IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 20 December 2023, accepted 2 January 2024, date of publication 4 January 2024,
date of current version 11 January 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3350036

RESEARCH ARTICLE

The Attention Mechanism Performance Analysis for Football Players Using the Internet of Things and Deep Learning

CHUAN MOU
Institute of Physical Education, Sichuan University, Chengdu, Sichuan 610065, China
e-mail: muchuan2012@gmail.com

This work was supported by the Fundamental Research Funds for the Central Universities with Sichuan University.

ABSTRACT This work proposes a novel Class Aware Network (CANet) for analyzing football player performance by decoding their body movements. Firstly, the role of the Internet of Things in football sports analysis and the advantages of deep learning techniques are introduced. Secondly, pyramid pooling modules and attention mechanisms are introduced. Moreover, the Group-split-bottleneck (GS-bt) module is employed, and the CANet is designed to extract and utilize multi-scale feature information and enhance the network's ability to perceive details. Finally, the effectiveness of the proposed model is validated through comparisons with other models. The results show that in image classification experiments, the mean accuracy of the GS-bt module is at least 2.79% higher than that of other models. In human body parsing experiments, results from two different datasets demonstrate that the CANet model achieves the highest mean Intersection over Union, improving by at least 6.02% compared to other models. These findings indicate that the proposed CANet model performs better in image classification and human body parsing tasks, presenting higher accuracy and generalization capabilities. This work provides new methods and technologies for analyzing football player performance, potentially promoting sports development and application in athletics.

INDEX TERMS Internet of Things, deep learning, attention mechanism, football player performance analysis, human body parsing.

I. INTRODUCTION

A. RESEARCH BACKGROUND AND MOTIVATIONS

Football, as a globally popular sport, has always received widespread attention and research. With the advancement and application of technology, the Internet of Things (IoT) and Deep Learning (DL), two cutting-edge technologies, are increasingly demonstrating powerful application potential in sports.

Traditional football sports analysis heavily relies on manual observation and manual recording. It is time-consuming, labor-intensive, and prone to subjective influences, limiting a comprehensive evaluation of athletes' performance [1], [2], [3]. However, with the progress of the IoT, motion sensors can now capture real-time data on athletes' movements and physiological parameters, providing more data support for sports analysis [4], [5], [6]. Nevertheless, efficiently and accurately processing such vast amounts of data remains challenging. As a popular technology in artificial intelligence (AI), DL has achieved significant breakthroughs in image recognition, speech processing, and other domains. DL can efficiently handle large-scale data in sports analysis, identify and extract key movements and sports details, and better understand athletes' performance [7], [8], [9], [10]. However, in complex sports scenarios, DL models may face issues like interference from data and information overload, leading to a decline in performance.

To address these issues, introducing attention mechanisms is a feasible approach. Attention mechanisms can help the model focus on key information and reduce interference from irrelevant data, thereby improving the model's accuracy and robustness [11], [12]. In sports analysis, attention

The associate editor coordinating the review of this manuscript and approving it for publication was Giacomo Fiumara.

mechanisms have the potential to optimize the DL model further, making it more adaptable to complex sports scenarios and accurately capturing athletes' key performances.

In this context, the research motivation of this work highlights three key aspects. Firstly, IoT technology offers a new dimension to real-time data acquisition. By obtaining movement data and physiological parameters of athletes through sensors, unprecedented detailed information can be obtained, which provides a richer data basis for sports analysis. This raises important questions about efficiently processing, analyzing, and applying these massive amounts of data. Secondly, DL, as a cutting-edge technology of AI, has shown amazing capabilities in image and voice. Introducing DL into football analysis can better identify and extract key action and detail information, and provide a more in-depth and accurate perspective for analyzing athletes' performance. Besides, it is also necessary to face the problems that DL may encounter in complex motion scenarios, such as data interference and information overload, which require targeted methods and technologies to deal with. Finally, the attention mechanism is introduced as a solution to optimize the effect of the DL model in motion analysis. By allowing the model to focus on key information and reduce redundancy and interference, it is expected to improve the accuracy and robustness of the model, making it better adapted to complex motion scenarios.

In general, the research motivation of this work is not only the pursuit of improvement of traditional methods but also an active attempt to apply emerging technologies in sports. Through the organic combination of IoT, DL, and attention mechanism, it is hoped to offer a new and more accurate solution for the sports performance analysis of football players. Meanwhile, it provides a useful reference for combining IoT and DL in other fields.

B. RESEARCH OBJECTIVES

The main research objective is to explore the application of IoT and DL in football player performance analysis, combined with attention mechanisms, to improve the efficiency and accuracy of sports analysis. The work utilizes a multi-label classification-guided human body parsing algorithm, leveraging IoT sensors to gather football players' motion data and physiological parameters. The DL model is then integrated with attention mechanisms to evaluate athletes' performance comprehensively. The pyramid pooling module (PPM) and attention mechanisms are introduced to achieve this goal, proposing a novel Class Aware Network (CANet) with skip connections in the algorithm. During the convolution process, different levels of detailed features are guided into the decoding process, enabling the analysis of football players' performance during their movements.

This work introduces a new CANet, which combines IoT technology and DL attention mechanism, as well as a Group-split-bottleneck (GS-bt) module for analyzing the sports performance of football players. This new network structure is still a novel attempt in the football player analysis field. By introducing the PPM and attention mechanism, CANet can effectively extract and utilize multi-scale feature information, thus enhancing the network's perception of details. This multi-scale feature utilization method provides more comprehensive data support for athlete performance analysis. By introducing a branch of multi-label classification for training, CANet can better handle human body parsing tasks, especially in parsing small targets. Based on the above innovation points, this work offers a new method in football players' sports performance analysis through an innovative network structure and comprehensive application of various technical means, to improve the accuracy and generalization ability of evaluating athletes' sports performance.

This work aims to provide a scientific basis for football players' training and competition strategies, thereby advancing the sports science field. Moreover, it offers valuable exploration for utilizing IoT and DL to address complex issues in other domains.

