Article - Engineering, Technology and Techniques

Image Segmentation Method for Athlete Knee Joint Injury Using Transformer Model by MIoT

Tao Jiang[1]
https://orcid.org/0009-0009-0148-4290

Feng Feng[1]*
https://orcid.org/0000-0002-1079-6677

Yu Cao[2]
https://orcid.org/0009-0008-9784-992X

Haitao Yang[3]
https://orcid.org/0000-0002-0381-1306

Raja Soosaimarian Peter Raj[4]
https://orcid.org/0000-0002-7216-2207

[1]Capital Normal University, Department of Physical Education, Beijing, China; [2]Renmin University of China, Physical Education Department, Beijing, China; [3]Beijing University of Technology, Physical Education Department, Beijing, China; [4]Vellore Institute of Technology, School of Computer Science and Engineering, Vellore, India.

*Correspondence: fengfeng6011@126.com (F.F.).

HIGHLIGHTS

- An image segmentation for athlete knee joint injury using the transformer model is proposed.

- Reducing the complexity of transformer model training.

- The comprehensive performance of this paper's method is verified by several evaluation criteria.

Abstract: The segmentation of athlete knee joint injury images can provide doctors with information about the location and extent of the athlete knee joint injury. Therefore, it is significant to segment the images of athlete knee joint injury. However, the traditional image segmentation method of athlete knee injury has the problems of low accuracy of mask region extraction, completion time of extraction and high error rate of segmentation. In the paper, we propose image segmentation method for athlete knee joint injury using the transformer model by the medical Internet of Things (MIoT). First, the MIoT was used as a way to obtain images of knee joint injury of athletes, and the images of knee joint injury of athletes were derived using the MIoT. Second, the exported image is input into the shadow expansion layer of the transformer model, which performs shadow expansion on the athlete knee joint injury image to obtain its mask region, and then the image is input into the patch embedding layer. Finally, after the patch embedding layer extracts the mask patch of the athlete knee joint injury image, the mask patch is input into the transformer block for down-sampling and up-sampling processing, and then the athlete knee joint injury image segmentation result is

output using the end backpropagation layer. The results show that the proposed method has a low error rate in extracting the mask region from the knee joint injury image of athletes, and a short completion time for extracting the mask region, the most detailed and comprehensive segmented athlete knee joint injury image, the maximum error rate of image segmentation is only 6.8%, and the maximum value of segmentation time is only 3.96s. It has important research value in the field of athlete knee joint injury diagnosis.

INTRODUCTION

In all sports or competitions, the knee joint plays an important role, the biggest role is the role of the top and bottom. The knee joint is divided into four parts: femur, tibia, shyness, and fat bone [1-2]. Its interior is mainly composed of nerves and muscles, articular cartilage, associated ligaments, muscle keys, meniscus, and synovial layer. The muscles, the medial and lateral collateral ligaments, the anterior and posterior cruciate ligaments, the patella, and the meniscal tissues that surround the knee joint form an important system that maintains the stability of the knee joint. Therefore, it can be seen that it is impossible for the knee joint to perform relevant movements on other axes than the decubitus movement on the frontal axis, and if this scientific principle is violated, then it is bound to suffer relevant injuries [3]. At present, the diagnosis of knee injuries in athletes is using imaging techniques, such as CT scans and MRI [4], which can generate multi-angle knee joint injury images to provide images for the diagnosis of knee injuries in athletes and facilitate the diagnosis of the extent of their injuries. It is significant to segment the images of knee injuries in athletes, which can provide doctors with relevant information about the location and extent of knee injuries in athletes. Therefore, it is important to segment the image of athlete knee joint injury, which can provide information about the location and extent of the athlete knee joint injury [5]. However, in the process of athlete knee joint injury image segmentation, the difficulty of image segmentation increases due to the difficulty of determining the mask region, which causes a very serious obstacle to the athlete knee joint injury image segmentation. Thus, it is important to study a new athlete knee joint injury image segmentation method.

There are many scholars studying image segmentation methods for knee injuries. Zhang, Y. and Tian, Y [6] proposed an image segmentation method based on fractional variable-order differential, which is based on fractional variable-order partial differential equations to establish a knee joint injury image segmentation model and segment the knee joint injury image based on its different pixel grayscale values. However, this method is affected by the iterative generalization performance of the model in the application process, and the extraction accuracy of the mask region of the knee joint injury image of athletes is low. Basar S and coauthors [7] proposed a defocus image segmentation method. The method used the local binary mode to present the sharpness measurement of the athlete knee joint injury image after image enhancement pretreatment of the athlete knee joint injury image, and then used the pulse-coupled neural network to output the segmentation results of the athlete knee joint injury image. However, this method is affected by the blurred region of the athlete knee joint injury image, which leads to a long completion time for the extraction of the mask region of the athlete knee joint injury image. Panfilov E and coauthors [8] proposed a medical image segmentation method based on deep learning. This method uses 3D dual-echo steady-state (DESS) MRI data set to obtain experimental sample data, uses bilateral filtering method to enhance the sample image, uses deep learning to segment the sample image, uses multi-spectral registration to partition it, and extracts the volume and thickness of the partition, which lays a solid foundation for subsequent research. However, in practical application, it is found that the segmentation image of knee joint injury of athletes is not detailed and comprehensive. Shi Y and coauthors [9] proposed a medical image segmentation method based on semi-supervised learning. The medical image segmentation network framework is built under the support of semi-supervised learning. The network framework is composed of three parts: conservative radical module, specific region segmentation network and uncertain region segmentation network. The network framework is trained by self-training strategy, and sample images are input into the network framework to obtain relevant segmentation results. However, this method has the problem of high segmentation error rate of athletes' knee joint injury image, and the actual application effect is not good. Wen Y and coauthors [10] proposed a medical image segmentation method based on pixel by pixel triad learning. This method is aimed at traditional medical image segmentation methods, which mainly use complex structures to enhance the boundary details, which will increase the computational burden of reasoning, and cannot be embedded into the latest segmentation architecture as a general boundary intensifier. Therefore, pixel-level triplet loss is used to enable the segmentation model to learn more discriminative feature representations at the boundary, and it can be easily incorporated into the latest segmentation network as a general boundary enhancer to complete the relevant

medical image segmentation tasks. However, this method is too complex and exists, the problem of long segmentation time of athletes' knee joint injury image. The medical Internet of Things (MIoT) is the product of digital management in the medical industry. It is a management system based on wireless communication network and integrated with computer intelligent identification and management technology. It makes the medical business gradually turn to the perspective of object-oriented management and further promotes the development of information in the medical industry. The MIoT plays a role in business management, information standardization and sharing in the medical field. Therefore, MIoT is the basis for the development of the medical industry. The transformer model is the encoder responsible for image segmentation in the convolutional neural network model. The transformer model can maintain the high-precision image segmentation capability without changing the number of model parameters. Therefore, this paper combines the advantages of the fast computing speed of the MIoT and the high segmentation accuracy of the transformer model, deeply integrates the two, applies them to the segmentation process of the athlete knee joint image, and proposes the segmentation method of the athlete knee joint injury image of the transformer model under the MIoT.

