SIN-CHN SCIENTIFIC PRESS PTE.LTD.

Molecular & Cellular Biomechanics 2025, 22(2), 1006.
https://doi.org/10.62617/mcb1006

Article

Applications and challenges of artificial intelligence-driven 3D vision in biomedical engineering: A biomechanics perspective

Lei Wang*, **Zunjie Zhu**

Hangzhou Dianzi University, Hangzhou 310005, China
*** Corresponding author:** Lei Wang, 221060021@hdu.edu.cn

CITATION

Wang L, Zhu Z. Applications and challenges of artificial intelligence-driven 3D vision in biomedical engineering: A biomechanics perspective. Molecular & Cellular Biomechanics. 2025; 22(2): 1006. https://doi.org/10.62617/mcb1006

ARTICLE INFO

Received: 4 December 2024
Accepted: 7 January 2025
Available online: 20 January 2025

Abstract: This paper explores the applications and challenges of artificial intelligence (AI)-driven 3D vision technology in biomedical engineering, with a specific focus on its integration with biomechanics. 3D vision technology offers richer spatial information compared to traditional 2D imaging and is increasingly applied in fields like medical image analysis, surgical navigation, lesion detection, and biomechanics. In biomechanics, AI-driven 3D vision is used for analyzing human movement, modeling musculoskeletal systems, and assessing joint biomechanics. However, challenges persist, including image quality, computational resource demands, data privacy, and algorithmic bias. This paper reviews the development of 3D vision technology and AI, discusses its applications in biomedicine and biomechanics, and addresses the key technical obstacles, offering insights into the future development of these technologies in the context of biomedical and biomechanical research.

Keywords: artificial intelligence; 3D vision; biomedical engineering; biomechanics; medical imaging; deep learning; surgical navigation; musculoskeletal modeling; joint biomechanics; image processing

1. Introduction

With the rapid development of Artificial Intelligence (AI) technology, it is making breakthroughs in several fields, especially in biomedical engineering. Biomedical engineering, as an interdisciplinary field, combines advanced technologies from a number of disciplines, including medicine, engineering, computer science, and biology, with the aim of solving healthcare problems through innovative technological solutions [1–3]. And artificial intelligence, especially its application in computer vision and deep learning, is revolutionizing many aspects of medical image analysis, diagnostic decision-making, surgical navigation and disease prediction [4].

In conventional medical image analysis, physicians usually rely on 2D images for lesion detection, diagnosis and treatment planning [5–7]. However, such 2D images do not provide sufficient spatial information in some cases, especially in the case of complex organ structures or 3D lesions. With the advent of 3D vision technology, medical image processing is able to provide more comprehensive and accurate spatial information. Three-dimensional images can not only better simulate and present the structure of the human body, but also help doctors carry out accurate disease diagnosis and treatment plan design, especially in the diagnosis of tumors, neurological diseases, cardiovascular diseases and other complex lesions, which show great advantages [8,9].

3D vision technology [10] is an important research topic in the field of computer vision, which mainly refers to how to recover the depth information lost during the projection of 3D vision into 2D images, i.e., reconstructing the 3D coordinates of pixel points used for depth estimation. This technique makes up for the limitations of 2D images by providing rich spatial information, especially when dealing with complex biological structures in the human body, which can more accurately present the shape, size, location and other key factors of the disease. For example, in the diagnosis of tumors, 3D images can help doctors better assess the growth pattern of the tumor, the degree of invasion, and the relationship with the surrounding tissues, so as to better formulate a treatment plan [11–13]. In surgery, 3D vision technology can provide surgeons with real-time spatial data to improve the precision and safety of surgery, especially in minimally invasive surgery. The application of 3D vision is not only limited to image analysis and diagnosis, it is also widely used in surgical navigation, robot-assisted surgery, personalized medicine and many other fields. Through real-time 3D reconstruction and 3D visualization technology, doctors are able to obtain precise spatial positioning information in complex surgical environments, reduce risks during surgery and improve the success rate of surgery.

In this paper we mainly summarize and study the applications and challenges of 3D vision technology in biomedical engineering, in Chapter 2 we review the development of 3D vision technology and the combination of AI in this field, in Chapter 3 we discuss in detail its main application areas in biomedicine, and finally we analyze the main technical challenges faced by AI-driven 3D vision technology in the field of biomedical engineering and future development directions.

2. The basic principle and development of 3D vision technology

3D stereo vision technology [14] is an important research topic in the field of computer vision: how to recover the depth information lost in 3D vision during projection into 2D images, i.e., to reconstruct the 3D coordinates of the pixel points used for depth estimation. In order to realize a 3D image sensor, it can be obtained either directly by active radar or with the help of passive 2D image computation. Among them, the main representatives of active depth sensors are time-of-flight (ToF) [15] radars and structured light cameras. Unlike passive stereo vision cameras, such devices are sensors that emit signals and receive them actively. ToF-based radars are widely used for scene reconstruction and geospatial information analysis, but are difficult to apply to embedded platforms due to their weight, power consumption, and field-of-view limitations. Structured light based sensors have good accuracy for indoor close range scenes. Passive-based stereo vision technology [16] does not require special lighting projection devices, and only utilizes the camera to capture the image of the object to be measured and establish the relative position relationship between the object and the camera, so as to obtain the three-dimensional information of the surface of the object to be measured. The hardware required for passive vision measurement is relatively simple. Depending on the number of cameras used, passive vision measurement, also known as stereo vision, can be categorized into monocular vision measurement, binocular vision measurement, and multicameral vision measurement. Three-dimensional vision technology is a technology that simulates and reconstructs objects, scenes, or the human body by acquiring, processing, and analyzing three-dimensional spatial data. Compared with traditional 2D vision, 3D vision can provide rich spatial information to help us understand and analyze complex objects and environments more accurately. With the advancement of sensor technology and the continuous development of computer vision algorithms, 3D vision technology has gradually become an important tool in modern scientific research, medical image analysis, robotics, virtual reality and other fields.

2.1. Introduction to the realization and principles of 3D vision

2.1.1. Stereoscopic vision

Stereo vision technology [14] imitates human binocular vision, and acquires multi-view images by simultaneously shooting the same scene or object from different angles with two or more cameras. Multiple cameras are extensions of two cameras, so binocular stereo vision is the basis of various stereo vision systems. By comparing the pixel differences (parallax) in these images, the depth information of the scene or object can be deduced to generate a 3D model. In this section we focus on the principle of stereo vision implementation modeled on binocular vision measurements.