II. LITERATURE REVIEW

The application of IoT in football sports analysis has become one of the current research hotspots. Wang utilized computer vision and IoT to design an intelligent image-processing method for recognizing and processing football images. The research results demonstrated the promising prospects of this intelligent recognition technology in sports and fitness [13]. Wang and Liu aimed to enhance the sports industry's operational efficiency and athletes' safety using technologies such as big data, computer vision, and IoT. They collected image data through smart cameras and radio frequency technology, and utilized DL and motion information for recognizing and extracting attacking actions [14]. Xue and Liu guided the training concept update of football players by analyzing the combined effects of embedded sensor networks and various football performance indicators. The results demonstrated that the experimental group showed better muscle and neural activation training outcomes, presenting clear advantages through embedded sensor network technology [15]. The application of IoT significantly improved the efficiency and accuracy of collecting football player performance data, offering more information support for sports analysis.

With the rapid development of DL, its advantages in football sports analysis are increasingly evident. Yin et al. proposed a DL-assisted motion capture system to reduce animation costs and improve motion capture data for football matches through bidirectional motion analysis. Simulation analysis showed that this framework was highly reliable [16]. Zanganeh et al. focused on computer vision analysis of football videos and introduced a public dataset called IAUFD to address the issue of insufficient data for DL methods. They evaluated the dataset effectiveness using deep neural networks (DNNs) like VggNet-13 and ResNet-18 as baselines [17]. Liu et al. presented a DL-based American football player tracking system using a two-stage network design. It first solved player detection in crowded scenes through an object detection network and then identified players' jersey numbers using an auxiliary convolutional neural network (CNN), synchronizing the results with the game clock [18]. These models achieved significant results in motion recognition and motion trajectory prediction tasks.

In addition, some related research has also attracted people's attention to the IoT and AI environment and the problem of big data processing. Kumari et al. analyzed how to handle secure streaming data from different devices. They highlighted the major threats and risks in big data processing and explored the security approach to handling big data in the IoT environment, as well as the architectural details and security approach required at various stages of the big data processing lifecycle [19]. These ideas can be used for reference for data processing and security problems in this work. In terms of building a more secure and privacy-protected Industrial Internet of Things (IIoT) system, Kumari et al. introduced a blockchain-based decentralized IIoT model. The interaction between nodes was realized through a secure P2P network to solve the security and privacy problems of the existing IIoT system with a centralized architecture [20]. The idea of this model can inspire people to explore more reliable data interaction and protection methods. Kumari and Tanwar proposed a big data analysis scheme called ρ Reval based on the AI energy price prediction model, using the bidirectional long short-term memory model. Through comparison with existing methods, the research proved the superiority of the ρ Reveal scheme in forecasting accuracy [21]. This AI-based approach to big data analysis offers inspiration for this work, especially in data processing when predicting and analyzing athletes' sports performance. These findings provide context for this work concerning IoT, AI, and big data processing. At the same time, it also offers valuable experience for exploring better and safer methods and technologies.

The foregoing survey of research underscores the critical roles that the IoT and DL assume in football player performance analysis. These technologies furnish new avenues for achieving comprehensive, efficient, and accurate motion analysis. However, current research faces challenges like data processing complexity and model robustness. In light of these considerations, this work will investigate elevated methodologies for analyzing motion performance by synergistically amalgamating IoT, DL, and attention mechanisms, affording more valuable information and suggestions for football players' training and competition.

III. RESEARCH METHODOLOGY

A. THE APPLICATION OF IOT TECHNOLOGY IN FOOTBALL SPORTS ANALYSIS

IoT connects various physical devices, sensors, and other objects through the internet for intelligent interactions and data transmission. These IoT devices can perceive environmental information and accomplish automation, intelligence, and remote control through data transmission and communication over the Internet [22], [23], [24]. As an advanced technology for connecting and transmitting data, it plays an increasingly significant role in football sports analysis. IoT sensors allow real-time acquisition of football players' motion data and physiological parameters, thereby providing more comprehensive and accurate information support for performance analysis [25], [26].

IoT relies on various types of sensors embedded in football shoes, players' bodies, or other equipment to real-time sense and collect data on players' activities [27], [28]. Common types of sensors include accelerometer, gyroscope, heart rate, and others, as shown in Figure 1.

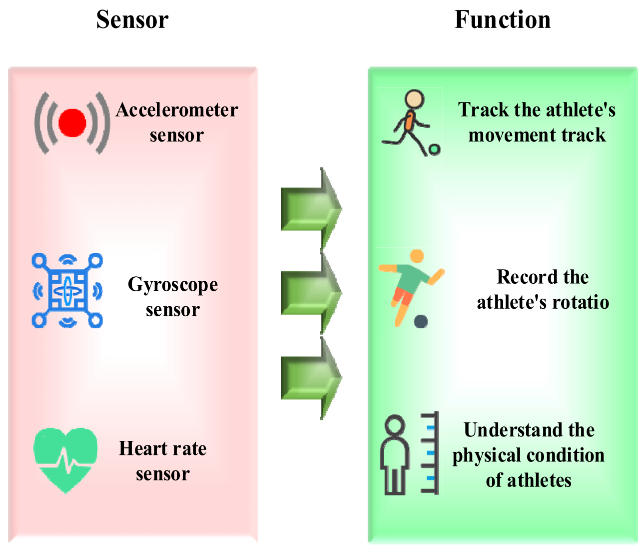

FIGURE 1. Common IoT sensors in football sports and their functions.

Figure 1 displays that in the context of football sports performance analysis, accelerometer sensors are employed to measure the acceleration and velocity of football players, thereby tracking their movement trajectories and action paths [29], [30]. Gyroscope sensors are utilized to measure the angular velocity of players, enabling the recording of their rotational and turning actions, thus refining the motion data of players [31], [32], [33]. Heart rate sensors monitor the changes in players' heart rates, providing coaches and athletes with timely physiological health references to understand their physical condition and fatigue levels [34]. By applying these sensors, IoT can collect real-time and high-precision motion data and physiological parameters of football players, thereby providing a robust data foundation for performance analysis.