The main contributions of this paper are as follows:(1) The image shadow expansion method is introduced to increase the shadow area of the athlete knee joint image and obtain its final mask region, which addresses the problems of low accuracy and long extraction time of the mask region of the athlete knee joint injury image existing in traditional methods; (2) Using a variable-length sliding window approach, envelope patch subsets are obtained. By reducing the scale of the mask patch, it is possible to mitigate the memory overhead and training complexity to some extent. This enables faster segmentation of athlete knee joint injury images, thereby addressing issues associated with inadequate and incomplete segmentation in traditional methods, as well as high segmentation error rates and lengthy segmentation times. (3) To validate the comprehensive performance of the proposed method in image segmentation of athletes' knee injuries from five aspects based on two data sets.

METHODOLOGY

The proposed method uses the MIoT as a way to obtain the image of the athlete knee joint injury. After exporting the image of the athlete knee joint injury using the MIoT, the image is input into the shadow expansion layer of the transformer model, which performs shadow expansion processing on the athlete knee joint injury image to obtain its mask region, and then inputs the image into the patch embedding layer. After the patch embedding layer extracts the mask patch of the athlete knee joint injury image, the mask patch is input to the transformer block for down-sampling and up-sampling processing, and the athlete knee joint injury image segmentation result is output using the end backpropagation layer.

Construction of image segmentation model for knee joint injury in athletes

Based on the MIoT and combined with the transformer model, the image segmentation model of knee joint injury was built. The overall structure of the model is shown in Figure 1. This paper takes MIoT as the basis, in MIoT, the wireless network of the hospital is deployed using the ultra-wide band signal distribution system and connected to the hospital intranet switch to realize the overall network coverage within the hospital. The hospital intranet switch is connected to WiFi/IoT base station, and the application layer of MIoT is connected through WiFi/IoT base station. Within the application layer of MIoT, there exist various application services such as mobile nursing, cold chain management, infusion monitoring and so on. The data intelligent control management unit acquires the relevant athlete knee joint injury images and realizes image transmission and calculation by the ultra-wide band signal distribution system, the network switch in the hospital and the WiFi/IoT base station to ensure that the subsequent image segmentation processing of the athlete knee joint injury can be performed smoothly. The MIoT architecture is shown in Figure 2
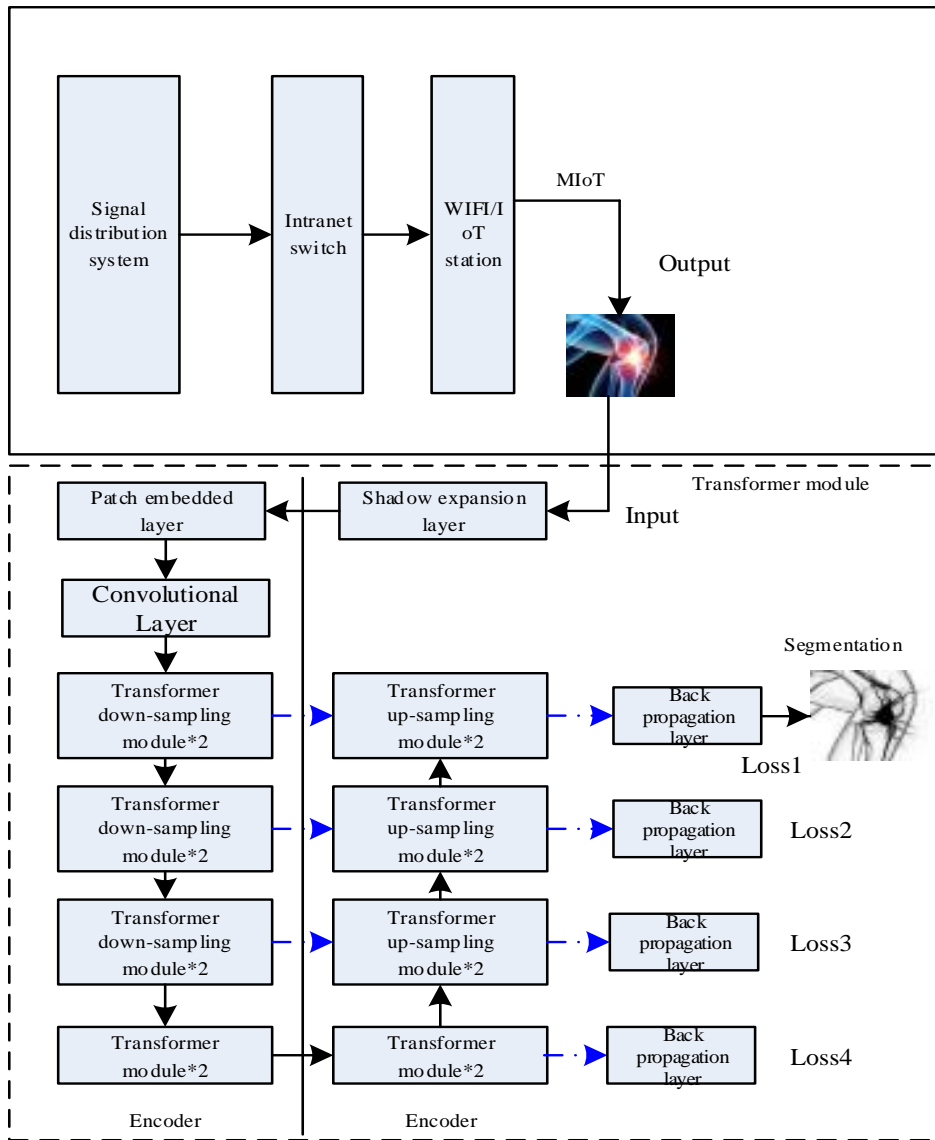
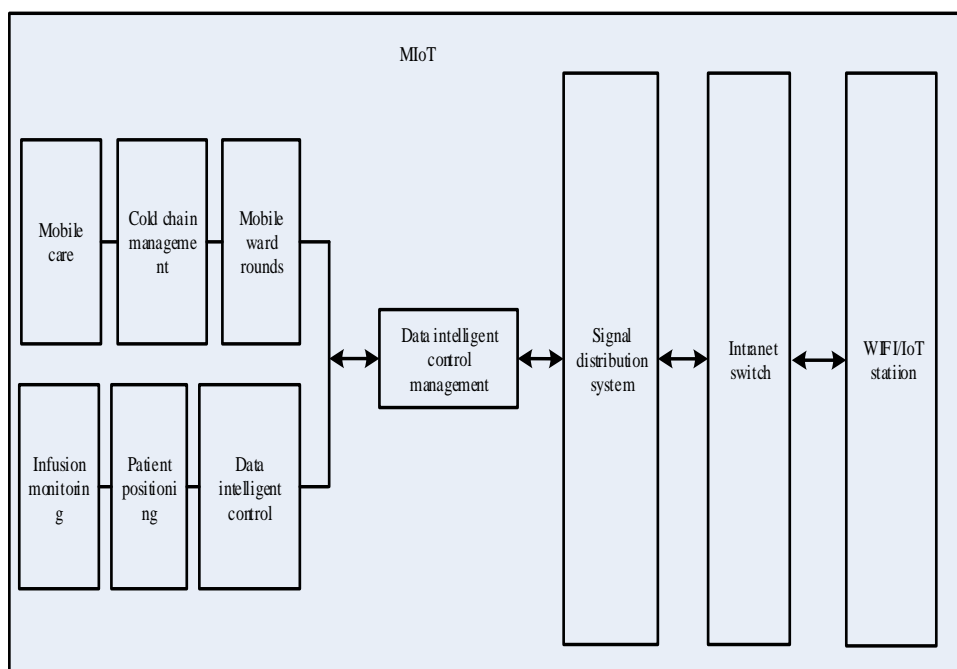**Figure 1.** The overall structure of the model