The stereo vision system [17] is calibrated to obtain an ideal binocular imaging model (the schematic diagram is shown in **Figure 1**), i.e., the left and right cameras are located in the same plane (the optical axis is parallel), and the camera parameters (focal length, image origin, and coefficient of aberration) are the same, a three-dimensional spatial point $P(X_c, Y_c, Z_c)$ will be imaged in two image planes, and the pixel longitudinal coordinates of these two imaging points are equal, so that the effect of the longitudinal coordinate can be ignored, and simplified As shown in **Figure 2**, the projected positions of the above two imaging points in the pixel coordinate system are determined by stereo matching, i.e., the left and right matching point pairs (x^l, y) and (x^r, y), and the parallax $D = x^l - x^r$ is obtained. Based on the pixel focal length of the imaging model in the camera calibration information and the optical centers of the two imaging systems, the focal lengths of the pixels are determined. f and the optical centers of the two imaging systems O_L and O_R the baseline distance between the optical centers of the two imaging systems B, the value of the three-dimensional point $P(X_c, Y_c, Z_c)$ can be calculated by the geometric principle of similar triangles. Specifically, the relationship between parallax D and depth Z_c is as follows.

$$Z_c = \frac{f \times B}{x^l - x^r} = \frac{f \times B}{D} \tag{1}$$

In the formula, the lateral focal length f_x is in pixels.

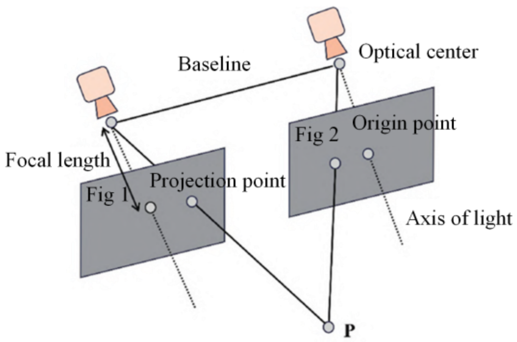

Figure 1. Binocular stereo vision model.

After the depth Z_c is known, X_c and Y_c in the 3D coordinates of the point P are calculated by the camera's internal reference, as shown in Equations (2) and (3). In the equation, the c_x, c_y denotes the left camera optical center O_L is the coordinate of the left camera optical center in the left pixel coordinate system.

$$X_c = (x^l - c_x) \times \frac{Z_c}{f} \tag{2}$$

$$Y_c = (y - c_y) \times \frac{Z_c}{f} \tag{3}$$

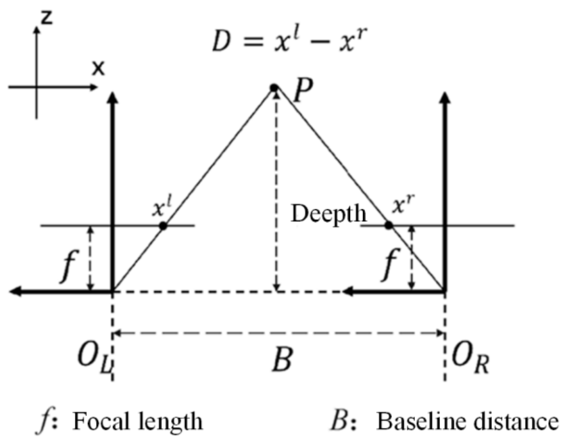

Figure 2. Simplified diagram of binocular stereo vision model.

The basic principle of stereo vision is similar to that of the human eye, but in machine vision, images are usually processed and analyzed by computer algorithms to obtain depth data. The advantage is that the equipment is simple and less costly, but the accuracy and stability may be affected when the light changes or the viewing angle is too small.

2.1.2. Laser scanning and structured light

Laser scanning technology [18] utilizes a laser beam to scan the target object through three-dimensional space and obtains distance information by calculating the reflection time of the laser beam, which is a method to realize three-dimensional vision technology based on the active depth sensing method, and its imaging principle is the same as that of structured light imaging. The measurement principle of laser scanning technology is that the laser is emitted to the surface of the object to be measured in the form of a point or line laser at a certain angle, and then through a certain angle of reflection or scattering, it is imaged on the photosensitive device, and according to the proportionality between the displacement of the object point on the surface of the object to be measured and the offset produced by the corresponding image point of the object point, the actual displacement of the object point is calculated, and then the three-dimensional information of the object under test is obtained as a whole. information of the object to be measured. The realization of the schematic diagram is shown in **Figure 3**.

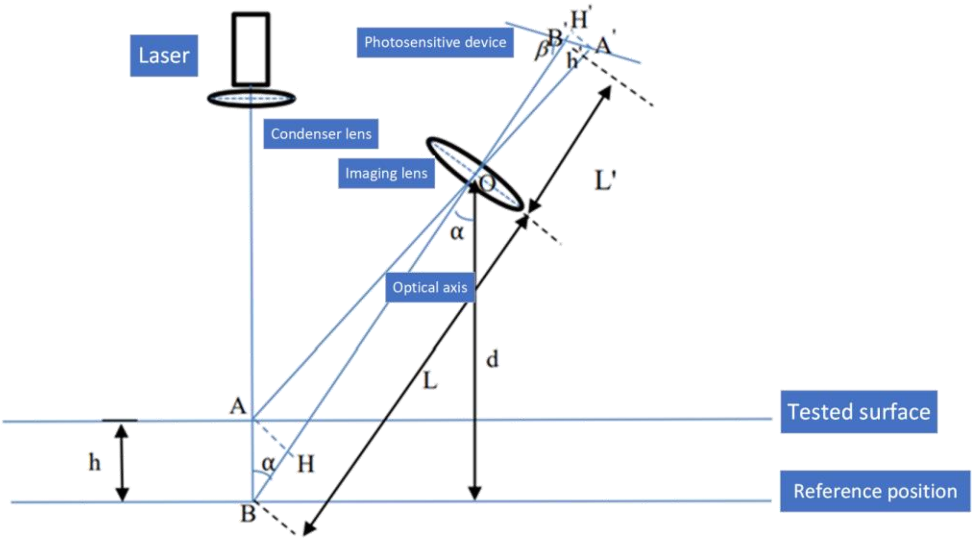

Figure 3. Schematic diagram of laser scanning technology.

Assuming that B is a measurement point on the reference plane, is the laser emitted by the laser incident line formed by a reference plane of light incident point, and then the light incident point on the surface of the object occurs at an angle of α diffuse reflection in the photosensitive device imaging for the B′ point, point A is a point on the surface of the object to be measured, and its imaging point on the photosensitive surface is A′, AB is the incident beam, AB is the reflected beam, AB is the incident beam and AB is the reflected beam. BB′ AB is the incident beam, AB is the reflected beam, AB is the incident beam and AB is the reflected beam. AB and BB′ The angle between the α, β for BB′ and A′B′ angle, B point, B′ and the perpendicular distance between the imaging surface for the object distance L, image distance L'.