IoT can offer extensive data support for applying DL models and attention mechanisms by enabling real-time data collection and efficient data processing. Further integrating DL is expected to achieve a comprehensive assessment and deeper understanding of football players' performance, providing more scientifically grounded guidance and a decision-making basis for training and competition strategies.

B. THE ADVANTAGES OF DL IN FOOTBALL SPORTS ANALYSIS

DL is a machine learning technique based on neural networks that has garnered significant accomplishments across diverse domains, including image recognition and natural language processing, over the past years [35], [36]. DL has also demonstrated unique advantages in football sports analysis, opening up new possibilities for analyzing athletes' performance. These advantages encompass motion recognition, pose estimation, motion trajectory prediction, data fusion, and transfer learning, as elucidated in Figure 2.

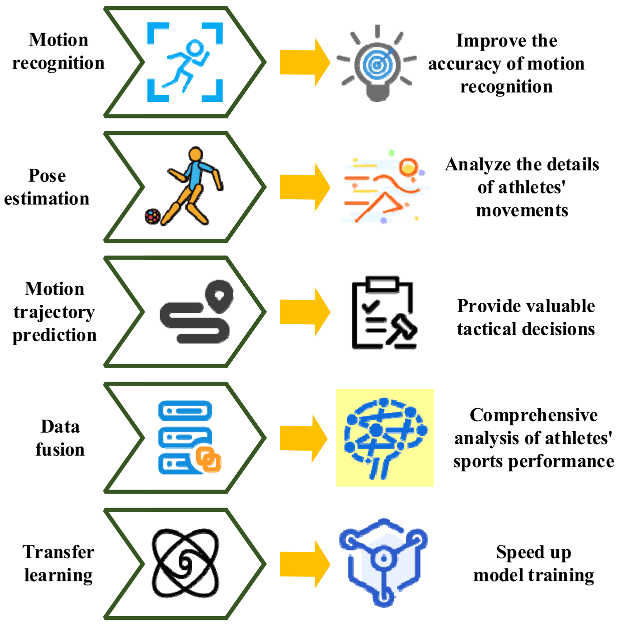

FIGURE 2. The advantages of DL in football sports analysis.

Figure 2 illustrates that in motion recognition, DL can automatically learn feature representations, enhancing motion recognition's accuracy and generalization ability. DL analyzes athletes' motion details and posture changes in pose estimation and motion trajectory prediction, providing valuable tactical decisions and match planning information. Transfer learning can expedite model training speed and improve prediction accuracy, while data fusion integrates data from various sources, enabling a comprehensive analysis of athletes' performance [37], [38], [39].

The advantages of DL in football sports analysis are evident. The application of DL models can achieve efficient recognition of athletes' movements, accurate pose estimation, and precise motion trajectory prediction [40], [41]. The dynamic evolution of DL brings forth both opportunities and challenges in football sports performance analysis. DL attention mechanisms are utilized to explore multi-label classification-guided human body parsing algorithms, enabling a more detailed categorization of human body movements in football sports.

C. MULTI-LABEL CLASSIFICATION-GUIDED HUMAN BODY PARSING ALGORITHM

In contrast to semantic segmentation, which primarily concerns broader categories, human body parsing necessitates an intricate classification of individual elements within images or videos. It requires annotating various body parts, such as the head, torso, or relevant attire like shoes and pants, and performing pixel-level annotations on semantically consistent regions in the image [42]. A novel CANet is proposed to address these intricate details. The algorithm integrates skip connections to guide varied hierarchical detailed features into the decoding process during convolution. Furthermore, PPM is introduced to extract and fuse information from multiple scales. The attention mechanism is employed to tackle the feature selection issue in the feature map channels. It enables the network to selectively enhance specific channel features' response to the final decision throughout the end-to-end training process. The algorithm adeptly manages human body parsing tasks through these strategies by extracting and effectively utilizing intricate information. Figure 3 depicts the architecture of the proposed CANet.

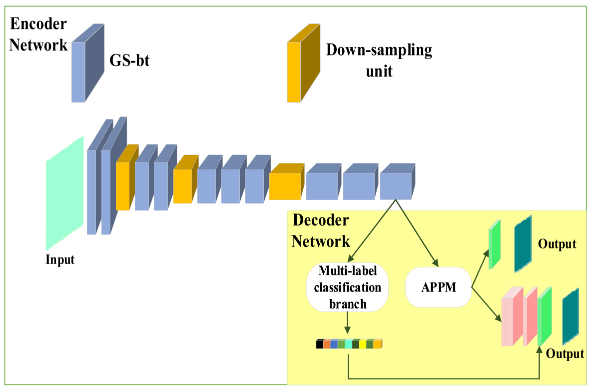

FIGURE 3. The overall architecture of CANet.

In Figure 3, CANet is a network with an asymmetric encoder-decoder structure. The encoder network adopts GS-bt modules, which modify the original Residual Network (ResNet) backbone, remove fully connected layers, and output a feature map of 1/8 the size of the input image. Simultaneously, a multi-label classification branch is introduced. The decoder network comprises the Attention Pyramid Pooling Module (APPM) and multiple convolutional layers. The entire network is trained end-to-end, incorporating the multi-label classification task to extract information about the presence or absence of categories in the image, affording prior knowledge for subsequent human body parsing tasks.

Figure 4 portrays the GS-bt module.

In Figure 4, from left to right, the first module is the bottleneck module, the second is the basic unit of ShuffleNet, and the third is the GS-bt module [43], [44]. The GS-bt module replaces the basic feature computation module in the baseline with two simple grouped convolutions and channel shuffling.