Figure 2. Architecture of MIoT

After acquiring and exporting the athlete knee joint injury image using the Data Intelligence Management Unit, it is input into the shadow expansion layer of the transformer model, which performs shadow expansion on the athlete knee joint injury image, and after obtaining the mask region of the athlete knee joint injury image, the athlete knee joint injury image is input into the patch embedding layer. The athlete knee joint injury image is divided into image blocks with uniform resolution using the embedding layer, and all the athlete knee joint injury image blocks are mapped into the corresponding dimensions using the convolution layer. Then the athlete knee joint injury image blocks are input into the transformer down-sampling module, and the contextual features of the athlete knee joint injury image blocks are obtained after three transformer down-sampling modules, and then the features are input into the three transformer up-sampling modules

The transformer up-sampling module is used to up-sample the athlete knee joint injury image. Each up-sampling module has a back-propagation layer. The back-propagation layer is used to output the segmentation result of the athlete knee joint injury image of the current module. Each back-propagation layer has a loss value. The accuracy of the output result of the transformer up-sampling module is judged by the loss value, the output result of the back propagation layer at the end is taken as the segmentation result of the athlete knee joint injury image.

Shadow expansion for knee joint injury image of athletes

After the athlete knee joint injury image is obtained using MIoT, the athlete knee joint injury image shadow expansion is the basis of this image segmentation [11], here the athlete knee joint injury image shadow expansion is implemented based on HSV color space algorithm to generate the athlete knee joint injury image mask region, and its detailed process is as follows.

The RGB colors of an athlete knee joint injury image are converted to HSV colors with the following equation.

$$H = \begin{cases} \theta, B \le G \\ 360° - \theta, B > G \end{cases} \tag{1}$$

$$S = \frac{R+G+B-3}{R+G+B} \min(R,G,B) \tag{2}$$

$$V = \frac{(R+G+B)}{3} \tag{3}$$

where R . G and B represent the gray values of red, green and blue pixels, respectively; H . S and V represent hue, saturation and brightness, respectively; θ is the hue measurement value.

In the knee joint injury image of athletes, the shade of the shadow area of the injury position is slightly higher, and its brightness will be lower [12,13]. Therefore, the threshold method is used to distinguish the shadow area and non-shadow area of the knee joint injury image of athletes.

Let $R(x, y)$ be the ratio diagram of knee joint injury image of athletes. In which, x and y represent coordinate pixel points, and their calculation formula is as follows:

$$R(x, y) = \left(H(x, y)+1\right)\frac{1}{V(x, y)+1} \tag{4}$$

where $V(x, y)$ denotes the normalized brightness image of knee joint injury image of athletes. $H(x, y)$ is the normalized tone map of the knee joint injury image of athletes.

Let T be the shadow detection threshold of knee joint injury image of athletes, and use OTSU method[14] to calculate the threshold, and its expression formula is as follows:

$$T = \arg\min\left[\sum_{i=1}^{T} P(i)(i - \Phi_1)^2\right] + \arg\min\left[\sum_{i=T+1}^{255} P(i)(i - \Phi_2)^2\right] \tag{5}$$

where i is the value in $R(x, y)$. $P(i)$ is the probability of occurrence of value i in $R(x, y)$. Φ_1 and Φ_2 are the average weighted value of the target and background in the knee joint injury image of athletes, respectively; T represents the maximum value of the value i .

Taking the result of formula (5) as the judgment basis, the result of equation (4) is compared and analyzed with the result of (5).When it is higher than the shadow detection threshold, it is determined that the current area is the shadow area of the athlete knee joint injury image [15]. After dividing the shadow area of the athlete knee joint injury image, the mask region is obtained.

Transformer model-based image segmentation method for knee joint injury of athletes

Based on the above steps, the segmentation method process of knee joint injury image of athletes is designed, as follows:

Input: athlete knee joint injury image

Output: segmentation result of knee joint injury image

After the athlete knee joint injury image is exported using the MIoT, the shadow expansion processing and the mask patch extraction of the transformer model are carried out on it, and then the encoder and decoder of the transformer model are used to carry out the down-sampling and up-sampling processing respectively, and the segmentation result of the athlete knee joint injury image is output using the back-propagation layer.

The transformer model [16] is the most important part in the segmentation process of athletes' knee joint injury image. It is mainly responsible for extracting the characteristics and segmentation of athletes' knee joint injury image. The transformer model is composed of encoder, decoder and back-propagation layer.

(1) In the patch embedding layer of the transformer model, the mask region of the athlete knee joint injury image is divided into several mask patch in the form of variable-length sliding window, which is the block of the athlete knee joint injury image. First set the length of the sliding long window and the sliding short window, use the sliding long serial port to expand the mask segment of the athlete knee joint injury image [17-18], and then use the sliding short window to obtain the patch block of the athlete knee joint injury image. The expression formula is as follows:

$$N_{(m-patch)} = \left\lfloor f^{(L)}(\frac{U_1 - w_l}{s_l} + 1) f^{(S)}(\frac{U_0 - w_s}{s_s} + 1) \right\rfloor \quad （6）$$

where $N_{(m-patch)}$ represents the number of large patch in the knee joint injury image of athletes under the action of sliding window. U_0 and U_1 are the width of the initial mask segment and the mask segment after the window slides, respectively; w_l and s_l represent the width and step length of the sliding long window, respectively. w_s and s_s represent the width and length of the sliding short window, respectively. $f^{(L)}$ and $f^{(S)}$ are sliding long window and sliding short window functions, respectively.

Let y be the patch tag value of the athlete knee joint injury image, and set the constraint conditions of formula (6) as follows：

When y_i is1, there are

$$f^{(L)}(\Omega) = \Omega \quad （7）$$

When y_i is not1, there are

$$f^{(L)}(\Omega) = 1 \quad （8）$$

where Ω denotes variable in sliding window.

According to the constraint conditions of Formula (7) and (8), using Formula (6), we can obtain several patches of the mask of the knee joint injury image of athletes.