From **Figure 3**, we can see that h is the relative height between point A and point B, and h' is the displacement of h in the image plane, i.e., the distance between point A and point B. Therefore, according to the triangular relationship between the object and the image points, we can get the distance from the point of the surface of the object being measured to the reference plane, and then find out the three-dimensional information of the surface of the object being measured, and obtain the measurement data we need. and then find out the three-dimensional information of the surface of the object to be measured, and obtain the measurement data we need.

In **Figure 3**, AH is perpendicular to BB′ and A′H′ is perpendicular to BB′, then it can be seen that ΔOHA is similar to $\Delta OHA'$, and according to the triangle similarity relation, then we have: $\frac{AH}{A'H'} = \frac{OH}{OH'}$, which means that

$$\frac{h\sin\alpha}{h'\sin\beta} = \frac{L - h\cos\alpha}{L' + h'\cos\beta}$$

Then the height h can be expressed as

$$h = \frac{Lh'\sin\beta}{L'\sin\alpha + h'\sin(\alpha+\beta)}$$

The advantage of this technology is that it provides highly accurate depth data and does not rely on an external light source, so it can work efficiently in low-light environments. Laser scanners typically acquire 3D data of the entire object or scene by adjusting the scanning frequency and scanning area to generate point cloud data.

Structured light technology analyzes the three-dimensional shape of an object's surface by projecting a known light pattern (e.g., a stripe or checkerboard grid) onto the surface and using a camera to capture the deformed light pattern. The principle of its realization is the same as that of laser scanning technology and will not be repeated. Structured light technology can accurately obtain the depth information of the object surface, which is widely used in facial recognition, human body scanning and industrial inspection.

2.1.3. Time-of-Flight (ToF) technology

Time of Flight (ToF) technology [15] is a distance measurement method based on light pulses. ToF sensors calculate distance by emitting a light pulse and measuring the time difference between the emission and return of the light pulse. This technique provides fast, real-time 3D data and is especially outstanding in dynamic environments. The TOF method [19] can be generally categorized into two types according to the modulation method: pulsed modulation (Pulsed Modulation) and continuous wave modulation (Continuous Wave Modulation). The schematic diagram is shown in **Figure 4**. The advantage of the ToF sensor is that it can quickly capture the 3D information of the whole scene, which is suitable for real-time applications, such as smartphones, autonomous driving and other fields.

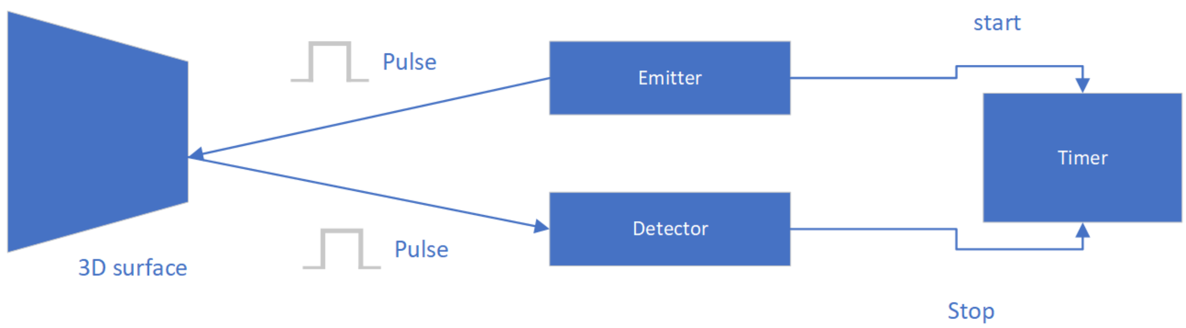

Figure 4. Schematic framework of time-of-flight (ToF) technology.

2.2. Deep learning and 3D vision

In recent years, deep learning techniques, especially convolutional neural networks (CNNs), have played an important role in the application of 3D vision. Deep learning significantly improves the accuracy and efficiency of 3D vision systems by automatically learning complex features in the data. The application of deep learning technology has made the processing of 3D vision more intelligent and automated, especially in the fields of medical imaging, robot perception, and automatic driving, which have made great progress.

2.2.1. Convolutional neural networks (CNN) in image processing

Convolutional Neural Network (CNN) [20] is a deep learning model that can automatically learn features from images. Its core idea is to automatically extract local features (e.g., edges, texture, shape, etc.) in an image through a convolutional layer, and perform feature compression through a pooling layer, and finally perform classification or regression through a fully connected layer. Its framework diagram is shown in **Figure 5**. In 3D vision, CNN is widely used in the following aspects:

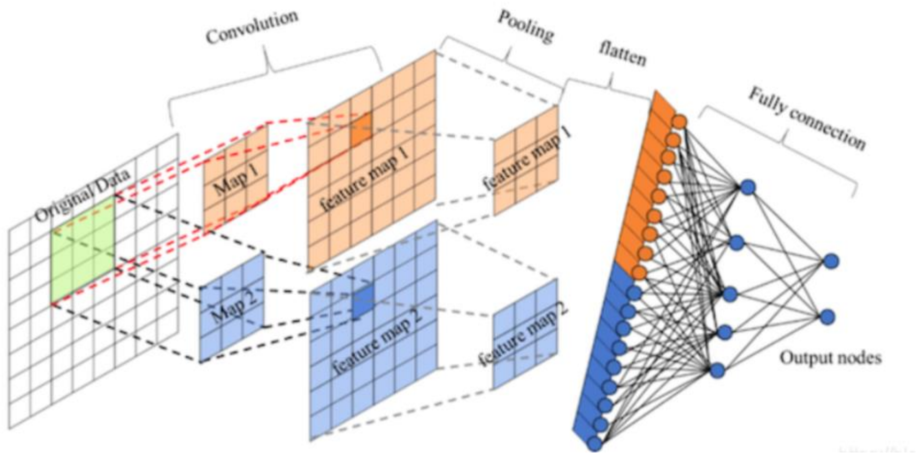

Figure 5. Convolutional neural network framework diagram.

- Image segmentation: automated segmentation of medical images or other 3D data using CNNs to extract specific structures (e.g., tumors, blood vessels, organs, etc.). For example, CNN-based 3D segmentation can perform automated segmentation of MRI or CT images to help doctors diagnose diseases faster.
- Feature extraction and matching: CNNs are able to automatically extract feature points in multi-view images or depth images and match them for tasks such as 3D reconstruction and target recognition.
- Target detection and localization: CNNs can detect targets in 3D scenes and accurately localize their spatial locations, which are widely used in automatic driving, robot perception, and other fields.