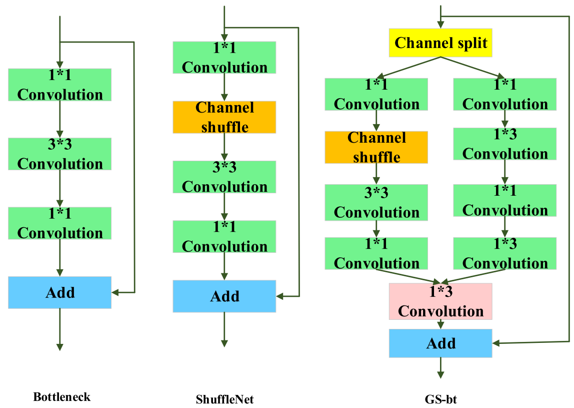

FIGURE 4. The comparison between the GS-bt module and other modules.

It aims to push the convolution to its limits, especially in balancing performance and efficiency.

Next, improvements have been made to the CNN architecture Pyramid Scene Parsing Network (PSPNet) to address pixel-level semantic segmentation in images [45], [46]. Figure 5 plots the modified APPM based on attention mechanisms.

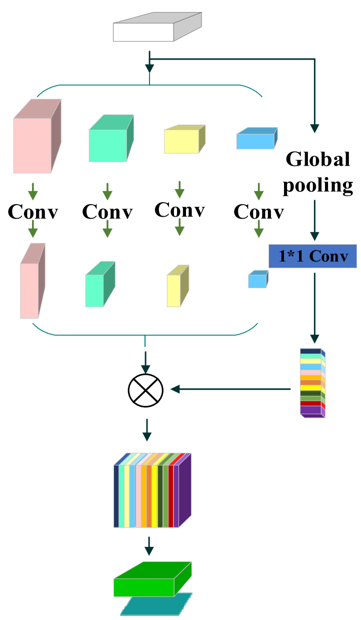

FIGURE 5. APPM structure.

In Figure 5, the attention mechanism called APPM is employed to optimize the analysis of football players' body movements. APPM is formed by incorporating channel attention operations into the original PPM, showcasing a functionality akin to feature selection. The role of channel attention operations is to enable the model, to learn relationships between different feature maps during training, facilitating the learning and utilization of global information. Specifically, channel attention operations enhance the network's understanding of the importance of various features by considering interactions between channels. Throughout the learning process, these operations weight and emphasize features that are more meaningful for the task of analyzing the body movements of football players. The channel attention operation in APPM is represented by Eq (1).

$$Y = \sigma(FC\ (AvgPool\ (X))) \cdot X \tag{1}$$

$X \in R^{C\times H\times W}$ is the input feature map, C is the number of channels, H is the height, and W is the width. Y represents the output of channel attention. $AvgPool$ denotes global average pooling applied to the input, resulting in a feature map of size $C \times 1 \times 1$. FC represents a fully connected layer that maps the feature map obtained from global average pooling to a new feature map of size $C \times 1 \times 1$. σ denotes the sigmoid activation function. The effect of this operation is to learn weights for each channel, thereby achieving selective attention to different channels.

During the training process, the channel attention operation automatically learns relationships between different feature maps through the backpropagation algorithm. By minimizing the loss function, the network adjusts the weights in the channel attention operation to enhance its performance in body movement analysis tasks. This adaptive learning process enables the network to dynamically adjust channel attention based on task requirements, better adapting to complex motion scenarios. The operation allows the network to model relationships between different feature maps, facilitating the learning and utilization of global information. It selectively emphasizes information-rich features and suppresses less important ones. In the research process, APPM is integrated into the CANet, replacing the original PPM as a module. This means that channel attention operations are applied in every branch of the network, comprehensively enhancing the network's perception of features at different scales. By introducing this attention mechanism, the analysis performance of the network for football player body movements is improved. This attention mechanism is essentially a "neural network attention mechanism," with the primary goal of establishing associations between feature maps at different levels. This enables the network to better understand and utilize global information, enhancing the effectiveness of body movement analysis. The introduction of this attention mechanism allows the network to adaptively focus on and emphasize specific features during training, thereby improving its analytical performance in body movement. Overall, the role of APPM in the entire network structure is to enhance the network's perception and utilization of global information, contributing to the accuracy and robustness of football player body movement analysis.

D. INTEGRATED APPLICATION OF IOT, DL, AND CANET MODELS IN REAL FOOTBALL

IoT, DL, and CANet models are vital in improving athletes' performance analysis and training effectiveness in real football applications.

First, IoT technology obtains athletes' movement data and physiological parameters in real-time through motion sensors, which provide rich information support for sports analysis. For example, the athlete's running speed, stride frequency, heart rate, and other data can be collected in real-time through IoT sensors, allowing in-depth analysis of the athlete's body state and movement characteristics. Real-time data acquisition reduces reliance on human observation and manual recording and provides coaches and athletes with more accurate and comprehensive performance information, enabling them to adjust training and competition strategies promptly.

Second, DL technology achieves finer motion analysis by processing large-scale data in real football applications. The DL model can identify and extract key movements and details, such as passing, shooting, defensive movements, etc., to comprehensively assess an athlete's performance. This technology can help coaches better understand athletes' strengths and room for improvement, and develop personalized training plans. Furthermore, DL can predict athletes' possible movement direction and strategy, providing coaches and teams with more basis for tactical decisions.

Most importantly, the CANet model, as a novel network combining IoT and DL, achieves a higher level of analysis accuracy for the human movement analysis of football players. By introducing PPM and attention mechanism, the CANet model can effectively extract and utilize multi-scale feature information, and enhance the network's ability to perceive details. This allows CANet to more accurately parse the athlete's movement, involving the movement of various parts and changes in posture. Therefore, in real football applications, the CANet model can offer coaches and analysts a more accurate and comprehensive analysis of athletes' movements, and provide a scientific basis for training and competition decisions.

To sum up, the IoT, DL, and CANet models are combined in real football applications and strongly support athletes' performance analysis and training. These technologies have revolutionized scientific training and competitive decision-making in football through real-time data acquisition and in-depth and precise motion analysis.

