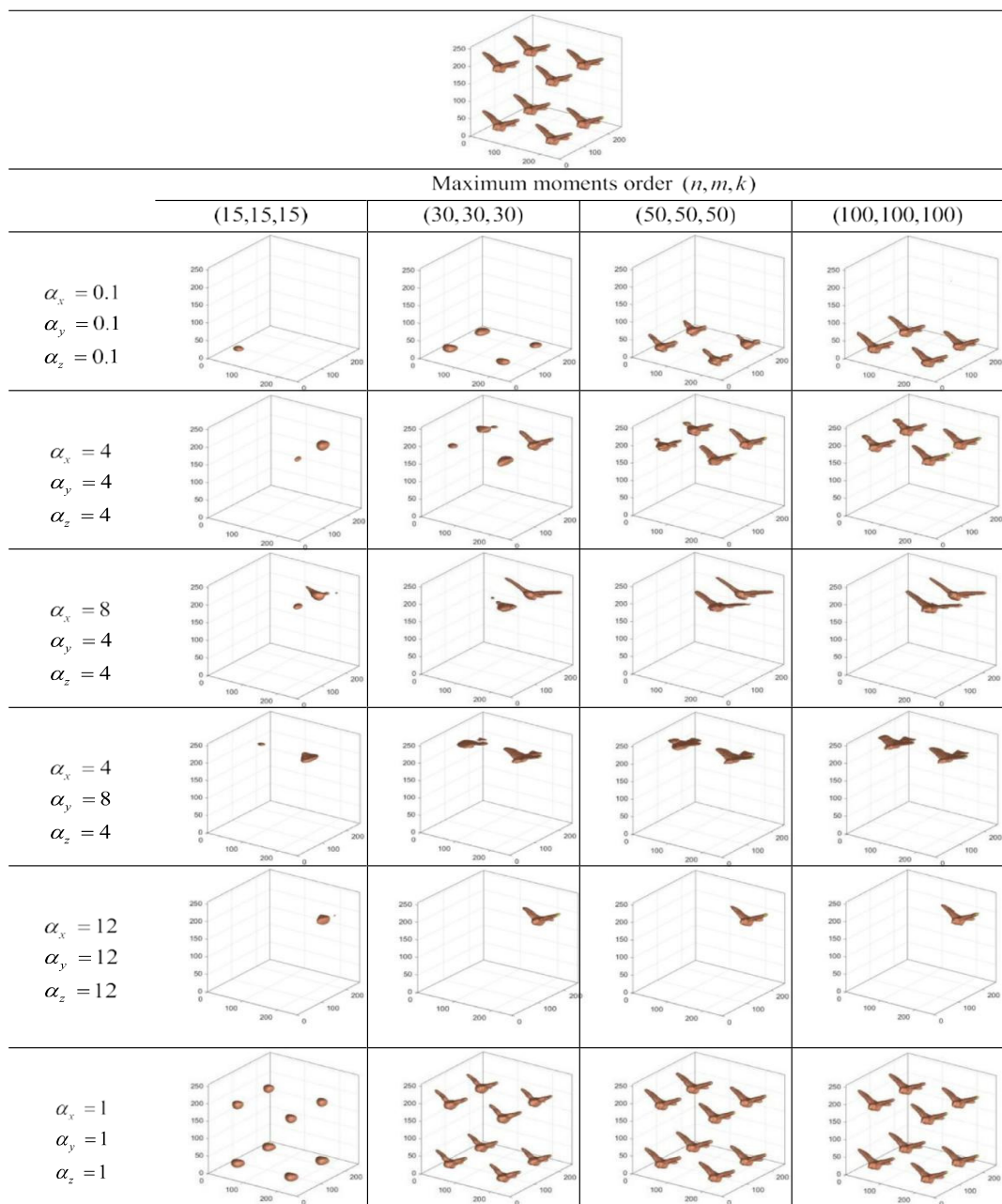


Figure 4. The local feature extraction results of FrCMs.

The noise robustness of the FrCMs is tested by rotating, scaling, and translating the dataset objects with transformations. The classification accuracy of the FrCMs is evaluated by adding various densities of Gaussian noise. The performance of FrCMs in 3D object classification is compared with that of FrFMMs, FrZMs, FrLMs, and GMs. The fractional parameters of FrCMs are set up as follows:

- (1) $\alpha_x = 1.4, \alpha_y = 1.4, \alpha_z = 1.4$; (2) $\alpha_x = 1.4, \alpha_y = 1.0, \alpha_z = 0.8$;
- (3) $\alpha_x = 0.8, \alpha_y = 1.4, \alpha_z = 1.0$; (4) $\alpha_x = 0.8, \alpha_y = 0.8, \alpha_z = 0.8$.

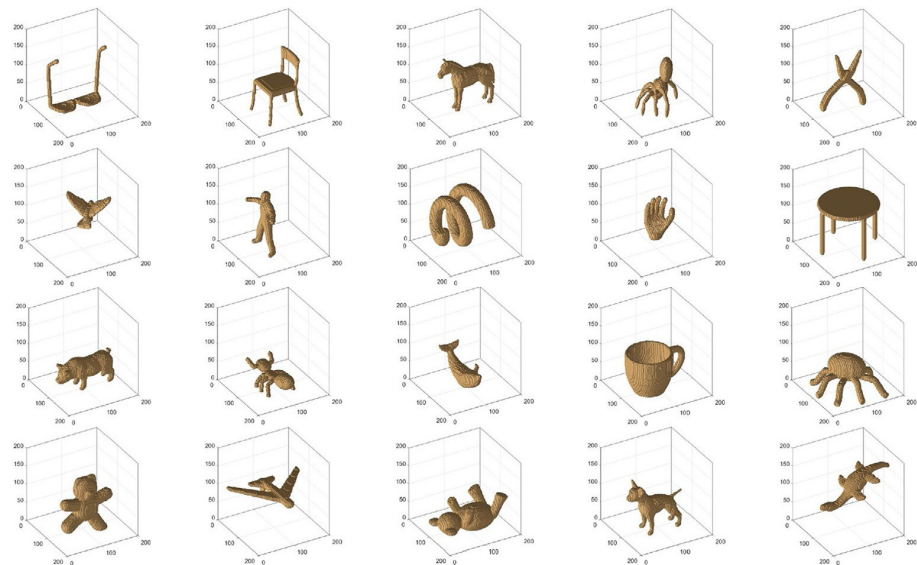


Figure 5. The PSB dataset.

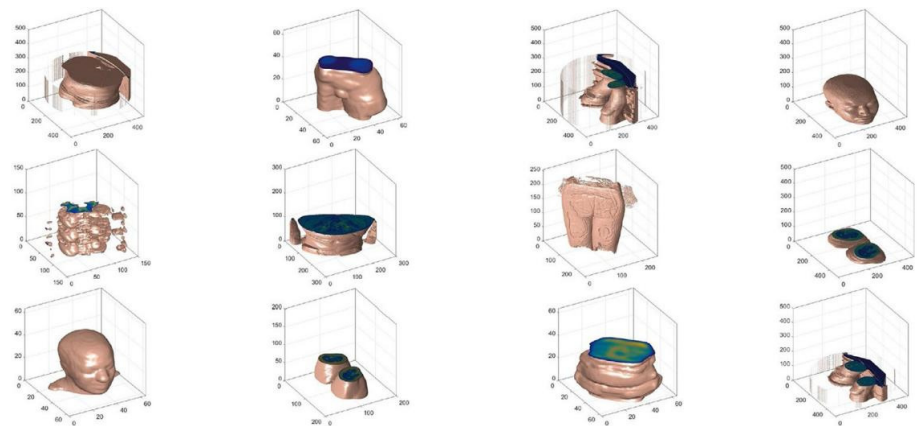


Figure 6. The medical images dataset.