2.2.2. How deep learning improves 3D vision accuracy and efficiency

Deep learning, especially convolutional neural networks, can significantly improve the accuracy and efficiency of image analysis in 3D vision. Trained with a large amount of labeled data, deep learning models are able to automatically extract important features from images, avoiding the complexity of traditional manual feature extraction and reducing the possible bias of human intervention. The advantages of deep learning include:

- Efficient automated processing: Traditional 3D vision processing requires a lot of human intervention, such as feature design and labeling. Deep learning, on the other hand, can automate all steps from image input to result output through end-to-end training, greatly improving processing speed.
- Stronger generalization ability: deep learning models are trained with large-scale data and can effectively learn from different types of 3D images, improving the generalization ability of the model and thus increasing accuracy.
- Real-time processing and efficient computation: with the development of deep learning hardware (e.g., GPUs), 3D vision can be processed in real-time to meet

the high demand for real-time performance such as autonomous driving and surgical navigation.

2.2.3. Automated feature extraction and data labeling

Traditional 3D vision techniques rely on manual feature extraction and data labeling, which usually require expert knowledge and a lot of manual operations with limited quality and efficiency of data labeling. In contrast, the introduction of deep learning, especially self-supervised learning and unsupervised learning, makes the feature extraction and data labeling processes automated. For example, methods based on self-supervised learning are able to train models with unlabeled data, thus reducing the reliance on labeled data. In addition, deep learning is able to automate tasks such as image segmentation, target recognition, and anomaly detection, further improving the efficiency of 3D vision systems.

2.3. 3D reconstruction and image fusion techniques

3D reconstruction and image fusion techniques are two very important research directions in 3D vision, and they are often used in combination to generate high-precision 3D models and perform in-depth analysis.

2.3.1. Three-dimensional reconstruction

3D reconstruction techniques are used to acquire 2D or 3D data of an object or scene by various means (e.g., laser scanning, stereo vision, ToF techniques, etc.), and then convert these data into a 3D model. The key challenge of 3D reconstruction technology is how to perform accurate depth measurement and spatial modeling in complex environments. The application of deep learning in 3D reconstruction, especially in the processing of multi-view images and point cloud data, improves the accuracy and speed of reconstruction. For example, CNN-based depth map reconstruction methods can accurately recover 3D scenes by analyzing multiple 2D images.

2.3.2. Image fusion techniques

Image fusion technology refers to the synthesis of image information from different sensors or different viewpoints to obtain more complete and accurate 3D data. For example, fusion of image data acquired by different sensors (e.g., laser scanning, stereo vision, ToF camera, etc.) can overcome the limitations of a single sensor and improve the accuracy and detail of 3D modeling. Image fusion technology is widely used in medical imaging (e.g., fusion of CT and MRI data), remote sensing, industrial inspection, and other fields. Through multimodal image fusion, more comprehensive 3D features can be extracted from the rich information obtained from different sources, further improving the accuracy and usefulness of the model.

3. Artificial intelligence-driven 3D vision in biomedical engineering

With the rapid development of artificial intelligence (AI) technology, especially deep learning and computer vision, the application of 3D vision technology in the biomedical field has gradually become an important tool for research and clinical treatment. AI-driven 3D vision not only shows great potential in medical image analysis, but also provides accurate and efficient solutions in surgical navigation, pathology analysis, and personalized treatment. These technologies not only improve the accuracy of medical diagnosis, but also help doctors achieve more precise surgical operations and personalized treatment plans.

3.1. Medical image analysis and diagnosis

Medical image analysis is one of the most widespread and promising areas for the application of AI-driven 3D vision technologies. Traditional medical imaging mainly relies on 2D images. Although 2D images can provide certain information about lesions, they often have limitations in the diagnosis of complex diseases, such as the localization, size and shape of tumors. By providing more comprehensive spatial information, 3D imaging technology can more accurately reflect the anatomical structure and lesion areas of the human body. The application of AI in this, especially with the support of technologies such as deep learning algorithms and convolutional neural networks (CNNs), significantly improves the accuracy and efficiency of image analysis. This is especially true for the analysis and 3D reconstruction of CT and MRI images. Traditional medical imaging technology mainly relies on 2D images for diagnosis, but with the combination of 3D vision and AI technology, the depth and breadth of medical image analysis has been greatly expanded.

3.1.1. CT/MRI image analysis and 3D reconstruction

CT and MRI scans provide high-resolution images of the inside of the human body, and 3D reconstruction technology provides doctors with a more comprehensive view of anatomical structures by combining these 2D images into 3D images. AI processes and analyzes 3D images through deep learning algorithms, which are capable of automatically segmenting different tissues and organs, and identifying areas of pathology, such as tumors, vascular lesions, and soft tissue injuries.

- 3D reconstruction and accurate diagnosis: While traditional CT/MRI images usually require detailed analysis and diagnosis by specialized radiologists, AI-driven 3D reconstruction technology can automatically convert 2D image data obtained from scans into 3D data, thus improving doctors' diagnostic efficiency. For example, in brain CT or MRI image analysis, AI algorithms can accurately segment brain regions and detect brain tumors or blood vessel abnormalities, helping doctors make faster diagnoses.
- Tumor detection and quantitative analysis: In tumor diagnosis, CT, MRI, PET, and other imaging technologies are able to generate three-dimensional image data to help doctors identify the location, volume, morphology, and other characteristics of tumors. AI technologies, especially deep learning algorithms, can provide more accurate diagnosis by automatically identifying and segmenting tumor regions through training on a large amount of image data. For example, tumor detection using convolutional neural networks (CNN) can automatically identify the boundaries of a tumor in 3D image data and calculate its volume, providing doctors with a more accurate assessment of the disease.
- Applications in cardiology: In the field of cardiology, 3D imaging technology combined with AI can be used for automatic segmentation of cardiac anatomy and functional assessment. For example, 3D images based on CT and MRI can be used to automatically recognize the morphology and function of various parts

of the heart, such as the left ventricle, right ventricle, and coronary arteries, using deep learning models. During diagnosis and treatment, AI-driven 3D image analysis helps doctors gain a comprehensive understanding of the structure and function of the heart and provide more accurate treatment plans.

In 2019, Stanford University and Google Health jointly conducted a study using convolutional neural networks (CNNs) to analyze 3D image data of breast cancer. The study found that deep learning algorithms were able to automatically identify breast tumors and predict tumor growth patterns and aggressiveness by processing MRI scans. Compared with traditional manual image analysis methods, the AI models showed significantly higher accuracy, especially in early tumor diagnosis and grading, and demonstrated the ability to outperform human experts.