IV. EXPERIMENTAL DESIGN AND PERFORMANCE EVALUATION

A. EXPERIMENTAL MATERIALS

The research experiments are divided into two parts. The first part is image classification experiments, used to validate the performance of the proposed GS-bt module. Qualitative experiments are conducted to analyze and compare the strengths and weaknesses of the proposed algorithm against the baseline algorithm. The baseline algorithm used here is the ResNet-50 model, and the comparison models cover Xception [47] and DenseNet [48]. The second part is human body parsing experiments conducted to demonstrate the effectiveness of the proposed CANet model. The comparison models include local-global long short-term memory (LG-LSTM) [49], Whole-Body Human Pose Estimation (WSHP) [50], PSPNet [51], Pose Guided Person Image Generation (PGN) [52], DeepLab V2 [53], and Attention [54].

The Canadian Institute for Advanced Research 100 (CIFAR-100) dataset is used in image classification experiments. It consists of 100 classes with approximately 600 images per class. Each class encompasses 500 training images and 100 validation images. The datasets used in the human body parsing experiments are PASCAL-Person-Part and Look Into Person (LIP). The PASCAL-Person-Part dataset contains 3533 images, with 1716 and 1817 images in the training and test sets. It categorizes body parts and background into 7 classes: background, head, torso, upper arms, lower arms, upper legs, and lower legs. The LIP dataset comprises human body images from various scenes, covering diverse body poses, clothing, and actions. The training set includes 30,462 images, while the validation set contains 10,000 images.

Additionally, real-time datasets are adopted in the training of football players in X City to verify the performance of the CANet model in practical application scenarios. In actual training, various sensors and cameras are employed to collect images and relevant data of athletes in real-time, encompassing movement trajectory, posture information, speed, acceleration, heart rate, etc. The images and data of the athletes are captured through sensors and cameras and transmitted to the data center in real-time. The acquired images are preprocessed, including image denoising, size adjustment, etc., to ensure the data quality and consistency of the input model. The CANet model is utilized to train the acquired real-time images. The model is trained as a multi-label classification model to predict different athletes' movements, postures, and parts. The trained model is applied to the validation set of real-time datasets to analyze and predict the athletes' sports performance.

B. EXPERIMENTAL ENVIRONMENT AND PARAMETERS SETTING

This experiment uses the PyTorchDL framework, employing the Python language, with the GTX 1080Ti graphics processor on the computer. Table 1 exhibits the experimental parameter settings.

The evaluation criterion for the image classification experiments is mean accuracy. In the human body parsing experiments, mean Intersection over Union (mIoU) across all categories emerges as the evaluation metric. IoU represents the ratio of the intersection to the union between the predicted result set of a certain category and the corresponding category set in the dataset [55], [56]. The calculation method for mIoU is as follows [57]:

$$mIoU = \frac{1}{n+1} \sum_{i=0}^{n} \frac{M_{ii}}{\sum_{j=0}^{n} M_{ij} + \sum_{j=0}^{n} M_{ji} - M_{ii}} \quad (2)$$

TABLE 1. Experimental parameter settings.

		Image classification experiment	Human body parsing experiments
Network structure	Number of modules for basic convolution modules	GS-bt	GS-bt
	Number of modules for basic convolution modules	16	12
	Number of layers	50	/
Training parameters	Batch size	128	4
	Learning rate	0.1	0.0005 for the Person-Part dataset and 0.001 for the LIP dataset
	Number of training rounds	200	300 for the PASCAL-Person-Part dataset and 180 for the LIP dataset
	Weight decay	0.0005	0.0001

n represents the number of categories in the dataset, and since counting starts from 0, $n+1$ indicates the total number of categories. M_{ii} means the number of values with a true value of i and predicted as i. M_{ij} refers to false positives, and M_{ji} stands for false negatives.

C. PERFORMANCE EVALUATION

1) ANALYSIS OF IMAGE CLASSIFICATION EXPERIMENT RESULTS

The experimental results of the CIFAR-100 dataset are suggested in Figure 6.

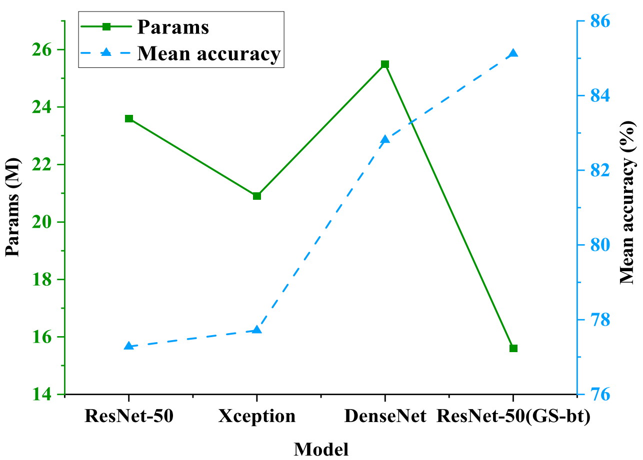

FIGURE 6. Image classification results.

Figure 6 presents that compared to the baseline ResNet-50 model, the mean accuracy of the GS-bt module has increased by 10.14%. Moreover, compared to the Xception and DenseNet models, the ResNet-50 model using the GS-bt module has enhanced its mean accuracy by 9.54% and 2.79%, respectively. It indicates that the proposed GS-bt module performs better in image classification.

2) ANALYSIS OF THE RESULTS FROM HUMAN BODY PARSING EXPERIMENTS

Figure 7 denotes the classification results of the LG-LSTM, WSHP, PSPNet, PGN, and Attention models on the PASCAL-Person-Part dataset.