(2) The transform model encoder part calculates the transform window boundary of each athlete knee joint injury image using self-attention, and interacts with the upper sampling module in its decoder [19-20]. There are two continuous transformer models in each down-sampling module in the encoder. The first transformer model is responsible for the general window configuration, and the second transformer model is responsible for the sliding window configuration. The image features of knee joint injury of athletes outputted from the two windows in the l transformer model are represented by C^l and $\hat{C}^l$, respectively. The calculation formula is as follows：

$$\hat{C}^l = W - MSA(LN(C^{l-1}) + C^{l-1}) \quad （9）$$

$$C^l = MLP(LN(\hat{C}^l) + \hat{C}^l) \quad （10）$$

where $W - MSA$ is the general window of the transformer model; $MLP(\cdot)$ is the back propagation layer function; $LN(\cdot)$ denotes the natural logarithm function returned; C^{l-1} is the image features of knee joint injury of athletes output from the first window in the $l-1$ transformer model; The image features of knee joint injury of athletes outputted from the two windows in the $l+1$ transformer model are represented by C^{l+1} and $\hat{C}^{l+1}$,

respectively. The calculation formula is as follows：

$$\hat{C}^{l+1} = SW - MSA(LN(C^l) + C^l)$$

(11)

$$C^{l+1} = MLP(LN(\hat{C}^{l+1}) + \hat{C}^{l+1})$$

（12）

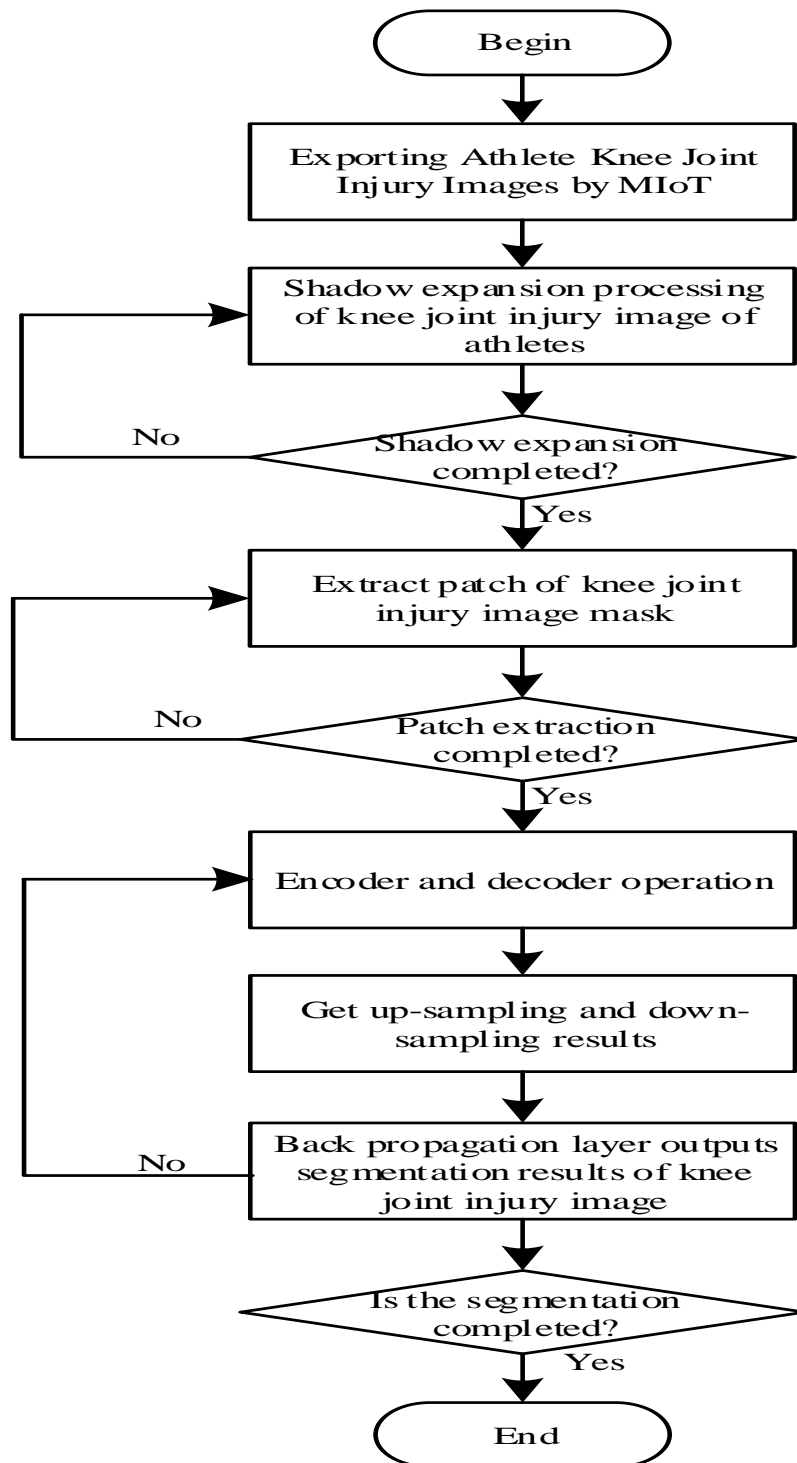where $SW - MSA$ represents the sliding window of the transformer model.

Figure 3. Process of image segmentation model for knee joint injury of athletes

(3)In the back propagation layer of the transformer model, there are two full connection layers. The first full connection layer is responsible for increasing the image channel of the athlete knee joint injury, and the second full connection layer is responsible for reducing the channel [21-22]. After the two full connection layer channels are added and reduced, the features of the athlete knee joint injury image are obtained.

Let q, k, and v be the query, key, and value matrices, respectively. The detailed calculation process of the back-propagation layer is as follows：

$$[q,k,v] = D\left[\varpi_q, \varpi_k, \varpi_v\right] \qquad (13)$$

where ϖ_q. ϖ_k and ϖ_v are the weight matrix of their corresponding matrix, respectively. D refers to the matrix packed together with patch of the knee joint injury image mask of athletes;

(4) In the back propagation process, the matrix normalization function is as follows：

$$SA(D) = Soft\max(\frac{qk^T}{\sqrt{\vartheta}})v \qquad (14)$$

where ϑ represents the feature dimension of mask patch; The superscript T is the normal distribution; $Soft\max(\cdot)$ represents the normalized exponential function.

(5) The output of the final segmentation result of the athlete knee joint injury image is as follows：

$$MSA(D) = \left[SA_1(D), SA_2(D), L, SA_t(D)\right] \qquad (15)$$

where $SA_1(D), SA_2(D), L, SA_t(D)$ represents the segmentation results of the normalized mask patch of the knee joint injury image of athletes.

The process of image segmentation method for knee joint injury of athletes is shown in Figure 3.

EXPERIMENTAL RESULTS AND ANALYSIS

Data sets

The experimental environment of this paper is based on windows10 64-bit system, memory is 16GB, and the GPU version is NVIDIA GeForce RTX 4080 16GB. Deep learning framework is TensorFlow to implement the training of models. The experimental data set is divided into two parts, one is MRNet data set and the other is SKI10 data set. **MRNet data set**: This data set includes 1370 magnetic resonance imaging (MRI) examination data of knee joint. There were 1104 (80.6%) cases of abnormal examination, including 319 (23.3%) cases of anterior cruciate ligament (ACL) tear and 508 (37.1%) cases of meniscus tear. This data set is further divided into training set (1130 images), verification set (120 images) and test set (120 images). **SKI10 data set:** This dataset aims to compare different algorithms for cartilage and bone segmentation in knee MRI data. Knee cartilage segmentation is a clinically related segmentation problem, which has received considerable attention in recent years. In which, it is used to quantify cartilage degeneration to diagnose osteoarthritis and optimize the surgical plan of knee implants. The data set consists of 2650 digital X-ray images of knee joint.