3.1.2. Monitoring and diagnosis of early lesions

The combination of AI and 3D vision has revolutionized the monitoring and diagnosis of early lesions. By automatically recognizing lesion areas in images, AI can assist doctors in detecting early-stage lesions such as cancer, cardiovascular disease, and neurological disease, especially small lesions that are not easily detected in the clinic.

Early cancer detection: AI technology has demonstrated significant advantages in the early detection of lung cancer, breast cancer, prostate cancer and so on. For example, through the 3D analysis of CT scan images of the lungs, AI is able to detect tiny lung nodules, thus improving the early diagnosis rate of lung cancer. In addition, AI can assist doctors in evaluating the malignant degree and development trend of tumors based on the morphology, size and other characteristics of the lesions.

3.2. Surgical navigation and robot-assisted surgery

Surgical navigation is another key application that utilizes 3D vision and artificial intelligence technologies. Traditional surgical navigation relies on 2D images to guide surgeons in surgical operations, but the limitations of 2D images make complex surgical operations more difficult, especially in minimally invasive surgery or neurosurgery, where accurate spatial positioning is critical. 3D vision technology combined with artificial intelligence can not only provide real-time 3D images, but also accurately localize the surgical area to help surgeons perform precise surgeries.

3.2.1. 3D vision applications in robot-assisted surgical systems

Robot-assisted surgical systems are able to provide surgeons with real-time 3D image views and precise positioning information by combining 3D vision technology. ai algorithms are able to analyze and adjust the surgical area in real time, helping surgeons make the most accurate decisions and reduce human error.

Precise positioning and surgical path planning: Robot-assisted surgery systems utilize 3D image data and AI technology to provide surgeons with accurate surgical path planning. For example, in spine surgery, through 3D image data, AI can analyze the patient’s spine anatomy in real time and automatically mark key areas, thus helping surgeons make precise cuts to avoid damaging important nerves or blood vessels.

3.2.2. Real-time surgical navigation and precise localization

The combination of AI and 3D vision technology can provide real-time navigation during surgery and adjust surgical strategies in real time to ensure precision. For example, in brain surgery, 3D imaging and AI technology can accurately analyze brain anatomy and guide surgical tools in real time to ensure that surgeons accurately position and operate in complex surgical areas.

- Improving surgical accuracy and safety: AI-driven 3D vision not only provides detailed information on anatomical structures before surgery, but also monitors the patient's physiological changes in real time during surgery and adjusts the surgical plan based on real-time images. The application of this technology significantly reduces risks during surgery, shortens recovery time, and improves the success rate of surgery.
- Specific applications of surgical navigation: In brain surgery, precise localization and navigation are key to ensuring a safe and successful operation. Using 3D vision technology and AI, surgeons can generate 3D brain images based on a patient's brain MRI or CT scan, and accurately segment and label brain tissue with deep learning algorithms. For example, brain tumor segmentation using AI technology can help surgeons accurately identify the boundaries and location of tumors, thereby optimizing the surgical path and reducing damage to surrounding normal brain tissue. AI and 3D vision technology are also widely used for spine surgery navigation. By acquiring 3D spine image data through CT scanning, AI can automatically extract various parts of the spine, including vertebrae, intervertebral discs, spinal cord, etc., and perform precise segmentation and analysis. During surgery, the navigation system based on 3D data can display the spatial position of the spine in real time, helping surgeons to perform precise guidance and positioning, thus improving the safety and success rate of surgery.

3.3. Pathology and cellular imaging

In the field of pathology and cellular imaging, AI-driven 3D vision technologies provide powerful support, especially in the analysis of tissue sections and 3D reconstruction at the cellular level, which greatly enhances the automation and precision of pathology analysis.

3.3.1. Three-dimensional microscopic imaging techniques

Three-dimensional microimaging technology is able to acquire three-dimensional structural information of cells and tissues, which can then be analyzed for detailed pathological analysis. by analyzing these high-resolution three-dimensional images, AI is able to identify cellular morphology, tissue distribution, and areas of pathology, providing pathologists with an aid in diagnosis.

3D reconstruction and analysis of cells and tissues: Using 3D microimaging, AI is able to perform automated 3D reconstruction of tissue sections and analysis of cells and tissues. This is important for cancer research, immunology research, etc. Through 3D reconstruction, AI can accurately identify the spatial relationship between cells and tissues, helping pathologists to identify cancerous cells, areas of inflammation and other pathological changes.

3.3.2. AI-driven automated pathology analysis

AI-driven automated pathology analysis systems can automatically detect and classify lesion areas based on 3D images. For example, the AI system can identify cancer cells, tumor grading, and assess the spread of cancer in pathology section images of breast cancer. Compared with traditional manual analysis, AI offers significant improvements in accuracy and efficiency, especially in high-throughput pathology analysis, where the application of AI greatly saves time and reduces human error.

3.4. Disease prevention and health surveillance

In addition to the diagnostic and therapeutic areas, AI-powered 3D vision technologies are increasingly being used in disease prevention and health monitoring, especially in the areas of personalized health management and disease prediction.

3.4.1. 3D vision-based health data collection

The combination of AI and 3D vision technology enables large-scale health data collection and analysis to help doctors monitor patients' health in real time. This data includes 3D scan information of various body parts, 3D images of the heart and lungs, and analysis of movement and posture.

Real-time health monitoring and early warning system: Utilizing 3D vision and AI analysis technology, it is able to monitor the health status of the human body in real time. For example, through 3D scanning and deep learning algorithms to continuously monitor the patient's spine, joints and other parts of the body, early detection of possible fractures, joint lesions and other problems, so as to provide early warning and reduce the occurrence of disease.

3.4.2. Personalized health management and disease prediction

AI combined with 3D vision technology can help doctors develop personalized health management plans based on a patient's specific situation. By analyzing a patient's 3D images and physiological data, AI can predict the health risks a patient may face in the future and thus provide targeted prevention programs.

Personalized disease prediction and prevention: Through the collection and analysis of long-term health data, AI can predict an individual's risk of developing certain chronic diseases, such as heart disease, diabetes, and tumors. Combined with 3D vision technology, AI can detect lesions or potential health problems at an early stage, helping doctors develop personalized prevention plans and optimize patient health management.

The application of AI-driven 3D vision technology in biomedical engineering has moved from theoretical research to clinical practice, showing strong potential. From medical image analysis, surgical navigation to pathology analysis, the combination of AI and 3D vision not only improves the accuracy and efficiency of diagnosis, but also greatly promotes the development of precision medicine and personalized treatment. With the continuous progress of technology, AI-driven 3D vision will play an even more important role in the field of healthcare in the future, providing stronger decision support for doctors and ultimately benefiting more patients.

3.5. Analysis of human movement and body state

3D vision technology can be widely used in motion analysis, health monitoring, rehabilitation training and other fields by acquiring motion data and shape information in the three-dimensional space of the human body, combined with artificial intelligence algorithms. The following are the main application scenarios and technical features.