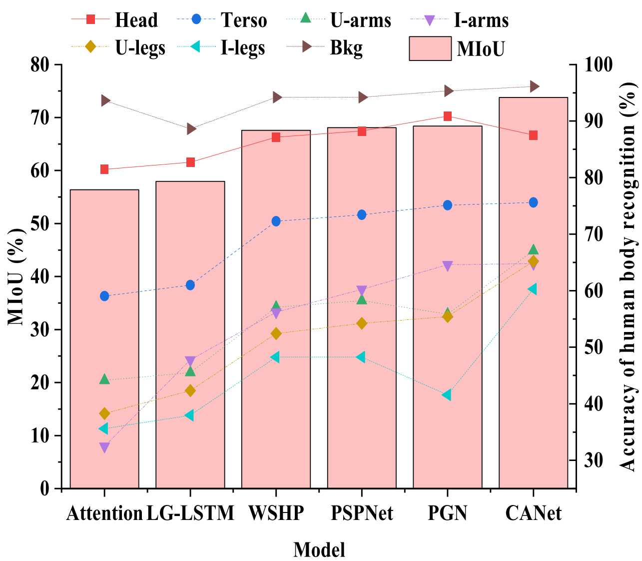

FIGURE 7. The experimental results on the PASCAL-Person-Part dataset.

Figure 7 demonstrates that the proposed CANet model achieves a mIoU of 73.78%. In comparison to the LG-LSTM, WSHP, PSPNet, PGN, and Attention models, the CANet model exhibits notable enhancements in mIoU results, with improvements of 27.29%, 9.16%, 8.34%, 7.88%, and 30.86%, respectively, in human body parsing tasks. This indicates that the CANet model performs better in human body parsing tasks. To validate the algorithm's robustness, the performance of the proposed CANet model is evaluated on the LIP dataset. The results are presented in Figure 8.

Figure 8 indicates that the mIoU of the CANet algorithm reaches 49.11% on the LIP dataset. Compared to the DeepLab V2, Attention, and PSPNet models, the mIoU of the CANet model has increased by 11.72%, 14.45%, and 6.02%, respectively. This result once again demonstrates the effectiveness of the CANet model.

The performance of the CANet model in recognizing human body parts in real training is portrayed in Figure 9.

Figure 9 suggests that on the real dataset obtained by actual training, the average accuracy of the CANet model on the head, torso, U-arm, L-arm, U-leg, and L-leg is 86.2%, 75.7%, 67.9%, 63.6%, 64.9%, and 59.8%, respectively. This is not much different from the test results on the open dataset, with a maximum difference of 1.3%. This indicates that the performance of the CANet model in practical application scenarios is close to that on idealized datasets, and it has good

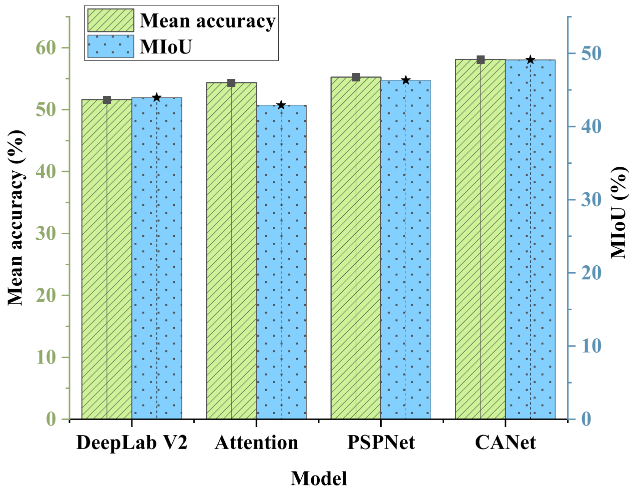

FIGURE 8. Experimental results on the LIP dataset.

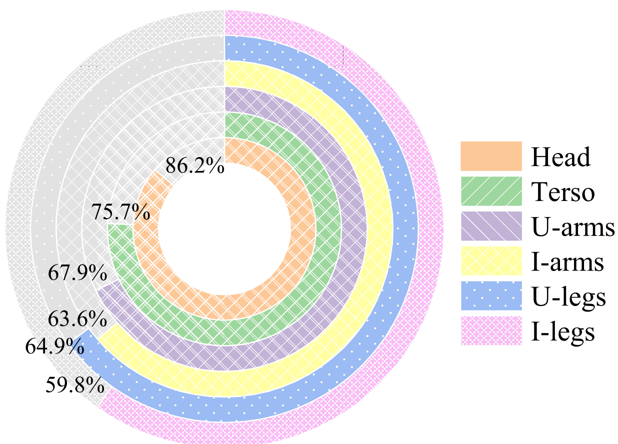

FIGURE 9. Average accuracy of human body parts recognition on real datasets.

generalization ability. The accuracy of the CANet model on different human body parts is relatively stable, and there is no large performance fluctuation, illustrating that the CANet model has certain stability and reliability for the feature learning and analysis of various human body parts.

D. DISCUSSION

Moreover, research articles published in DNNs in recent years have demonstrated that neural network models incorporating attention mechanisms demonstrate better performance. For instance, Huang et al. devised a lightweight tongue image segmentation network for automated tongue diagnosis and introduced attention mechanisms to enhance crucial features and suppress irrelevant ones. The results showcased that this network exhibited competitive performance on two tongue image datasets and accurately extracted tongue bodies, fully meeting practical application requirements [58]. To address the problem of tobacco leaf maturity recognition, Zhang et al. proposed a lightweight tobacco-leaf maturity recognition method founded on the MobileNetV1 model, feature pyramid network (FPN), and attention mechanism. By incorporating the FPN structure and attention mechanism, this model effectively combined high-level semantic information and low-level positional information, successfully identifying tobacco leaf boundaries [59]. This proposed model significantly contributes to tobacco leaf maturity recognition in real-world applications. Ma et al. introduced an improved DL model, CTR_YOLOv5n, for detecting common maize diseases such as corn leaf blight, gray leaf spot, and rust in mobile applications. This model combined coordinate attention mechanisms and the Swin Transformer detection head. It enhanced the accuracy and global information acquisition capability by introducing coordinate attention modules into the lightweight model YOLOv5n. Experimental results revealed that CTR_YOLOv5n achieved an average recognition accuracy of 95.2%, representing a 2.8 percentage point improvement over the original model [60]. In summary, an increasing number of studies have shown that incorporating attention mechanism modules into other DNNs can achieve superior outcomes, thus enhancing the feasibility of practical applications.