Fractional-order moment invariants are used to process 3D objects with different densities of Gaussian noise on the PSB dataset and medical image dataset. Gaussian noise is caused by the random noise of an image sensor. It is random and follows Gaussian distribution. It will make the brightness and color of the image have slight random changes, as well as cause blurring and distortion.

In the image processing experiments, the Gaussian noise density is chosen to be at most 10%, and according to the previous experimental rules, Gaussian noise with a density of 1–5% will be selected for experimental verification, as shown in Figures 7 and 8, for the PSB dataset and medical dataset, with different densities of Gaussian noise processing for 3D object recognition.

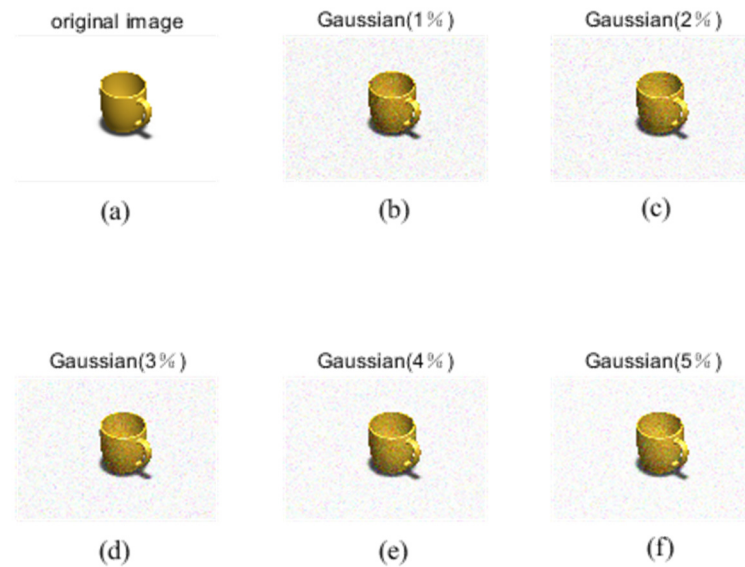


Figure 7. Gaussian noise plots for different densities of the PSB dataset.

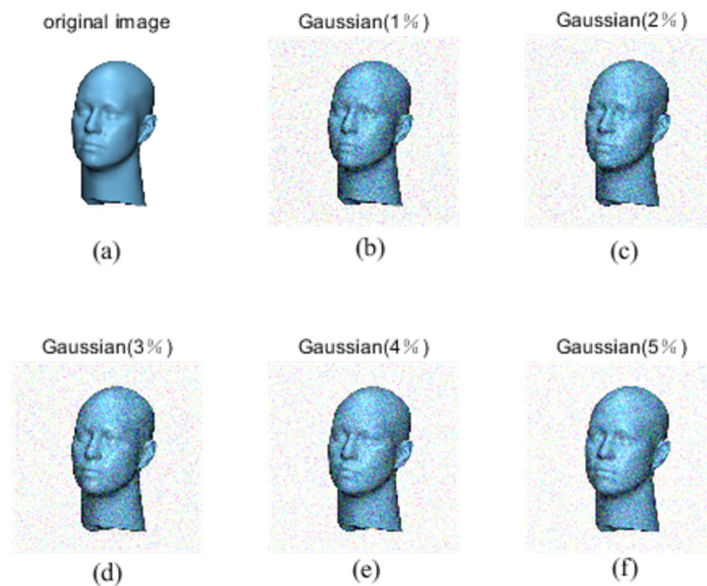


Figure 8. Gaussian noise plots for different densities of medical image dataset.

Tables 3 and 4 depict the object recognition rate results for the PSB dataset and the medical images dataset, respectively. The description of the followed methods is shown in Appendix A.

Table 3. Comparison of object recognition rates for the PSB dataset.

Moment Invariance	Noiseless	Gaussian Noise (1–5%)					Mean Value
		1%	2%	3%	4%	5%	
FrCMIs(1)	99.95	71.72	68.50	59.65	48.12	39.10	57.418
FrCMIs(2)	99.88	72.20	68.79	59.89	49.53	39.74	58.03
FrCMIs(3)	99.97	73.50	68.92	60.50	49.33	40.80	58.61
FrCMIs(4)	99.98	72.42	69.57	59.96	49.76	40.68	58.478
FrFMMIs	80.38	58.60	40.15	35.90	30.10	20.92	37.134
FrLMIs	98.30	66.30	58.95	50.25	36.32	30.87	48.538
FrGLMIs	97.87	65.79	57.90	53.48	35.92	29.75	48.568
FrZMIs	82.55	59.56	46.28	42.41	34.44	26.25	41.788
GMI	75.60	33.23	23.35	18.85	16.70	14.57	21.34

Table 4. Comparison of object recognition rates for the medical images dataset.

Moment Invariance	Noiseless	Gaussian Noise (1–5%)					Mean Value
		1%	2%	3%	4%	5%	
FrCMIs(1)	99.35	66.15	54.76	42.71	36.57	32.76	46.59
FrCMIs(2)	99.92	67.36	56.45	45.45	35.70	33.84	47.76
FrCMIs(3)	99.92	67.40	57.25	46.89	40.05	37.16	49.75
FrCMIs(4)	99.37	69.57	55.01	44.52	41.10	36.16	49.272
FrFMMIs	79.90	46.85	34.59	30.85	27.02	21.35	32.132
FrLMIs	97.75	56.24	53.65	40.85	39.90	30.58	44.244
FrGLMIs	96.74	53.56	50.37	42.45	38.70	31.45	43.306
FrZMIs	81.90	45.02	34.15	30.91	29.10	22.12	32.26
GMIIs	74.35	34.35	25.47	18.35	17.05	15.33	22.11

As shown in Tables 3 and 4, there is a difference between the object recognition rate presented by adding Gaussian noise and no noise in the 3D image, and the recognition effect of the 3D object treated with no noise is close to the exact recognition compared with the 3D object treated with Gaussian noise; according to the different densities of the 3D object treated with Gaussian noise, the higher the density, the greater the error caused by the recognition of the 3D object in the recognition process. The higher the density, the greater the error caused in the recognition process, resulting in a lower recognition rate in the end; by comparing different types of fractional-order moment invariants in the recognition process of 3D objects without noise and Gaussian noise, it can be seen that the proposed FrCMIs recognition effect is the best, and the recognition rate is the highest among all fractional-order moment invariants.