Gait analysis and diagnosis of movement disorders

Gait analysis refers to the observation and analysis of an individual's walking or running patterns to assess their athletic fitness. Traditional gait analysis often relies on video surveillance or ground-based sensors, but these methods often struggle to provide detailed three-dimensional information and dynamic changes. AI-powered 3D vision technology is able to capture stereoscopic movement data of the human body through depth cameras or sensors, providing more accurate and comprehensive gait analysis. The diagnosis of movement disorders is to help doctors diagnose various movement disorders, such as Parkinson's disease and movement disorders after stroke, by capturing and analyzing the subtle changes of the human body during movement.

In 2023, a research group of Wang Yizhou at Peking University proposed the MotionBERT method, which learns the general representation of human movement from large-scale and diverse data, and then completes various human-centered downstream video tasks with a unified paradigm. The authors propose a new way to learn representations of human movement and then apply this learned representation to different tasks. In the pre-training phase, the authors trained a model to recover 3D human movements from noise and partial 2D observations. This model can then be fine-tuned in subsequent tasks to suit a particular task. The method uses a two-flow space-time converter network that captures the relationship between human bones in space and time, allowing the model to show very low 3D pose estimation errors during training, which helps in subsequent human state analysis. As shown in the **Figure 6**, artificial intelligence based on 3D vision can describe human body information digitally, better infer pathological conditions, and provide good treatment means

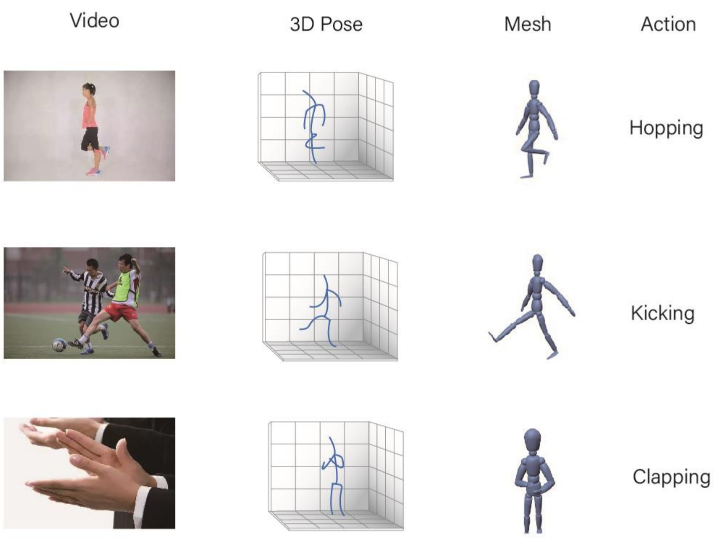

Figure 6. 3D visual digital effect display diagram.

4. Challenges for artificial intelligence-driven 3D vision

Although Artificial Intelligence (AI)-powered 3D vision technology has shown great potential in the biomedical field, it still faces a number of technical, ethical, and resource challenges in its practical application. The following are the main challenges:

4.1. Image quality and data preprocessing issues

The quality of medical image data directly affects the training effect and prediction ability of AI models, especially in 3D visual analysis, the image quality issue is particularly important.

4.1.1. Noise, resolution and contrast issues

Noise, low resolution, or low contrast are often present in medical imaging data, and these problems can lead to misjudgment by AI models or failure to extract critical information from images. Noise is usually caused by factors such as the performance of the scanning equipment, the patient's movement, or disturbances during the image acquisition process, while low resolution and contrast affect the presentation of details, especially in the detection of tiny lesions. Continued research is needed on relevant advanced denoising algorithms (e.g., convolutional neural network (CNN) denoising) and super-resolution techniques that can mitigate the noise and lack of resolution to some extent. In addition, the use of contrast-enhancing techniques (e.g., histogram equalization or adaptive filtering) can help improve the visibility of image details.

4.1.2. Data standardization and consistency issues

The degree of standardization of medical images varies widely across devices and hospitals. Due to the different scanning protocols and image formats among CT, MRI, ultrasound and other devices, the quality and features of images may vary, which makes data preprocessing and analysis more complicated. AI models may exhibit different effects on images from different sources, which can affect their generalization ability. A standardized image preprocessing process can be constructed through large-scale datasets across institutions and devices, adopting a unified image format (e.g., the DICOM standard) and performing image calibration and normalization, which can effectively reduce the inconsistency problems caused by devices and data sources.

4.1.3. Bias in the dataset and inadequacy of the training set

The training of AI algorithms relies on a large number of labeled datasets. However, datasets in medical imaging often suffer from sample imbalance and bias, especially in rare diseases, special populations (e.g., the elderly, children, etc.), and cases from different regions, and the training data is often insufficient, which limits the pervasiveness and accuracy of AI models. The insufficiency of training sets can be compensated by data augmentation techniques, migration learning and unsupervised learning. In addition, increasing labeled data for diverse populations and diseases and cross-regional data collaboration can effectively reduce dataset bias and improve the generalizability of AI algorithms.

4.2. Computing resources and real-time processing requirements

AI-driven 3D vision technology requires a lot of computational resources, especially when performing real-time data processing, which has higher performance requirements.

4.2.1. Hardware requirements for high-performance computing

3D vision techniques and deep learning models often require powerful computational capabilities, especially when processing high-resolution 3D medical images. The training and inference process of Deep Neural Networks (DNN), Convolutional Neural Networks (CNN), and other complex models requires a lot of GPU or TPU support, which puts high demands on hardware resources.

4.2.2. Balancing real-time data processing with response time

In clinical environments, AI-powered 3D vision systems need to analyze patient image data and make decisions in real time, especially in scenarios such as surgical navigation and emergency medicine, where delays can lead to serious consequences. However, the high-dimensional nature of 3D data makes real-time processing a challenge.

4.2.3. Edge computing and cloud computing in healthcare

Edge computing and cloud computing technologies provide new solutions for processing medical images. Edge computing can transfer data processing from remote servers to local devices, reducing the latency of data transmission, while cloud computing is able to process large amounts of medical image data by providing powerful remote computing capabilities. Combining the advantages of edge computing and cloud computing can optimize the performance of 3D vision systems while ensuring data privacy and real-time performance. For example, computationally intensive tasks can be processed through the cloud while real-time responsive tasks can be processed on edge devices, enabling efficient and secure data analysis.

4.3. Data privacy and ethical issues

In the biomedical field, AI-powered 3D vision technologies need to process large amounts of personal medical data, making data privacy and ethical issues particularly important.