The data in the MRNet data set and the SKI10 data set were integrated and processed, and some of the lower quality data were removed to obtain three thousand images of knee joint injuries of athletes. Two thousand and four hundred images were used as the training set, and the rest six hundred images were used as the test set. Before the experimental test, it is necessary to unify the size of the input image specifications of the transformer model. The image size is unified to 720 * 720. Input the image of the test set to the simulation software for debugging before the experiment to obtain the optimal operating parameters. This parameter is used as the experimental parameter in the subsequent testing process, and the data from the experimental set is input to the simulation software for relevant testing, with the iteration parameter set to 3000, the number of encoders and decoders and the number of reverse replay layers set to 4, and the number of sliding windows set to 10, to obtain more reliable experimental results.

A certain athlete knee joint injury image in MRNet dataset was used as the experimental object, and its mask region was delineated using the proposed method, and the delineation results are shown in Figure 4.
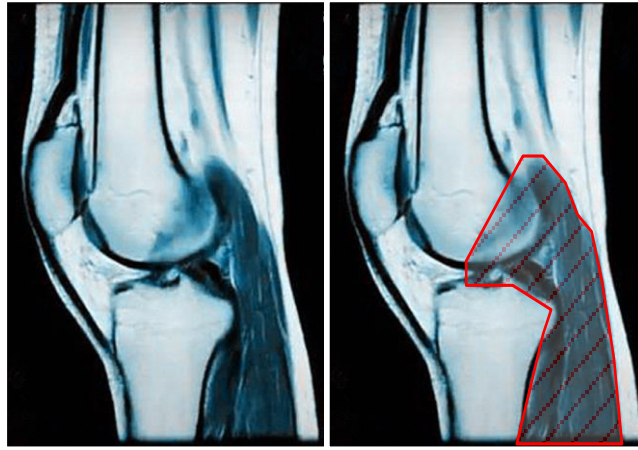
Figure 4. The result of the mask region division of the knee joint injury image of athletes

It can be seen from the analysis of Figure 4 that the proposed method can effectively expand the shadow area of the knee joint injury image of athletes, and obtain its mask region. The large knee joint injury image mask region of proposed method for athletes completely overlaps with their original image shadow area. The results show that proposed method can obtain the mask region of the knee joint injury image of athletes, and the results are more accurate, which lays a good foundation for the subsequent acquisition of mask patch.

The mask patch nugget of the above image is obtained using the proposed method, and the results are shown in Figure 5. It can be seen from the analysis of Figure 5 that when the proposed method extracts the mask patch of the knee joint injury image of athletes, the original image is divided into four small blocks after sliding through the window, and the four effects are divided into four small mask patch after sliding again. The results show that the proposed method has a good ability to obtain mask patch.

Figure 5. Mask patch block acquisition

Evaluation criteria

To verify the effectiveness of the transform model athlete knee joint injury image segmentation of proposed method under the MIoT, the ISFVO [6] method, the NODIS [7] method, the DLSK [8] method, the IAUEIS [9] method and the PWMIS [10] method are used as experimental comparison methods. By comparing different methods of athlete knee joint injury image mask region extraction accuracy, athlete knee

joint injury image mask region extraction completion time, athlete knee joint injury image segmentation effect, athlete knee joint injury image segmentation error rate, and athlete knee joint injury image segmentation time as experimental indexes, we verify the practical application effect of different methods.

The formula for the indicator of the extraction accuracy of the masked region of the athlete knee joint injury image is as follows

The calculation formula for the error rate of extracting the mask region from the knee joint injury image of athletes is as follows:

$$R = \frac{|r_i - r_j|}{r_j} \times 100\%$$ （16）

where r_i represents the extracted mask region; r_j is the actual mask region.

The completion time of extracting the mask region of the knee joint injury image of athletes. The calculation formula is as follows

$$T = \sum_{i=1}^{n} t_i$$ （17）

where t_i represents the time taken to extract the mask region of the i–th athlete knee joint injury image.

The segmentation effect of athlete knee joint injury image refers to that the closer the segmentation result is to the actual athlete knee joint injury area, the better the segmentation effect is.

The calculation formula of segmentation error rate of knee joint injury image of athletes is as follows

$$E = \frac{Y - C}{Y} \times 100\%$$ （18）

where Y represents the total number of experiments. C is the number of correct segmentation.

The calculation formula for segmentation time of knee joint injury image of athletes is as follows

$$G = |g_2 - g_1|$$ （19）

where g_1 and g_2 are the start and end time of the split, respectively.

RESULTS AND DISCUSSION

The comparison of error rates in extracting the mask region of knee joint injury images for athletes is shown in Figure 6. With an increasing number of experiments, the error rates of mask region extraction for the six methods fluctuated. Among them, when the number of experiments reached 10, the proposed method achieved the lowest error rate of 4% in mask region extraction. In comparison, the error rates of mask region extraction were 36% for the method in ISFVO [6], 13% for NODIS [7], 18% for DLSK [8], 10% for IAUEIS [9], and 7% for PWMIS [10]. The proposed method showed a 32%, 9%, 14%, 6%, and 3% lower error rate compared to the methods in ISFVO [6], NODIS [7], DLSK [8], IAUEIS[9], and PWMIS [10]， respectively. These results indicate that the proposed method achieves a lower error rate in mask region extraction for athlete knee joint injury images, demonstrating its superior segmentation capability in this context.
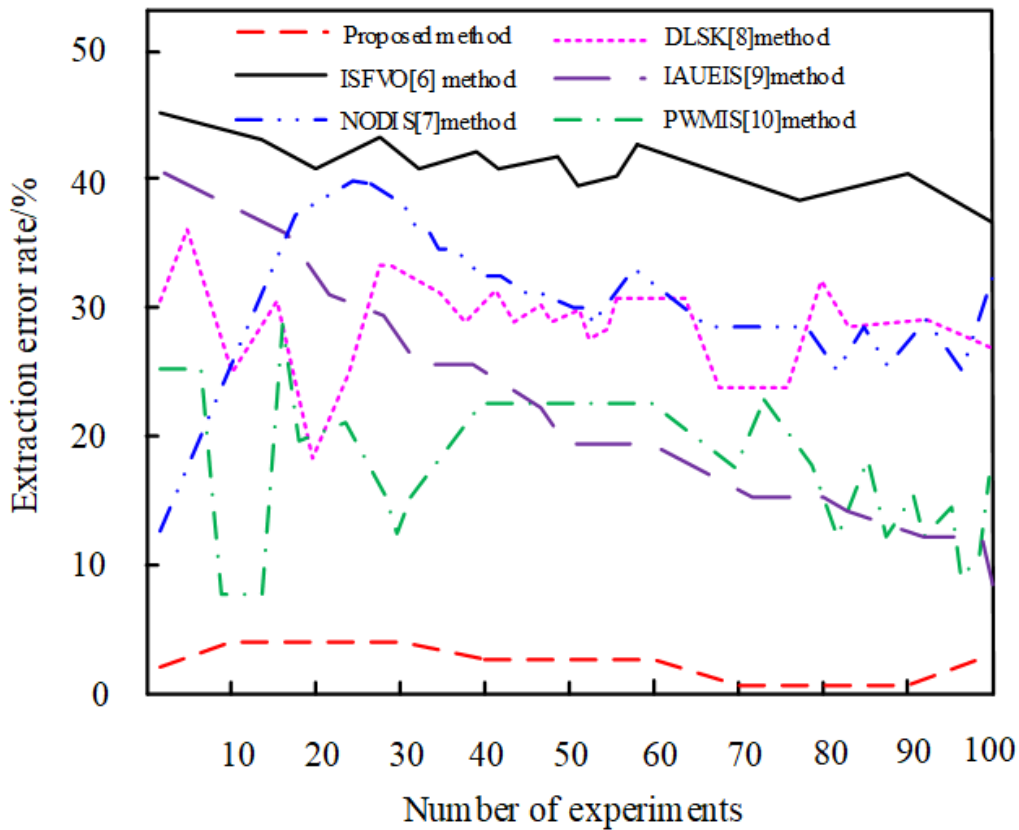
Figure 6. Comparison of error rates in extracting mask regions from knee joint injury images of athletes