Experiment 2: The datasets are constructed by performing a series of transformations on selected objects. The PSB dataset includes 10 categories: airplane, ant, bird, cup, fish, hand, octopus, spider, glasses, and teddy bear. The medical images dataset includes 5 categories: head, abdomen, hip, knee, and leg. In the classification task of the PSB dataset, 240 (40%) objects are randomly selected as the training set, and 360 (60%) objects are selected as the test set. In the classification task of the medical images dataset, 150 (50%) objects are randomly selected as the training set, and 150 (50%) objects are selected as the test set.

As depicted in Figure 9, the recognition accuracy of FrCMs-DNNs is higher than FrCMIs, and the FrCMs-DNNs model has the best classification results. Figure 10 presents the confusion matrix of the fractional Chebyshev moments models for the PSB dataset [35] and the medical images dataset [36]. Most of the confused objects are nearly completely recognized, with a little amount of confusion between the bird/airplane and ant/octopus categories in the PSB dataset, and a small amount of confusion between the abdomen/hip and knee/leg categories in the medical images dataset, since these categories have similar shapes.

The FrOLMs, FrOFMMs, FrZMs, FrCMs, and FrGLMs are adopted as input layers to demonstrate the classification capabilities of the FrCMs unit, respectively. The accuracy of corresponding classification results is obtained by adding the order of different fractional-order moments. As shown in Figure 11, compared with other methods combining fractional-order moments and DNNs, FrCMs-DNNs has the highest object recognition accuracy.

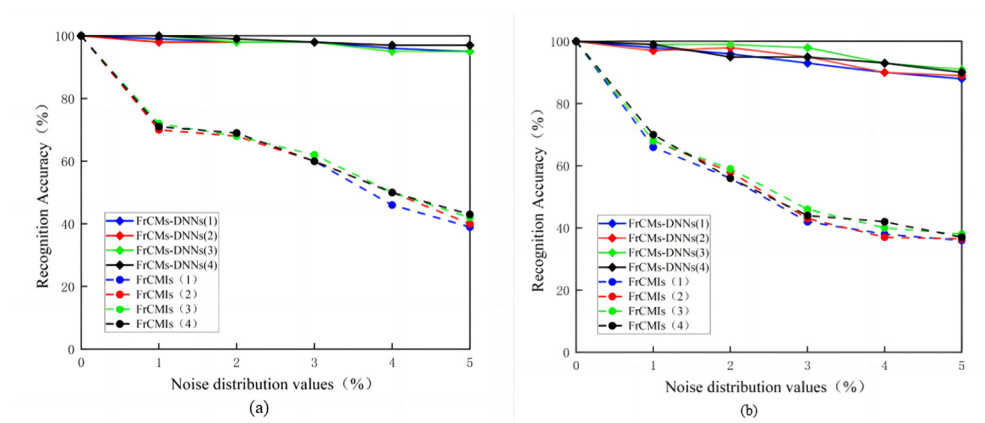


Figure 9. Comparison of the recognition accuracy of FrCMs-DNNs and FrCMIs with different score parameters: (a) the PSB dataset; (b) the medical images dataset.

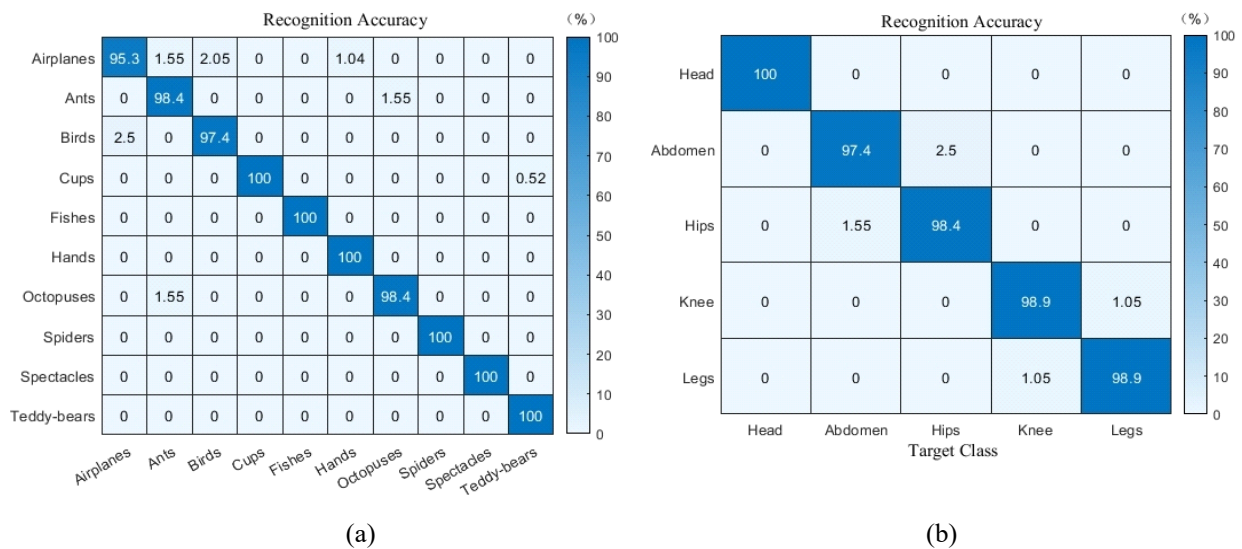


Figure 10. The confusion matrix of object recognition: (a) the PSB dataset; (b) the medical images dataset.

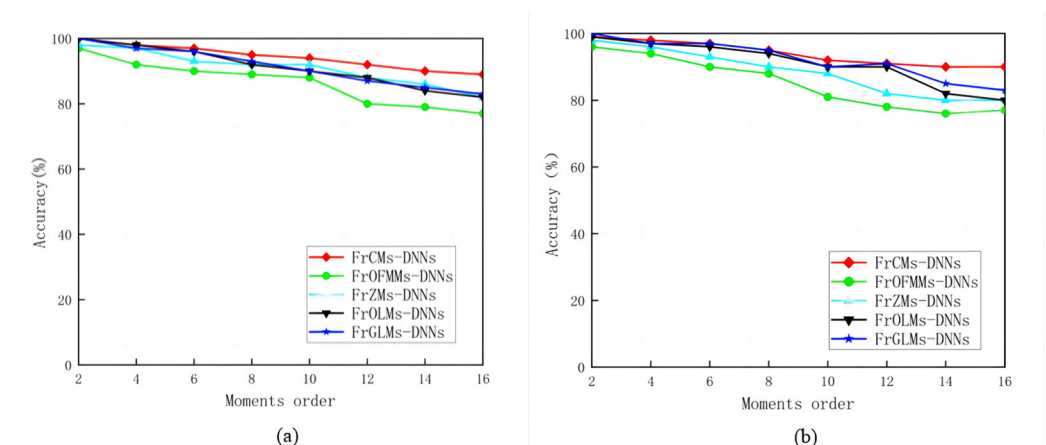


Figure 11. Comparison of the recognition accuracy: (a) the PSB dataset; (b) the medical images dataset.

3.4. Ablation Experiment

The FrCMs-DNNs model consists of FrCMs and DNNs. In order to demonstrate the effectiveness of each module of the FrCMs-DNNs, the accuracy of FrCMs-DNNs was verified by ablation experiments.