4.3.1. Privacy protection of medical data

Patients' medical data involves personal privacy, especially sensitive information such as genetic data, medical records and imaging data. How to ensure data privacy and security during data sharing and analysis is a key issue.

Technologies such as data encryption, anonymization and differential privacy need to be adopted to protect patients' privacy. In addition, the establishment of comprehensive data security regulations to ensure the legitimacy and security of medical data is also an important safeguard to address privacy issues.

4.3.2. Transparency and trustworthiness issues for AI algorithms

The "black box" nature of AI algorithms, especially deep learning models, makes their decision-making process opaque. It is critical for doctors and patients to understand how AI reaches a certain conclusion. Lack of transparency and interpretability affects physicians' trust in AI systems.

Develop explainable AI technology to enhance the transparency and interpretability of models. By providing visual analysis of the decision-making process and the chain of model inference, it helps doctors and patients understand the decision-making basis of AI, thus enhancing their trust in AI technology.

4.3.3. Ethical issues of AI in medical decision making

AI is playing an increasingly important role in medical decision-making, but this raises ethical concerns. For example, should AI replace physician judgment in clinical decision-making? Who should be held accountable if AI makes mistakes?

Establish a clear responsibility framework and regulatory mechanism to ensure the rational use of AI technology in medicine. Collaboration between doctors and AI should be based on the principle of AI as a supplement and doctors as the mainstay, and AI should be used as a decision-support tool rather than a complete replacement of the doctor's role.

4.4. Algorithm bias and accuracy

AI algorithms can be affected by training data, model design, and algorithmic assumptions, which can create biases that affect their accuracy and fairness, especially when dealing with specific populations or specific diseases.

4.4.1. Algorithm bias in specific populations (e.g., race, gender)

AI models may not perform as well with some groups as others, especially in terms of race, gender, or age. Due to imbalances in healthcare datasets, AI models may be less predictive of conditions in certain groups or even produce discriminatory decisions.

Algorithmic bias can be reduced by adding diverse samples to the training data, especially by balancing data from underrepresented populations (e.g., ethnic minorities, women, elderly, etc.). In addition, the use of fairness assessment indicators to detect and correct algorithmic bias is also an effective way.

4.4.2. Misdiagnosis and misjudgment in high-risk decision-making

Misdiagnosis and misclassification can lead to serious consequences in some high-stakes medical decisions. AI algorithms may not be able to identify all lesions perfectly, especially in the detection of complex or rare conditions that may be missed or misclassified.

Enhanced training of AI models to ensure they have high generalization capabilities, combined with human intervention, ensures that the models are supported by clinical experts when making decisions. Continuous learning and optimization of the model can also improve its accuracy in high-stakes decision making.

4.4.3. How to improve the pervasiveness and generalization of algorithms

In order for AI algorithms to adapt to different types of patients and diseases, their ability to generalize over diverse datasets needs to be improved. However, due to issues such as data bias and insufficient samples, AI algorithms may not be able to generalize to specific application

5. Future directions

In the future, AI-driven 3D vision technology will usher in breakthroughs and innovations in multiple technical fields, promote interdisciplinary integration of biomedical engineering, and lay the foundation for the construction of intelligent medical environments. The following is an in-depth discussion of technological breakthroughs, interdisciplinary integration and the construction of intelligent medical environments.

5.1. Technological breakthroughs and innovations

5.1.1. New sensor and image acquisition technologies

One of the core of 3D vision technology is high-quality image acquisition. Although traditional imaging technologies such as CT and MRI have made significant progress, there are still some limitations in terms of resolution, contrast, scanning speed and patient comfort. In the future, the introduction of new sensors will drive breakthroughs in image acquisition technology and improve the quality and real-time performance of medical imaging.

5.1.2. Construction of multimodal fusion and intelligent diagnostic system

Medical imaging data is usually multimodal, e.g., CT, MRI, ultrasound, PET, and other imaging technologies provide data that each has different advantages. Fusion of these different modalities of imaging data not only provides a more comprehensive assessment of the patient's health status, but also provides a more accurate diagnostic basis for the AI model.

5.1.3. Optimization and adaptive enhancement of deep learning algorithms

Deep learning (DL) has been widely used in medical image analysis, but the optimization of models still faces challenges in the case of high-dimensional data, complex structures and few samples. In the future, deep learning algorithms will evolve towards adaptive learning to improve the generalization ability of models across different data, diseases and patient populations.

5.2. Interdisciplinary integration of biomedical engineering

5.2.1. Cross-domain collaboration to promote the integration of artificial intelligence and biomedical technologies

Future technological breakthroughs cannot be achieved without interdisciplinary cooperation. The booming development of artificial intelligence and the innovation of biomedical technology require close cooperation among many fields, such as engineering, medicine, and computer science, etc. Advances in AI technology can provide more accurate tools for medical imaging, oncology, neuroscience, etc., while the actual needs of the biomedical field provide application scenarios for the development of AI technology.

5.2.2. Close collaboration between medicine, computer science and engineering

Close collaboration between medicine, computer science and engineering will accelerate the development of AI applications in biomedicine. Medical experts will be able to provide clinical knowledge of diseases and practical needs, computer scientists will be able to develop and optimize deep learning algorithms for medical imaging, and researchers with engineering backgrounds will be responsible for developing new imaging devices and sensors.

5.3. Construction of intelligent medical environment

5.3.1. Smart hospitals and the future of digital healthcare

With the combination of AI with Internet of Things (IoT) and big data technologies, the construction of smart hospitals and digital healthcare will become a trend in the future development of the healthcare industry. Smart hospitals will utilize 3D vision technology to achieve accurate patient data collection and diagnosis, combined with AI systems for condition monitoring and remote support, to improve the quality and efficiency of medical services.

5.3.2. Possibility of real-time telemonitoring and telesurgery

With the development of 5G networks, AI and telemedicine technologies, remote monitoring and remote surgery will become possible. Doctors will be able to monitor, diagnose and, if necessary, provide remote surgical guidance to patients remotely through real-time video streaming and 3D image data of the patient.

In the future, AI-driven 3D vision technology will usher in technological breakthroughs and innovations in biomedical engineering. Through new sensor and image acquisition technologies, the construction of multimodal fusion and intelligent diagnostic systems, the optimization of deep learning algorithms and the enhancement of self-adaptive capabilities, AI will greatly promote the accuracy and efficiency of medical image analysis. Interdisciplinary cooperation will become the key to the integration of AI and biomedical technology, while the construction of intelligent medical environments will provide patients with more personalized, precise and efficient medical services. With the continuous advancement of these technologies, the future of medicine will enter a new era of greater intelligence and personalization.