The comparison of the extraction completion time of the mask region of the knee joint injury image of athletes are shown in Table 1. By analyzing the data in Table 1, it can be seen that with the increase of the number of experimental samples, the extraction completion time of the mask region of the knee joint injury image of athletes in the six methods has an upward trend. The average value of the extraction completion time of the mask region of the proposed method is 1.57s, which is 0.4s lower than the method in ISFVO [6], 2.73s lower than the method in NODIS [7], 4.18s lower than the method in DLSK [8], 5.21s lower than the method in IAUEIS [9], and 1.67s lower than the method in PWMIS [10]. It shows that compared with the experimental comparison method, the proposed method has a shorter completion time and higher efficiency for the extraction of the knee joint injury image mask region of athletes.

Table 1. Comparison of extraction time of mask region of knee joint injury image of athletes

Number of samples	ISFVO [6] method /s	NODIS [7] method /s	DLSK [8] method /s	IAUEIS [9] method /s	PWMIS [10] method /s	Proposed method/s
500	1.26	1.96	2.64	3.97	2.21	0.78
1000	1.48	2.55	3.84	4.87	2.51	1.22
1500	1.69	3.69	4.58	5.74	2.98	1.64
2000	2.14	4.77	5.99	6.87	3.69	1.87
2500	2.58	5.84	7.84	8.99	3.94	1.92
3000	2.69	6.99	9.61	10.26	4.16	2.01
Average	1.97	4.30	5.75	6.78	3.24	1.57

Taking the knee joint injury image of an athlete in the MRNet dataset as the experimental object, different methods are used to segment it. It can be seen from the analysis of Figure 7 that among the six methods, the athletes' knee joint injury image segmented by proposed method is the most detailed and comprehensive, and the smaller and less obvious knee joint injury areas are effectively segmented. However, the method in ISFVO [6] has the problem of missing the details of the knee joint injury area, and the coverage of the segmentation area is insufficient. The method in NODIS [7] can divide most of the knee joint injury areas. However, the coverage is still not high. The method in DLSK [8] can only divide the largest part of the knee

joint injury area, and it is difficult to determine other knee joint injury areas. The methods in IAUEIS [9] and PWMIS [10] can only divide most of the knee joint injury areas. However, the coverage is still not high, and the image segmentation effect is poor. The above results show that the proposed method has a relatively significant ability of image segmentation of knee joint injury of athletes, and the application effect is better.
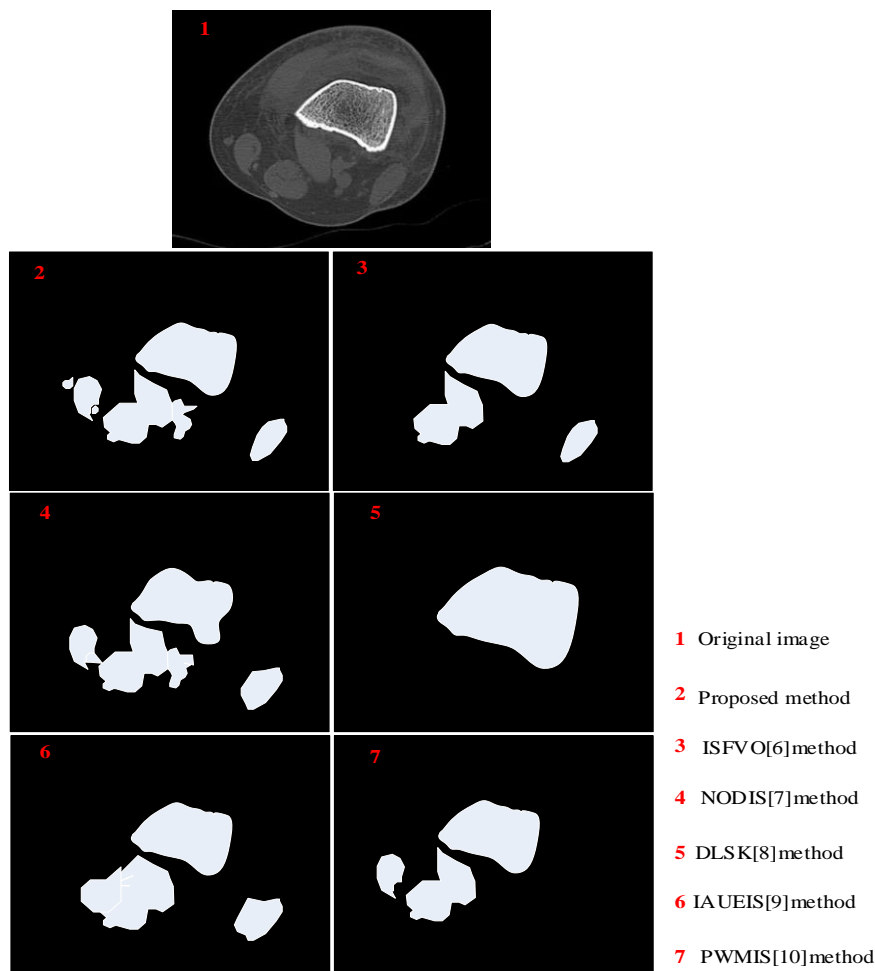

1 Original image

2 Proposed method

3 ISFVO[6]method

4 NODIS[7]method

5 DLSK[8]method

6 IAUEIS[9]method

7 PWMIS[10]method

Figure 7. Comparison of image segmentation of knee joint injury of athletes

The comparison of segmentation error rates of knee joint injury images of athletes using six methods are shown in Figure 8. According to the results in Figure 8, we can see that the maximum segmentation error rate of athlete knee joint injury image in this method is only 6.8%, and the maximum segmentation error rate of athlete knee joint injury image of proposed method in ISFVO [6] is 36.4%. The maximum segmentation error rate of athlete knee joint injury image in the method of NODIS [7] is 26.9%, the maximum segmentation error rate of athlete knee joint injury image in the method of DLSK [8] is 32.6%, the maximum segmentation error rate of athlete knee joint injury image in the method of IAUEIS [9] is 13.9%, and the maximum segmentation error rate of athlete knee joint injury image in the method of PWMIS [10] is 24.6%. The error rate of proposed method is 29.6%, 20.1%, 25.8%, 7.1% and 17.8% lower than that of ISFVO [6] method, NODIS [7] method, DLSK [8] method, IAUEIS [9] method and PWMIS [10] method, respectively. In general, the error rate of proposed method is lower and the segmentation accuracy is higher.