In this experiment, the performance of the FrCMs-DNNs model applied to 3D object recognition was verified. In this experiment, the FrCM and DNN models used for 3D object recognition were taken as the benchmark model, and the PSB dataset and medical image dataset were chosen as the experimental datasets to verify the accuracy and universality of the FrCMs-DNNs model.

3.4.1. Evaluation Methods and Indicators

In order to avoid the phenomenon of high accuracy due to the presence of all normal prediction sample data in the test data, $A_{Accuracy}$, precision rate $P_{Precision}$, recall rate R_{Recall} , and F_1 -score are selected in the experiment. Here are the calculation formulae:

$$A_{Accuracy} = \frac{T}{N_{ALL}}, \quad (34)$$

$$P_{Precision} = \frac{T_p}{T_p + F_p}, \quad (35)$$

$$R_{Recall} = \frac{T_p}{T_p + F_N}, \quad (36)$$

$$F_1 = \frac{2 \times R_{Recall} \times P_{Precision}}{R_{Recall} + P_{Precision}}, \quad (37)$$

where N_{ALL} is the total sample number; T is the number of samples that are predicted correctly; T_p is the normal and predicted normal sample number; F_p is the number of samples where normal prediction is abnormal; and F_N is the number of samples where the abnormal prediction is normal.

The accuracy value can directly reflect the overall accuracy of the method. The precision rate and recall rate can reflect whether the model is in the overfitting state. A low accuracy rate indicates that the model is biased to output abnormal labels, while a low recall rate indicates that the model is biased to output normal labels. The F_1 -score comprehensively reflects the precision rate and recall rate, and the higher the F_1 -score is, the better the model fitting effect is.

3.4.2. Ablation Experiment and Analysis

As is shown in Figure 12, Experiment 1 compared the detection effect of the FrCM, DNN, and FrCMs-DNNs models based on the PSB dataset, set the same parameters for their models, and compared the 3D object recognition to assess the performance of FrCMs-DNNs. In the process of training, the method of a 5-fold crossover experiment was used to verify the results.

As is shown in Figure 13, Experiment 2 compared the universality of the FrCM, DNN, and FrCMs-DNNs models based on the medical image dataset, set the same parameters of their models, and used the undersampling method to avoid the imbalanced data phenomenon. The dataset was also trained by the method of five-fold cross-validation.

In summary, according to the results of Experiment 1, with the same parameters, the accuracy of FrCMs-DNNs is improved by 0.07 and 0.05, respectively, compared with FrCMs and DNNs. Its precision rate, recall rate, and F_1 -score are also among the leading levels. Note that FrCMs-DNNs takes into account the advantages of both FrCM and DNN models. Compared with a single model, the performance is significantly improved.

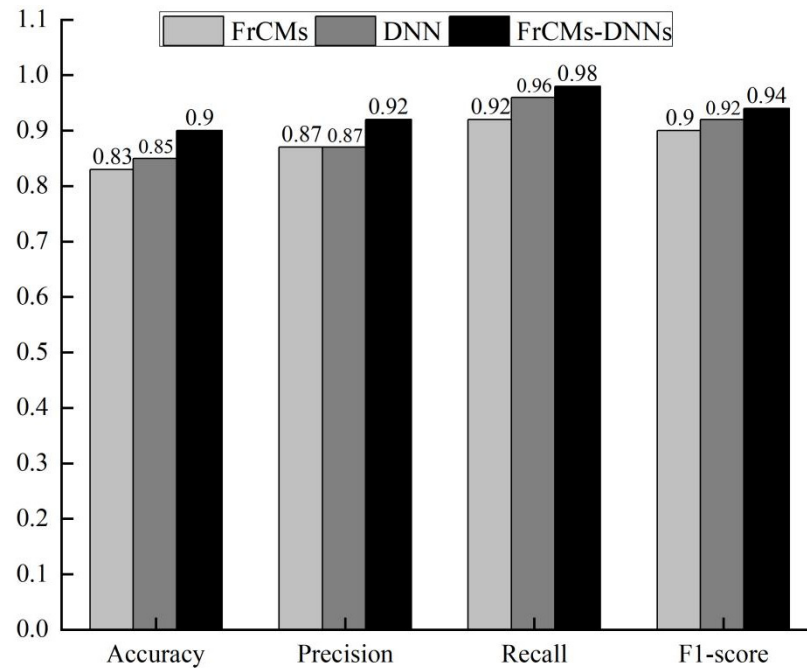


Figure 12. PSB dataset 3D object recognition results.

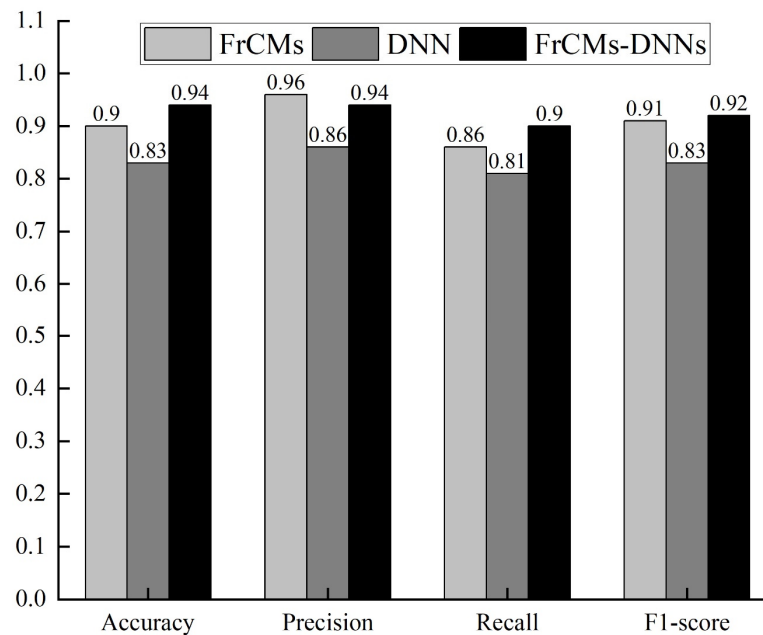


Figure 13. 3D object recognition results from medical images dataset.

According to the results of Experiment 2, in the model migration experiment from the PSB dataset to medical image dataset, the FrCMs-DNNs model still maintains a high performance advantage, and the accuracy and F1-score are 0.04 and 0.01 higher than the FrCMs model, respectively. It can be concluded that the FrCMs-DNNs model has high accuracy.

In Experiment 3, to evaluate the computational efficiency of FrCMs-DNNs, the model will be validated for computational efficiency on CPUs and GPUs using the PSB dataset and the medical images dataset, and the computer configurations chosen are RTX 2080Ti GPUs and an Intel(R) Xeon Silver 4112 CPU@2.60GHz. To fully evaluate the performance of the approach in a different hardware environment, the computational efficiency of FrCMs is evaluated by floating point operations per second (FLOPs), training or inference speed (frame per second, FPS), and Top-1/Top-5 accuracy (%), as shown in Tables 5 and 6.

Table 5. Comparison of CPU computing efficiency of FrCMs-DNNs model.

Datasets	#.Param.	CPU FLOPS	Training	Inference	Top-1	Top-5
PSB	24.40 M	3.86 G	1024 FPS	1850 FPS	75.20	92.30
Medical Images	24.40 M	3.87 G	958 FPS	1680 FPS	77.52	93.35

Table 6. Comparison of GPU computing efficiency of FrCMs-DNNs model.