V. CONCLUSION

A. RESEARCH CONTRIBUTION

To investigate the sports performance of football players, this work takes their human posture during exercise as the research object, analyzes the role of IoT and DL in football performance analysis, and proposes a CANet model. The model's effectiveness is verified through comparative experiments, and the following conclusions are obtained:

1) In the image classification experiments, the mean accuracy of the GS-bt module exhibits a notable upswing of 10.14% compared to the baseline ResNet-50 model. Furthermore, when compared against the Xception and DenseNet models, the ResNet-50 model using the GS-bt module improves its average precision by at least 2.79%, demonstrating the advantages of the GS-bt module in image classification.
2) In the human body parsing task on the PASCAL-Person-Part dataset, the proposed CANet model achieves a mIoU of 73.78%. The CANet model's performance markedly outshines the LG-LSTM, WSHP, PSPNet, PGN, and Attention models, yielding an enhanced mIoU by at least 8.34%. It indicates that the CANet model performs better in human body parsing tasks.
3) Similarly, the CANet algorithm's application to the human body parsing task on the LIP dataset yields a notable mIoU of 49.11%. Compared to the DeepLab V2, Attention, and PSPNet models, the CANet model showcases an incremental enhancement in mIoU results by a minimum of 6.02%. These outcomes further reinforce the efficacy of the CANet model.

In conclusion, the proposed GS-bt module exhibits superior performance in image classification. Meanwhile, the CANet model excels in human body parsing tasks, enabling better handling of pose estimation and body part segmentation problems. This furnishes strong support and improvement directions for football player performance analysis.

B. FUTURE WORKS AND RESEARCH LIMITATIONS

Although the proposed CANet model has achieved satisfactory results in football player performance analysis, some limitations still need to be considered. First, introducing multiple technical approaches in the CANet model has increased its computational complexity. Consequently, practical implementation may necessitate higher-performing hardware and substantial computational resources to facilitate real-time inference of the model. Then, the datasets used here may exhibit certain inherent limitations, potentially failing to encompass all different scenarios and conditions of football player performance. Future research could consider using more diverse and larger-scale datasets to validate the model's generalization ability. Further studies could also focus on optimizing the structure and design of the CANet model to explore more effective feature extraction and fusion methods and further enhance the model's performance and efficiency.

REFERENCES

[1] A. López-Valenciano, J. Raya-González, J. A. Garcia-Gómez, A. Aparicio-Sarmiento, P. S. D. Baranda, M. D. S. Croix, and F. Ayala, ''Injury profile in women's football: A systematic review and meta-analysis,'' *Sports Med.*, vol. 51, pp. 423–442, Jan. 2021.

[2] T. Gronwald, C. Klein, T. Hoenig, M. Pietzonka, H. Bloch, P. Edouard, and K. Hollander, ''Hamstring injury patterns in professional male football (soccer): A systematic video analysis of 52 cases,'' *Brit. J. Sports Med.*, vol. 56, no. 3, pp. 165–171, Feb. 2022.

[3] Z. Milanovic, N. Covic, E. W. Helge, P. Krustrup, and M. Mohr, ''Recreational football and bone health: A systematic review and meta-analysis,'' *Sports Med.*, vol. 52, no. 12, pp. 3021–3037, Jul. 2022.

[4] S. Lin, S. Hu, W. Song, M. Gu, J. Liu, J. Song, Z. Liu, Z. Li, K. Huang, Y. Wu, M. Lei, and H. Wu, ''An ultralight, flexible, and biocompatible all-fiber motion sensor for artificial intelligence wearable electronics,'' *Npj Flexible Electron.*, vol. 6, no. 1, p. 27, Apr. 2022.

[5] W. Liu, Z. Long, G. Yang, and L. Xing, ''A self-powered wearable motion sensor for monitoring volleyball skill and building big sports data,'' *Biosensors*, vol. 12, no. 2, p. 60, Jan. 2022.

[6] W. Cao, Y. Luo, Y. Dai, X. Wang, K. Wu, H. Lin, K. Rui, and J. Zhu, ''Piezoresistive pressure sensor based on a conductive 3D sponge network for motion sensing and human-machine interface,'' *ACS Appl. Mater. Interfaces*, vol. 15, no. 2, pp. 3131–3140, Jan. 2023.

[7] S. M. A. Akber, S. N. Kazmi, S. M. Mohsin, and A. Szczesna, ''Deep learning-based motion style transfer tools, techniques and future challenges,'' *Sensors*, vol. 23, no. 5, p. 2597, Feb. 2023.

[8] A. Pardos, A. Menychtas, and I. Maglogiannis, ''On unifying deep learning and edge computing for human motion analysis in exergames development,'' *Neural Comput. Appl.*, vol. 34, no. 2, pp. 951–967, Jul. 2021.

[9] E. Padovan, G. Marullo, L. Tanzi, P. Piazzolla, S. Moos, F. Porpiglia, and E. Vezzetti, ''A deep learning framework for real-time 3D model registration in robot-assisted laparoscopic surgery,'' *Int. J. Med. Robot. Comput. Assist. Surg.*, vol. 18, no. 3, p. e2387, Mar. 2022.

[10] R. Z. Weber, G. Mulders, J. Kaiser, C. Tackenberg, and R. Rust, ''Deep learning-based behavioral profiling of rodent stroke recovery,'' *BMC Biol.*, vol. 20, no. 1, pp. 1–19, Oct. 2022.