To further verify the ability of the proposed method to segment the images of knee joint injury of athletes, the segmentation time was used as a measure to test the speed of six methods to segment the images of knee injuries of athletes with the same number of images of knee joint injury of athletes, and the test results are shown in Table 2. The analysis of Table 2 shows that the more images of knee joint injury of athletes, the longer the segmentation time. In the six methods, when the number of knee joint injury images of athletes is the same, the maximum value of segmentation time of the proposed method is only 3.96s, which is 3.29s lower than the method of ISFVO [6], 4.79s lower than the method of NODIS [7], 5.4s lower than the method of DLSK [8], 6.29s lower than the method of IAUEIS [9], and 3.67s lower than the method of PWMIS [10]. It shows that the proposed method takes less time to segment the images of knee joint injury of athletes and its segmentation ability is stronger.
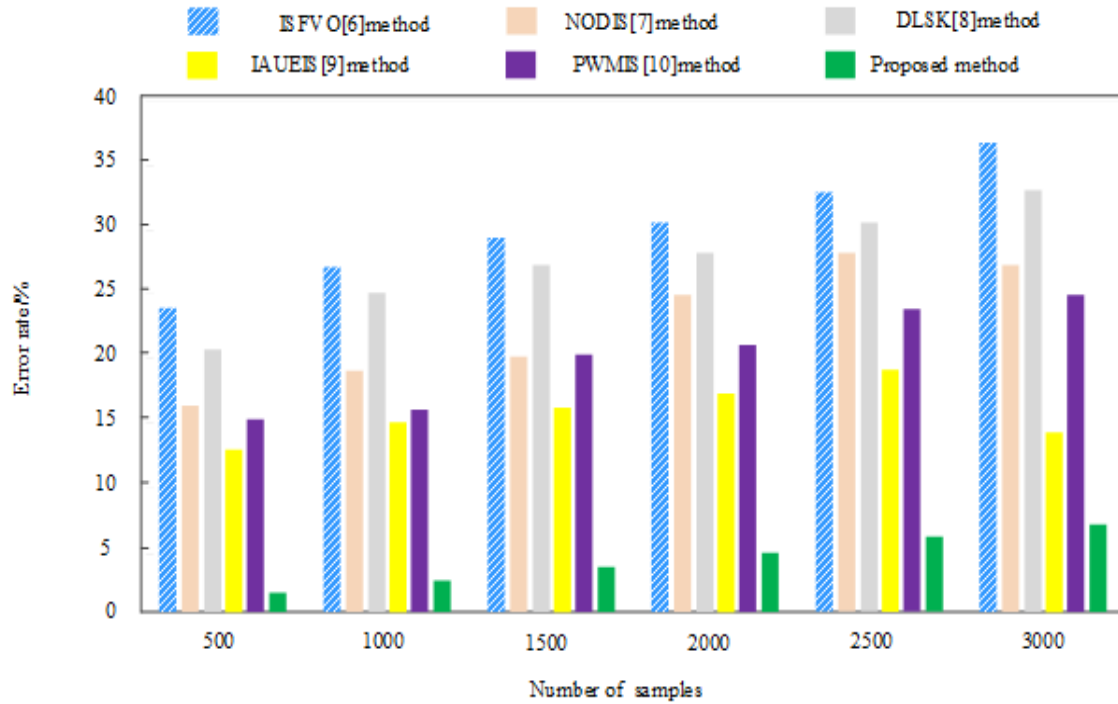
Figure 8. Comparison of segmentation error rate of knee joint injury images of athletes.

Table 2. Comparison of segmentation time of knee joint injury images of athletes

Methods	Number of samples					
	500	1000	1500	2000	2500	3000
Proposed	1.28	1.55	2.47	2.69	3.58	3.96
ISFVO[6]	1.46	1.92	3.48	5.84	6.17	7.25
NODIS[7]	1.29	2.01	3.59	4.78	5.93	8.75
DLSK[8]	1.33	1.84	3.11	5.32	6.35	9.36
IAUEIS[9]	1.82	2.16	4.41	5.91	8.69	10.25
PWMIS[10]	1.96	3.16	4.87	5.84	6.35	7.63

The comparison of the accuracy, Peak Signal to Noise Ratio (PSNR), Structural Similarity Index (SSIM), Image Quality Index (IQI), and Area Under the ROC Curve (AUC-ROC) between the proposed method and other methods is shown in Table 3.

Table 3. Comparison between this method and other methods in different evaluation indicators

Methods	Accuracy	PSNR	SSIM	IQI	AUC-ROC
Proposed	0.95	35.12	0.93	0.89	0.92
ISFVO[6]	0.92	32.14	0.89	0.85	0.87
NODIS[7]	0.89	30.89	0.87	0.82	0.84
DLSK[8]	0.91	31.78	0.88	0.84	0.86
IAUEIS[9]	0.88	29.72	0.86	0.81	0.82
PWMIS[10]	0.90	30.56	0.87	0.83	0.85

According to the data in Table 3. The proposed method achieved the highest accuracy of 0.95 among all methods, indicating its superior ability to accurately predict the injury regions in knee joint images. The ISFVO[6]method and DLSK[8] method also showcased high accuracies, scoring 0.92 and 0.91, respectively. Meanwhile, NODIS[7] method, IAUEIS[9] method, and PWMIS[10] method achieved slightly lower accuracies ranging from 0.89 to 0.90.The proposed method displayed the highest PSNR value of 35.12, indicating its better preservation of image details. Among the other methods, PSNR values ranged from 32.14 to

30.56.With an SSIM value of 0.93, the proposed method outperformed all other methods in terms of maintaining structural similarity in the segmented images. Other methods yielded SSIM values from 0.89 to 0.87. The proposed method achieved the highest IQI value of 0.89, indicating its ability to produce segmentations that possess good spatial consistency and brightness contrast. Other methods yielded IQI values ranging from 0.85 to 0.83.The proposed method attained an AUC-ROC value of 0.92, the highest among all methods. AUC-ROC measures the classification performance and balance between true positive and false positive rates at various thresholds. Thus, the proposed method demonstrates superior classification performance in knee joint injury image segmentation. In conclusion, based on the provided data table, the proposed method outperforms other methods in terms of accuracy, PSNR, SSIM, IQI, and AUC-ROC. It demonstrates superior classification performance, better preservation of image details, and improved image quality in knee joint injury image segmentation tasks. This highlights its research value in providing accurate and reliable information for medical professionals regarding injury location and extent.