Datasets	#.Param.	CPU FLOPS	Training	Inference	Top-1	Top-5
PSB	24.40 M	7.34 G	2538 FPS	3905 FPS	78.95	94.55
Medical Images	24.40 M	7.35 G	2365 FPS	3685 FPS	78.52	93.89

As can be seen from Tables 5 and 6, FrCMs-DNNs was implemented on different performance hardware for 3D image recognition with the same network parameters. The experimental results demonstrate that on GPU, the FLOPS value is approximately twice that of CPU, and the speed and accuracy of GPU training is better compared with CPU, which improves the computational efficiency more effectively; thus, the FrCMs-DNNs model can be better implemented in specific hardware environments. Therefore, for 3D image recognition, the FrCMs-DNNs model has better computational efficiency.

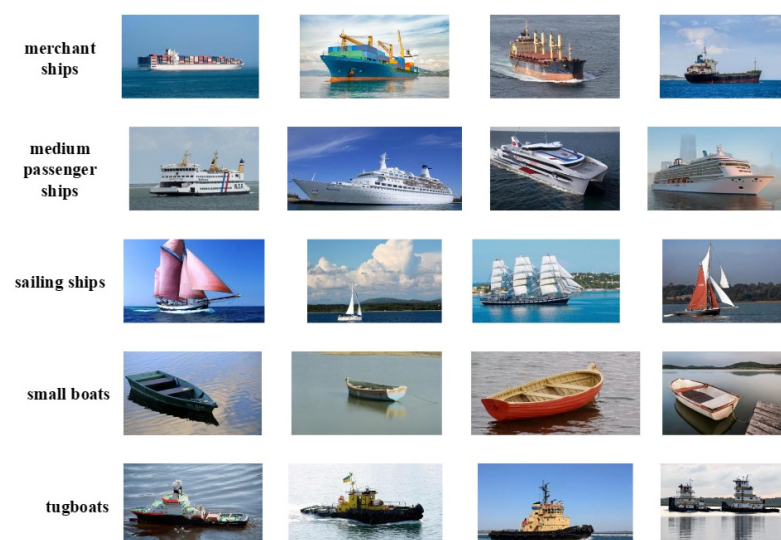
3.5. SAR Image Recognition

The feasibility of the FrCMs-DNNs model with high speed and accurate recognition is verified in the SAR image classification and detection experiments.

3.5.1. SAR Image Ship Classification

Synthetic aperture radar (SAR) is a remote sensing technology that uses radar signals and signal processing techniques to create high-resolution radar images. With synthetic aperture technology, SAR systems can achieve very-high-resolution imaging that can provide detail-rich images of the target and obtain high-quality images.

FrCMs-DNNs is applied to SAR images using the public VAIS ship dataset [41], as shown in Figure 14, which has 1088 images, mainly including 6 coarse-grained categories, 5 categories of which are selected as merchant ships, medium passenger ships, sailing ships, small boats, and tugboats. In this paper, 477 images are randomly chosen for training, while 473 images are designated for testing.

**Figure 14.** Category example images from the VAIS ship dataset.

In this experiment, to validate the performance of FrCMs-DNNs, which can be applied in image classification, fractional-order Chebyshev moments and deep neural network

features are fused for SAR ship classification. FrCMs-DNNs is used for SAR image ship classification on the VAIS ship dataset, and the feasibility and robustness of its model are verified by commonly used evaluation methods and metrics.

In order to avoid the phenomenon that the prediction sample data in the test data are all normal and the accuracy is too high, the above-mentioned accuracy A_{Accuracy} , inspection accuracy $P_{\text{Precision}}$, recall R_{Recall} , and F1-score are selected for the experiment. The results are presented in Table 7.

Table 7. Accuracy, precision, recall, and F1-score for different categories of VAIS ship dataset.

Evaluation Index	Merchant Ships	Medium Passenger Ships	Sailing Ships	Small Boats	Tugboats
Accuracy	0.8539	0.8395	0.9191	0.8497	0.9500
Precision	0.8950	0.7508	0.8652	0.8400	0.5000
Recall	0.8539	0.5300	0.9191	0.8497	0.9500
F1-score	0.8725	0.7175	0.8920	0.8440	0.6550

As evident from the table, FrCMs-DNNs is effective and robust for image recognition in SAR image ship classification using the VAIS ship dataset for the classification test, and the results show that the accuracy is above 80% and the model fits well when using the FrCMs-DNNs model as a feature vector to overcome the sensitivity of SAR images to orientation and effectively improve SAR image ship classification.

3.5.2. High-Speed SAR Image Ship Detection

To validate the effectiveness of the proposed FrCMs-DNNs, high-resolution SAR images are employed for comparison. The model is compared with the mask attention interaction and scale enhancement network, grid convolutional neural network, and depthwise separable convolutional neural network.

The experimental dataset is all the 3000 SAR images containing ships within the area of Dalian Port, as required for high-speed SAR ship detection, including high-speed SAR ship images and corresponding labels (ship's position, bounding box, etc.). Then, the dataset is pre-processed, including the operations of image denoising, normalization, cropping, etc., to ensure the quality and consistency of the input data. The effectiveness of the model is measured by evaluating the methods and metrics: accuracy A_{Accuracy} , accuracy check $P_{\text{Precision}}$, recall R_{Recall} , and F1-score.

From Figures 15 and 16, it can be seen that the high-speed SAR ship detection results of FrCMs-DNNs are better than those of the Mask attention interaction and scale enhancement networks; the grid convolutional neural network and the depthwise separable convolutional neural network have better detection results, and their accuracy rates are higher than those of the other three networks; and the proposed FrCMs-DNNs has good effectiveness.

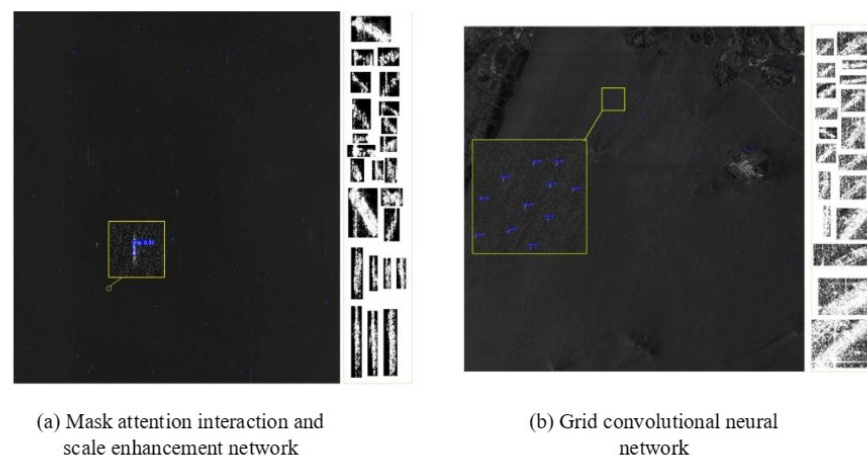


Figure 15. Cont.