6. Conclusion

Artificial intelligence (AI)-driven 3D vision technology in biomedical engineering is advancing rapidly and demonstrates immense potential. With ongoing breakthroughs in AI and 3D vision technology, their integration in medical imaging analysis, surgical navigation, pathology research, and health management has significantly enhanced precision, efficiency, and personalization in the medical field. By providing more detailed spatial information, 3D vision technology assists physicians in diagnosing diseases, planning surgeries, and formulating treatment strategies with greater accuracy, while AI plays a pivotal role in image analysis, pattern recognition, and automated tasks.

However, despite the remarkable advantages of AI-driven 3D vision technology in medicine, several challenges remain. Issues such as image quality and data preprocessing, computational resource demands, data privacy and ethical concerns, and algorithmic biases and accuracy must be addressed through technological innovation and sound policies. In the future, with the continuous advancement of novel sensors and image acquisition technologies, the optimization of deep learning algorithms, and close interdisciplinary collaboration, 3D vision technology is expected to achieve broader and deeper applications in the medical field.

Looking ahead, AI-driven 3D vision technology will play an increasingly critical role in the intelligent healthcare environment. Smart hospitals, digital healthcare, personalized treatment, and remote monitoring will become key directions for future medical development. Cross-disciplinary collaboration will foster the deep integration of AI and biomedical engineering, driving the realization of intelligent diagnostic systems, precise medical solutions, and efficient remote surgeries. With continuous technological progress and innovation, we have every reason to believe that the combination of AI and 3D vision technology will bring revolutionary changes to the global healthcare industry, benefiting more patients and contributing significantly to the sustainable development of the global medical system.

In conclusion, AI-driven 3D vision technology is not only a vital development direction in biomedical engineering but also a key pillar for building future intelligent healthcare systems. Facing both challenges and opportunities, researchers, technology developers, and policymakers should work hand in hand to promote the application and popularization of this technology, making greater contributions to human health.

Author contributions: Conceptualization, LW and ZZ; methodology, LW; software, LW; validation, LW and ZZ; formal analysis, LW; investigation, LW; resources, LW; data curation, LW; writing—original draft preparation, LW; writing—review and editing, LW; visualization, ZZ; supervision, LW; project administration, LW; funding acquisition, ZZ. All authors have read and agreed to the published version of the manuscript.

Ethical approval: Not applicable.

Conflict of interest: The authors declare no conflict of interest.

References

1. Liu, Y., et al. (2021). “Deep learning in 3D medical image analysis: a comprehensive review.” Journal of Healthcare Engineering, 2021, 1-18. DOI: 10.1155/2021/6675851
2. Wang, S., et al. (2021). “Artificial intelligence for medical image analysis in radiology: recent advances and future prospects.” Nature Biomedical Engineering, 5, 248-262. doi: 10.1038/s41551-020-00650-6
3. Roth, H. R., et al. (2021). “AI-driven three-dimensional reconstruction of medical images for personalized treatment planning.” Frontiers in Artificial Intelligence, 4,. 582709. DOI: 10.3389/frai.2021.582709
4. Liu, J., et al. (2021). “3D convolutional neural networks for enhanced medical image segmentation.” IEEE Transactions on Biomedical Engineering, 68(5), 1504-1513. doi: 10.1109/tbme.2020.3017021
5. Zhao, S., et al. (2022). “AI-based multi-modal 3D medical image analysis and its applications.” Medical Image Analysis, 74, 102229. DOI: 10.1016/j.media.2021.102229
6. Xu, S., et al. (2022). “Deep learning-based 3D medical image reconstruction and its clinical applications.” Frontiers in Bioengineering and Biotechnology, 10, 103628. doi: 10.3389/fbioe.2022.103628
7. Chen, C., et al. (2022). “AI-driven 3D image analysis for early-stage cancer detection.” Journal of Medical Imaging, 29(1), 115-124. doi: 10.1117/1.jmi.29.1.115
8. Gupta, A., et al. (2022). “Advanced deep learning methods in 3D medical image segmentation.” Computers in Biology and Medicine, 143, 105235. DOI: 10.1016/j.compbiomed.2021.105235
9. Zhou, J., et al. (2023). “Artificial intelligence for 3D medical image analysis in robotic surgery systems.” IEEE Transactions on Robotics, 39(1), 20-33. doi: 10.1109/tro.2022.3153913

10. Li, W., et al. (2023). “3D convolutional neural networks for medical image reconstruction and diagnostic prediction.” Medical Image Analysis, 90, 102777. DOI: 10.1016/j.media.2023.102777
11. Zhang, L., et al. (2023). “Deep learning-based 3D tumor segmentation and quantification in PET-CT images.” Journal of Nuclear Medicine, 64(4), 589-595. doi: 10.2967/jnumed.122.263828
12. Li, X., et al. (2023). “AI and 3D imaging: novel approaches to brain tumor segmentation.” Neurocomputing, 521, 108-118. DOI: 10.1016/j.neucom.2022.10.070
13. Wang, F., et al. (2023). “AI-driven 3D imaging for minimally invasive surgery.” Journal of Surgical Research, 298, 44-53. DOI: 10.1016/j.jss.2023.03.013
14. Kang, H., et al. (2023). “3D deep learning for the detection and analysis of breast cancer from mammographic images.” Artificial Intelligence in Medicine, 128, 102301. DOI: 10.1016/j.artmed.2023.102301
15. Xu, W., et al. (2024). “Multimodal 3D medical image fusion using deep learning techniques.” IEEE Transactions on Medical Imaging, 43(2), 678-689. doi: 10.1109/tmi.2023.3231069
16. Yang, Z., et al. (2024). “3D image reconstruction and AI-based automated feature extraction for lung cancer diagnosis.” Journal of Cancer Research and Clinical Oncology, 150(5), 1371-1382. DOI: 10.1007/s00432-024-04573-w
17. Zhou, Y., et al. (2024). “Deep learning for 3D cardiovascular image analysis: from segmentation to diagnosis.” Computerized Medical Imaging and Graphics, 97, 102097. DOI: 10.1016/j.compmedimag.2023.102097
18. Huang, Y., et al. (2024). “AI-enhanced 3D image-guided brain tumor resection: a review of current trends and future directions.” Journal of Neuro-Oncology, 160(1), 1-14. doi: 10.1007/s11060-024-04223-9
19. Zhang, H., et al. (2024). “AI-driven three-dimensional cardiac image analysis for personalized treatment planning.” IEEE Journal of Biomedical and Health Informatics, 28(3), 678-689. doi: 10.1109/jbhi.2023.3245949
20. Li, X., et al. (2024). “Robust AI-based 3D image analysis for automated diagnosis of osteoarthritis.” Journal of Medical Systems, 48(2), 78. doi: 10.1007/s10916-023-01914-x