CONCLUSION

In this paper, a method of image segmentation of athlete knee joint injury using the transformer model by the MIoT is proposed. The MIoT is used as a way to obtain the image of athlete knee joint injury, and the transformer model is used to realize the image segmentation of athlete knee joint injury by the MIoT. The results show that the minimum error rate for extracting the mask region of the athlete knee joint injury image using this method is 4%, and the average time of the mask region extraction is 1.57s; The knee joint injury image segmented by proposed method is the most detailed and comprehensive, and the smaller and less obvious knee joint injury areas are effectively segmented; The maximum segmentation error rate of knee joint injury image of proposed method is only 6.8%, and the maximum segmentation time value is only 3.96s.

The proposed method has significant advantages in comparison with Accuracy, PSNR, SSIM, IQI, AUC-ROC. The results show that the proposed method has a strong ability to mobilize the knee joint injury image mask region division and segmentation, and the application effect is better. However, the experimental sample images used in the experiment process are not massive enough, which makes the scientific and reliability of the results slightly decreased. Therefore, more experimental data need to be used for future research to test the application effect of the proposed method.

Conflicts of Interest: The authors declare that there is no conflict of interest with any financial organizations regarding the material reported in this manuscript.

REFERENCES

1. Bicer M, Phillips A, Modenese L. Altering the strength of the muscles crossing the lower limb joints only affects knee joint reaction forces. Gait Posture. 2022; 95(1):210-6. doi: 10.1016/j.gaitpost.2022.03.020
2. Wang X, Lin J, Li Z, Ma Y, Zhang X, He Q, et al. Identification of an ultrathin osteochondral interface tissue with specific nanostructure at the human knee joint. Nano Lett. 2022; 22(6):2309-19. doi.org/10.1021/acs.nanolett.1c04649
3. Huang C, Xu Z, Shen Z, Luo T, Li T, Nissman D, et al. DADP: Dynamic abnormality detection and progression for longitudinal knee magnetic resonance images from the Osteoarthritis Initiative. Med Image Anal. 2022; 77(14): 102343. doi:10.1016/j.media.2021.102343
4. Mehrtash A, Ghafoorian M, Pernelle G, Ziaei A, Heslinga F, Tuncali K, et al. Automatic needle segmentation and localization in mri with 3-d convolutional neural networks: application to mri-targeted prostate biopsy. Ieee T med Imaging. 2019; 38(4): 1026-36. doi: 10.1109/TMI.2018.2876796.
5. Naseem R, Khan Z, Satpute N, Beghdadi A, Cheikh F, Olivares, J. Cross-modality guided contrast enhancement for improved liver tumor image segmentation. Ieee Access. 2021; 9: 118154-67. doi: 10.1109/ACCESS.2021.3107473. doi: 10.1109/ACCESS.2021.3084905.
6. Zhang Y, Tian Y. A new image segmentation method based on fractional-varying-order differential. J Beijing Inst Technol (English Edition). 2021; 30(3): 254-264.doi:10.15918/j.jbit1004-0579.2021.028
7. Basar S, Ali M, Ochoa-Ruiz G, Waheed A, Rodriguez-Hernandez G, Zareei M, A Novel Defocused Image Segmentation Method Based on PCNN and LBP. Ieee Access. 2021;9: 87219-87240. doi: 10.1109/ACCESS.2021.3084905.

8. Panfilov E, Tiulpin A, Nieminen M, Saarakkala S, Casula V. Deep learning-based segmentation of knee MRI for fully automatic sub-regional morphological assessment of cartilage tissues: data from the osteoarthritis initiative. J Orthop Res. 2022;40(5): 1113-24. doi:10.1002/jor.25150

9. Shi Y, Zhang J, Ling T, Lu Jiwen, Zheng Y, Yu Q, et al. Inconsistency-aware uncertainty estimation for semi-supervised medical image segmentation. Ieee T Med Imaging. 2022; 41(3):608-20. doi: 10.1109/TMI.2021.3117888.

10. Wen Y, Chen L, Deng Y, Zhang Z, Zhou C. Pixel-wise triplet learning for enhancing boundary discrimination in medical image segmentation. Know-based Syst. 2022; 243(11): 108424.doi:10.1016/j.knosys.2022.108424

11. Liu X, Zhang Y, Jing H, Wang L, Zhao S. Ore image segmentation method using U-Net and Res_Unet convolutional networks. Rsc Adv.2020; 10(16): 9396-406. doi: 10.1039/C9RA05877J

12. Jyotheeswari P, Jeyanthi N . Hybrid encryption model for managing the data security in medical internet of things. Int J Internet Proto. 2020;13(1): 25-31. doi:10.1504/IJIPT.2020.105049

13. Shen M, Deng Y, Zhu L, Du X, Guizani N. Privacy-preserving image retrieval for medical iot systems: a blockchain-based approach. Ieee Network. 2019; 33(5): 27-33. doi: 10.1109/MNET.001.1800503.

14. Chai R. Otsu's image segmentation algorithm with memory-Based fruit fly 0ptimization algorithm. Complexity. 2021: 1-11. doi.org/10.1155/2021/5564690

15. Chernyshov A, Klimov V, Balandina A, Shchukin B. The application of transformer model architecture for the dependency parsing task. Procedia Computer Science. 2021; 190(2):142-5. doi:10.1016/j.procs.2021.06.018

16. Shen T, Xu H, Medical image segmentation based on transformer and HarDNet structures. IEEE Access.2023; 11: 16621-16630. doi: 10.1109/ACCESS.2023.3244197.

17. Li J , Ou X, Shen N , Sun J, Ding J, Zhang J, et al. Study on strategy of CT image sequence segmentation for liver and tumor based on U-Net and Bi-ConvLSTM. Expert Syst Appl. 2021; 180(6): 115008. doi:10.1016/j.eswa.2021.115008

18. Gao W, Li X, Wang Y, Cai Y. Medical image segmentation algorithm for three-dimensional multimodal using deep reinforcement learning and big data analytics. Front Public Health.2022; 10:879639. doi:10.3389/fpubh.2022.879639

19. Bates R, Irving B, Markelc B, Kaeppler J, Brown G, Muschel R, et al. Segmentation of vasculature from fluorescently labeled endothelial cells in multi-photon microscopy images. Ieee T Med Imaging. 2019; 38(1): 1-10. doi: 10.1109/TMI.2017.2725639.

20. Huang Y, Dou Q, Wang Z, Liu L, Jin Y, Li C, et al. 3-D ROI-aware U-NET for accurate and efficient colorectal tumor segmentation. Ieee T Cybernetics. 2021; 51(11):5397-5408. doi: 10.1109/TCYB.2020.2980145.

21. Trajanovski S, Shan C, Weijtmans P, Koning S, Ruers T. Tongue tumor detection in hyperspectral images using deep learning semantic segmentation. Ieee T Bio-med Eng. 2021; 68(4):1330-1340. doi: 10.1109/TBME.2020.3026683.

22. Zhang D, Huang G, Zhang Q, Han J, Yu Y. Cross-modality deep feature learning for brain tumor segmentation. Pattern Recogn. 2021; 110(11):107562. doi: 10.1016/j.patcog.2020.107562