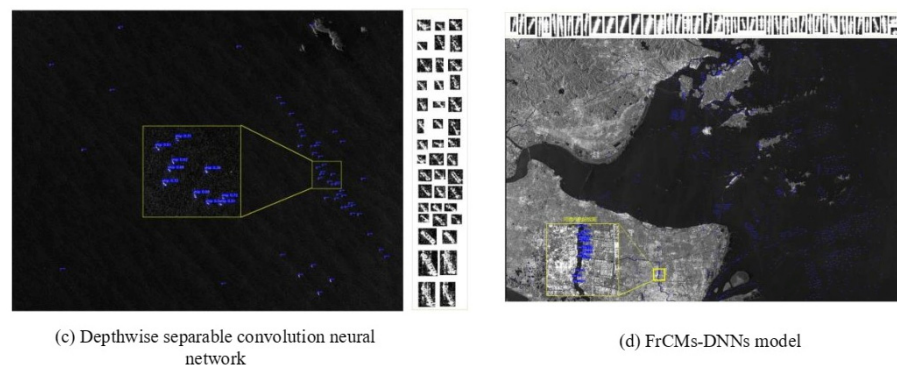


Figure 15. Chart of the high-speed SAR image ship detection results' comparison.

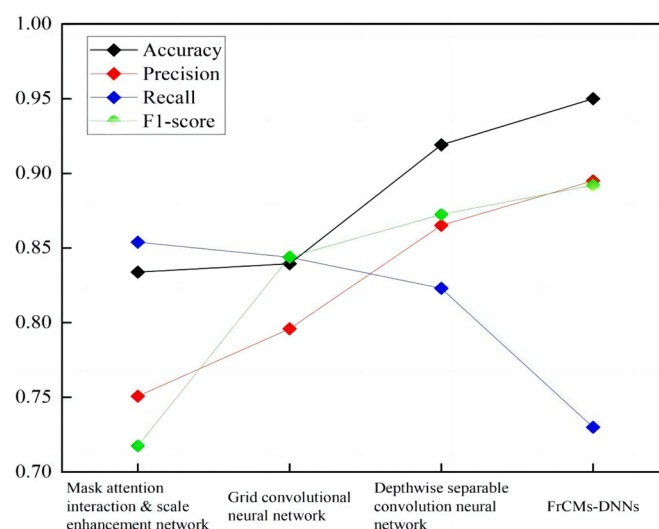


Figure 16. Comparison of accuracy, precision, recall, and F1-score of different network models.

The performance of the proposed FrCMs-DNNs is evaluated in image reconstruction and image classification experiments.

3.6. 3D Recognition Consumes Time

The efficiency of 3D recognition is also a matter of concern in the application. Being efficient and fast is also of great importance. In this section, different types of fractional moment DNN models are used to verify that the proposed FrCMs-DNNs model takes less time and recognizes well in the 3D recognition process.

The time consumption method for 3D recognition is influenced by many factors, such as the feature extraction method, classifier, and hardware conditions. In this section, the hardware condition method is used to calculate the time consumption. However, the time consumption is related to the configuration of the device, such as the GPU, CPU model, and memory size, which will affect the efficiency of 3D image recognition.

When running 3D object recognition experiments on a computer with performance parameters of AMD A10-7300 Radeon 610 Compute Cores 4C+6G, the elapsed time is recorded through the timeliness of the sensor as the order increases. The beginning and end times should be recorded for data processing, and the experiment should be repeated several times to avoid large errors.

In this experiment, 3D object recognition is performed using fractional-order moments-DNNs models of different orders with the fractional parameter $\alpha = 1.4$ under the PSB dataset of size $128 \times 128 \times 128$ and the medical images dataset, respectively, and the time used for the fractional-order moments-DNNs model consumption corresponding to their recognition process is collected, as shown in Tables 8 and 9.

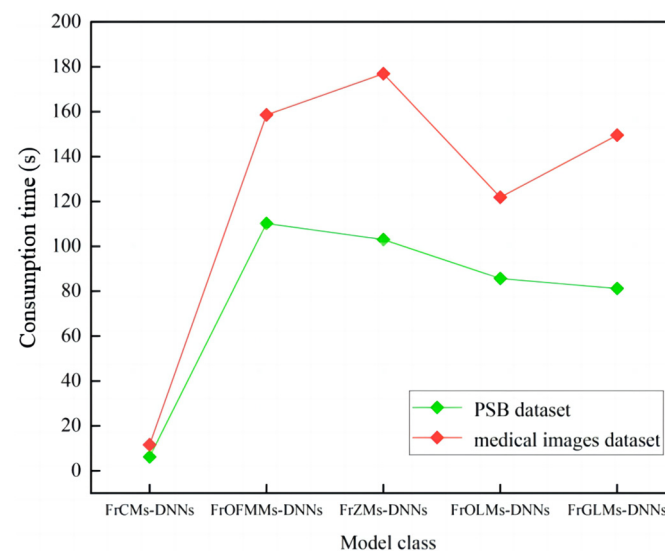
Table 8. The time consumption of 3D object recognition for PSB dataset (unit: second).

Order(n,m,k)	FrCMs-DNNs	FrOFMMs-DNNs	FrZMs-DNNs	FrOLMs-DNNs	FrGLMs-DNNs
(0,0,0)	0.020	0.284	0.108	0.060	0.085
(2,2,2)	0.150	0.765	0.520	0.655	0.592
(4,4,4)	0.433	8.250	6.017	7.650	5.520
(6,6,6)	0.840	26.270	18.120	19.755	17.529
(8,8,8)	1.950	45.382	35.460	39.155	30.367
(10,10,10)	4.660	78.500	80.230	62.832	60.630
(12,12,12)	8.200	114.725	119.735	94.735	90.670
(14,14,14)	14.025	197.115	222.100	170.235	185.420
(16,16,16)	25.088	520.521	445.150	375.850	340.150
Mean value	6.152	110.201	103.049	85.659	81.218

Table 9. The time consumption of 3D object recognition for medical images dataset (unit: second).

Order(n,m,k)	FrCMs-DNNs	FrOFMMs-DNNs	FrZMs-DNNs	FrOLMs-DNNs	FrGLMs-DNNs
(0,0,0)	0.045	0.520	0.230	0.085	0.189
(2,2,2)	0.230	1.395	1.120	0.920	1.279
(4,4,4)	0.855	15.170	10.785	10.980	10.952
(6,6,6)	1.635	48.172	38.520	27.850	38.953
(8,8,8)	3.755	85.575	73.242	55.520	65.785
(10,10,10)	8.585	144.048	172.661	89.012	131.387
(12,12,12)	15.605	210.450	252.406	134.205	196.455
(14,14,14)	25.000	361.565	493.210	247.985	400.520
(16,16,16)	48.305	560.150	550.175	530.220	500.658
Mean value	11.557	158.561	176.928	121.864	149.575

As shown in Tables 8 and 9 and Figure 17, it is evident that the time consumed by the fractional-order moments–DNNs model for recognizing 3D objects increases linearly with the gradual increase in moment order; the average time consumed by the proposed FrCMs–DNNs model is the shortest compared to other fractional-order moments–DNNs models, which can indicate that the FrCMs–DNNs model is efficient in 3D object recognition.

**Figure 17.** Comparison of consumption time of different classes of fractional-order moments–DNNs models under PSB dataset and medical images dataset.

As a nonlinear feature extraction method, FrCMs can capture higher-order statistical information in data, but it is not easy to extract results from 3D images. However, deep

neural networks have nonlinear activation functions, which can improve the extraction of complex structures and better adapt to the nonlinear data relationship. Utilizing fractional moments, FrCMs can reduce feature dimensions while preserving crucial information in the data. This aids in reducing the parameter count of the neural network model, lessening the computational burden of training and inference, and enhancing overall computational efficiency.

Therefore, the combination of fractional Chebyshev moments and deep neural networks can fully leverage the advantages of both, improve the extraction ability, adaptability, and robustness of FrCMs-DNNs, and reduce the dimensional requirements, making it more suitable for dealing with a variety of practical problems.

4. Limitations and Future Work

Fusing fractional Chebyshev moments and deep neural networks is a new image recognition technology, and its main characteristics include the following:

- (1) High efficiency: the method can recognize 3D images quickly and accurately, and the processing speed is fast;
- (2) High accuracy: the method combines the benefits of fractional Chebyshev moments and deep neural networks, effectively enhancing the accuracy of 3D image recognition;
- (3) High reliability: the method adopts the integration of multiple technologies to enhance the reliability of 3D image recognition and reduce misjudgment rates.

This method also has its limitations. It needs a lot of training data to enhance the recognition performance, and the quality of the training data is also high. High computing power and storage space are required to support model training and inference. It is necessary to set and optimize multiple sets of parameters, which requires a high technical level.

The development of image recognition algorithms will continue to explore the improvement of deep learning, multimodal recognition, and migration learning with fewer samples to enhance the accuracy, generalization, and adaptability of image recognition. Also, attempts can be made to combine fractional-order moments with other types of classifiers to build new image recognition algorithms.

5. Conclusions

The FrCMs-DNNs method for 3D image recognition is proposed by combining fractional Chebyshev moments and deep neural networks. The experimental results of 3D image reconstruction show that FrCMs have the smallest MSE and the highest accuracy of image reconstruction in comparison with other fractional moments. For the PSB dataset, the recognition accuracy of FrCMs is 32.1% higher than the mean of other fractional-order moment invariants. For the medical images dataset, the recognition accuracy of FrCMs is 27.3% higher than the mean of other fractional-order moment invariants. The recognition rate of the FrCMs-DNNs model surpasses that of other fractional-order moment-DNN models in the PSB and medical images datasets under the same parameters, and the average value of time consumed by the FrCMs-DNNs model for 3D object recognition in the PSB and medical images datasets is the smallest for different types of fractional-order moments-DNNs models under the condition of increasing order.

Author Contributions: Methodology, L.G. and X.Z.; software M.Z.; investigation, J.Z. and X.Z.; writing, X.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This study is supported by the Liaoning Provincial Department of Education Youth Project (No.1030040000560), National Natural Science Foundation of China (No.42071428), Liaoning Province Applied Basic Research Program (Youth Special Project) (2023JH2/101600038), Shenyang Youth Science and Technology Innovation Talent Support Program (RC220458) and Jinyi Zhang is funded by the China Scholarship Council (No. 202208210120).

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Acknowledgments: The authors thank the reviewers for their valuable comments on the paper.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

Table A1. Glossary of mathematical formulas.

Formula	Annotation
$\alpha_x, \alpha_y, \alpha_z$	Fractional parameter values for x, y and z axes.
$I_{X_p}^{\alpha_x}(x_i), I_{Y_q}^{\alpha_y}(y_j), I_{Z_r}^{\alpha_z}(z_k)$	Spatial moment invariants of the x, y and z axes.
$IT_{X_p}^{\alpha_x}(x_i), IT_{Y_q}^{\alpha_y}(y_j), IT_{Z_r}^{\alpha_z}(z_k)$	Spatial translation invariants of x, y and z axes.
FrFMMIs	Fractional-order Fourier-Mellin moment Invariants
FrLMIs	Fractional-order Legendre moment Invariants
FrGLMIs	Fractional-order Generalized Laguerre moment Invariants
FrZMIs	Fractional-order Zernike moment Invariants
GMIIs	Gegenbauer Moment Invariants

References

1. Song, L.B.; Ren, Z.J.; Fan, C.J.; Qian, Y.X. Virtual source for the fractional-order Bessel–Gauss beams. *Opt. Commun.* **2021**, *499*, 127307. [\[CrossRef\]](#)
2. Karmouni, H.; Jahid, T.; Sayyouri, M.; Alami, R.E.; Qjidaa, H. Fast 3D image reconstruction by cuboids and 3D Charlier’s moments. *J. Real-Time Image Process.* **2020**, *17*, 949–965. [\[CrossRef\]](#)
3. Babadian, R.P.; Faez, K.; Amiri, M.; Falotico, E. Fusion of tactile and visual information in deep learning models for object recognition. *Inf. Fusion* **2023**, *92*, 313–325. [\[CrossRef\]](#)
4. Xiao, M.; Yang, B.; Wang, S.L.; Zhang, Z.P.; Tang, X.L.; Kang, L. A feature fusion enhanced multiscale CNN with attention mechanism for spot-welding surface appearance recognition. *Comput. Ind.* **2022**, *135*, 103583. [\[CrossRef\]](#)
5. Wei, W.; Dai, H.; Liang, W.T. Regularized least squares locality preserving projections with applications to image recognition. *Neural Netw.* **2020**, *128*, 322–330. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Wang, Y.B.; You, Z.H.; Yang, S.; Yi, H.C.; Chen, Z.H.; Zheng, K. A deep learning-based method for drug-target interaction prediction based on long short-term memory neural network. *BMC Med. Inform. Decis. Mak.* **2020**, *20* (Suppl. S2), 49. [\[CrossRef\]](#)
7. Kaur, A.; Singh, C. Automatic cephalometric landmark detection using Zernike moments and template matching. *Signal Image Video Process.* **2015**, *9*, 117–132. [\[CrossRef\]](#)
8. Farokhi, S.; Sheikh, U.U.; Flusser, J.; Yang, B. Near infrared face recognition using Zernike moments and Hermite kernels. *Inf. Sci.* **2015**, *316*, 234–245. [\[CrossRef\]](#)
9. Ghazal, M.T.; Abdullah, K. Face recognition based on curvelets, invariant moments features and SVM. *TELKOMNIKA Indones. J. Electr. Eng.* **2020**, *18*, 733–739. [\[CrossRef\]](#)
10. Emam, M.; Han, Q.; Niu, X.m. PCET based copy-move forgery detection in images under geometric transforms. *Multimed. Tools Appl.* **2016**, *75*, 11513–11527. [\[CrossRef\]](#)
11. Wang, P.; Wang, L.G.; Leung, H.; Zhang, G. Super-Resolution Mapping Based on Spatial–Spectral Correlation for Spectral Imagery. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 2256–2268. [\[CrossRef\]](#)
12. Shang, X.; Song, M.; Wang, Y.; Yu, C.; Yu, H.; Li, F.; Chang, C.I. Target-Constrained Interference-Minimized Band Selection for Hyperspectral Target Detection. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 6044–6064. [\[CrossRef\]](#)
13. Pallotta, L.; Cauli, M.; Clemente, C. Classification of micro-Doppler radar hand-gesture signatures by means of Chebyshev moments. In Proceedings of the 2021 IEEE 8th International Workshop on Metrology for AeroSpace (MetroAeroSpace), Naples, Italy, 23–25 June 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 182–187.
14. Machhour, S.; Grivel, E.; Legrand, P.; Corretja, V.; Magnant, C. A Comparative Study of Orthogonal Moments for Micro-Doppler Classification. In Proceedings of the 2018 26th European Signal Processing Conference (EUSIPCO), Rome, Italy, 3–7 September 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 366–370.
15. Bolourchi, P.; Demirel, H.; Uysal, S. Target recognition in SAR images using radial Chebyshev moments. *Signal Image Video Process.* **2017**, *11*, 1033–1040. [\[CrossRef\]](#)
16. Neri, M.; Pallotta, L.; Carli, M. Low-Complexity Environmental Sound Classification using Cadence Frequency Diagram and Chebychev Moments. In Proceedings of the 2023 International Symposium on Image and Signal Processing and Analysis (ISPA), Rome, Italy, 18–19 September 2023; IEEE: Piscataway, NJ, USA, 2023; pp. 1–6.
17. Li, J.B.; Pan, J.S.; Lu, Z.M. Face recognition using Gabor-based complete Kernel Fisher Discriminant analysis with fractional power polynomial models. *Neural Comput. Appl.* **2009**, *18*, 613–621. [\[CrossRef\]](#)

18. Li, D.; Mathews, C.; Zamarripa, C.; Zhang, F.; Xiao, Q. Wound tissue segmentation by computerised image analysis of clinical pressure injury photographs: A pilot study. *J. Wound Care* **2022**, *31*, 710–719. [[CrossRef](#)] [[PubMed](#)]
19. Xiao, B.; Wang, G.Y.; Li, W.S. Radial shifted Legendre moments for image analysis and invariant image recognition. *Image Vis. Comput.* **2014**, *32*, 994–1006. [[CrossRef](#)]
20. Deepthi, V.H.; Swarna, K.; Kumar, C.M.S.; Kant, D.S.; Rao, A.K.; Kyamaky, K. A Novel Zernike Moment-Based Real-Time Head Pose and Gaze Estimation Framework for Accuracy-Sensitive Applications. *Sensors* **2022**, *22*, 8449. [[CrossRef](#)] [[PubMed](#)]
21. Shao, Z.H.; Shang, Y.Y.; Zhang, Y.; Liu, X.L.; Guo, G.D. Robust watermarking using orthogonal Fourier–Mellin moments and chaotic map for double images. *Signal Process.* **2016**, *120*, 522–531. [[CrossRef](#)]
22. Yang, H.Y.; Qi, S.R.; Wang, C.; Yang, S.B.; Wang, X.Y. Image analysis by log-polar Exponent-Fourier moments. *Pattern Recognit.* **2020**, *101*, 107177. [[CrossRef](#)]
23. Zhang, H.; Li, Z.; Liu, Z. Fractional orthogonal Fourier-Mellin moments for pattern recognition. In Proceedings of the Chinese Conference on Pattern Recognition, Chengdu, China, 5–7 November 2016; pp. 766–778.
24. El Ogri, O.; Daoui, A.; Yamni, M.; Karmouni, H.; Sayyouri, M.; Qjidaa, H. New set of fractional-order generalized Laguerre moment invariants for pattern recognition. *Multimed. Tools Appl.* **2020**, *79*, 23261–23294. [[CrossRef](#)]
25. Kaur, P.; Pannu, H.S.; Malhi, A.K. Plant disease recognition using fractional-order Zernike moments and SVM classifier. *Neural Comput. Appl.* **2019**, *31*, 8749–8768. [[CrossRef](#)]
26. Hosny, K.M.; Darwish, M.M.; Eltoukhy, M.M. New fractional-order shifted Gegenbauer moments for image analysis and recognition. *J. Adv. Res.* **2020**, *25*, 57–66. [[CrossRef](#)] [[PubMed](#)]
27. Vargas, V.H.; Camacho, B.C.; Rivera, L.J.S.; Noriega, E.A. Some aspects of fractional-order circular moments for image analysis. *Pattern Recognit. Lett.* **2021**, *149*, 99–108. [[CrossRef](#)]
28. Guo, B.Y.; Zhuang, Z.J.; Pan, J.S.; Chu, S.C. Optimal Design and Simulation for PID Controller Using Fractional-Order Fish Migration Optimization Algorithm. *IEEE ACCESS* **2021**, *9*, 8808–8819. [[CrossRef](#)]
29. Zhang, X.F.; He, H.; Zhang, J.X. Multi-focus image fusion based on fractional order differentiation and closed image matting. *ISA Trans.* **2022**, *129 Pt B*, 703–714. [[CrossRef](#)]
30. Andrushia, A.D.; Patricia, A.T. Artificial bee colony optimization (ABC) for grape leaves disease detection. *Evol. Syst. Interdiscip. J. Adv. Sci. Technol.* **2020**, *11*, 105–117. [[CrossRef](#)]
31. Smith, A.; Jones, B.; Wang, C. 3D object recognition using convolutional neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016.
32. Zhang, Z.; Song, Y.; Qi, H. Shape completion using 3D-encoder-predictor CNNs and shape synthesis. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017.
33. Liu, J.; He, Z.; Tang, J. 3D shape recognition using fractional Chebyshev moments. *J. Vis. Commun. Image Represent.* **2019**, *65*, 102634.
34. Wang, Y.; Lu, J.; Huang, Q. Fractional Chebyshev moments-based 3D shape analysis. *Signal Process.* **2021**, *181*, 107893.
35. Chen, X.; Zhang, Y.; Li, S. 3D shape analysis using fractional Chebyshev moments and convolutional neural networks. *Pattern Recognit.* **2018**, *79*, 150–162.
36. Chen, X.; Zhang, Y.; Li, S. Joint optimization of deep neural networks and fractional Chebyshev moments for 3D shape recognition. *Neural Netw.* **2022**, *145*, 148–160.
37. Liu, J.; Zhao, Y.; Wu, Q. Joint application of 3D convolutional neural networks and fractional Chebyshev moments for scene understanding. *Comput. Vis. Image Underst.* **2023**, *214*, 103121.
38. Hosny, K.M.; Darwish, M.M.; Aboelenen, T. New fractional-order Legendre-Fourier moments for pattern recognition applications. *Pattern Recognit.* **2020**, *103*, 107324. [[CrossRef](#)]
39. McGill 3D Shape Benchmark. Available online: <http://www.cim.mcgill.ca/~shape/benchMark/> (accessed on 9 August 2020).
40. Centre Hospitalier Universitaire Hassan II. Available online: <http://www.chu-fes.ma/ar/home-ar-2/> (accessed on 10 October 2020).
41. Zhang, M.M.; Choi, J.; Daniilidis, K.; Wolf, M.T.; Kanan, C. VAIS: A dataset for recognizing maritime imagery in the visible and infrared spectrums. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Boston, MA, USA, 7–12 June 2015; IEEE Computer Society: Washington, DC, USA, 2015.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.