[11] X. Li, M. Li, P. Yan, G. Li, Y. Jiang, H. Luo, and S. Yin, ''Deep learning attention mechanism in medical image analysis: Basics and beyonds,'' *Int. J. Netw. Dyn. Intell.*, pp. 93–116, Mar. 2023.

[12] Y. Wang, G. Yang, S. Li, Y. Li, L. He, and D. Liu, ''Arrhythmia classification algorithm based on multi-head self-attention mechanism,'' *Biomed. Signal Process. Control*, vol. 79, Jan. 2023, Art. no. 104206.

[13] T. Wang, ''Exploring intelligent image recognition technology of football robot using omnidirectional vision of Internet of Things,'' *J. Supercomput.*, vol. 78, no. 8, pp. 10501–10520, Jan. 2022.

[14] J. Wang and B. Liu, ''Analyzing the feature extraction of football player's offense action using machine vision, big data, and Internet of Things,'' *Soft Comput.*, vol. 23, pp. 1–16, Jun. 2023.

[15] M. Xue and Z. Liu, ''Internet football training teaching data analysis based on an embedded sensor network,'' *Wireless Commun. Mobile Comput.*, vol. 2022, pp. 1–13, Jan. 2022.

[16] X. Yin, C. C. Vignesh, and T. Vadivel, ''Motion capture and evaluation system of football special teaching in colleges and universities based on deep learning,'' *Int. J. Syst. Assurance Eng. Manag.*, vol. 13, no. 6, pp. 3092–3107, Jan. 2022.

[17] A. Zanganeh, M. Jampour, and K. Layeghi, ''IAUFD: A 100 K images dataset for automatic football image/video analysis,'' *IET Image Process.*, vol. 16, no. 12, pp. 3133–3142, May 2022.

[18] H. Liu, C. Adreon, N. Wagnon, A. L. Bamba, X. Li, H. Liu, S. MacCall, and Y. Gan, ''Automated player identification and indexing using two-stage deep learning network,'' *Sci. Rep.*, vol. 13, no. 1, p. 10036, Jun. 2023.

[19] A. Kumari, S. Tanwar, S. Tyagi, and N. Kumar, ''Verification and validation techniques for streaming big data analytics in Internet of Things environment,'' *IET Netw.*, vol. 8, no. 3, pp. 155–163, May 2019.

[20] A. Kumari, S. Tanwar, S. Tyagi, and N. Kumar, ''Blockchain-based massive data dissemination handling in IIoT environment,'' *IEEE Netw.*, vol. 35, no. 1, pp. 318–325, Jan. 2021.

[21] A. Kumari and S. Tanwar, ''ρReveal: An AI-based big data analytics scheme for energy price prediction and load reduction,'' in *Proc. 11th Int. Conf. Cloud Comput., Data Sci. Eng.*, Jan. 2021, pp. 321–326.

[22] Y. Ding, M. Jin, S. Li, and D. Feng, ''Smart logistics based on the Internet of Things technology: An overview,'' *Int. J. Logistics Res. Appl.*, vol. 24, no. 4, pp. 323–345, Apr. 2020.

[23] A. Rehman, T. Saba, M. Kashif, S. M. Fati, S. A. Bahaj, and H. Chaudhry, ''A revisit of Internet of Things technologies for monitoring and control strategies in smart agriculture,'' *Agronomy*, vol. 12, no. 1, p. 127, Jan. 2022.

[24] M. Paiola and H. Gebauer, ''Internet of Things technologies, digital servitization and business model innovation in BtoB manufacturing firms,'' *Ind. Marketing Manag.*, vol. 89, pp. 245–264, Aug. 2020.

[25] N. F. Abdulsattar, M. H. Hassan, S. A. Mostafa, H. S. Mansour, N. A. Alduais, A. Mustapha, and M. A. Jubair, ''Evaluating MANET technology in optimizing IoT-based multiple WBSN model in soccer players health study,'' *Factors Softw. Syst. Eng.*, vol. 61, pp. 1–9, Jul. 2022.

[26] Z. Jiang, ''WITHDRAWN: Real-time monitoring of track and field teaching based on Internet of Things and sensors,'' *Microprocess. Microsyst.*, vol. 28, Jan. 2021, Art. no. 103840.

[27] A. Farrokhi, R. Farahbakhsh, J. Rezazadeh, and R. Minerva, ''Application of Internet of Things and artificial intelligence for smart fitness: A survey,'' *Comput. Netw.*, vol. 189, Apr. 2021, Art. no. 107859.

[28] T. Wu, P. Yang, H. Dai, C. Xiang, X. Rao, J. Huang, and T. Ma, ''Joint sensor selection and energy allocation for tasks-driven mobile charging in wireless rechargeable sensor networks,'' *IEEE Internet Things J.*, vol. 7, no. 12, pp. 11505–11523, Dec. 2020.

[29] R. I. Alfian, A. Ma'arif, and S. Sunardi, ''Noise reduction in the accelerometer and gyroscope sensor with the Kalman filter algorithm,'' *J. Robot. Control (JRC)*, vol. 2, no. 3, pp. 180–189, 2021.

[30] J. Kang, J. Shin, J. Shin, D. Lee, and A. Choi, ''Robust human activity recognition by integrating image and accelerometer sensor data using deep fusion network,'' *Sensors*, vol. 22, no. 1, p. 174, Dec. 2021.

[31] X. Xie, Y. Chen, J. Jiang, J. Li, Y. Yang, Y. Liu, L. Yang, X. Tu, X. Sun, C. Zhao, M. Sun, and Z. Wen, ''Self-powered gyroscope angle sensor based on resistive matching effect of triboelectric nanogenerator,'' *Adv. Mater. Technol.*, vol. 6, no. 10, Aug. 2021, Art. no. 2100797.

[32] M. Webber and R. F. Rojas, ''Human activity recognition with accelerometer and gyroscope: A data fusion approach,'' *IEEE Sensors J.*, vol. 21, no. 15, pp. 16979–16989, Aug. 2021.

[33] V. V. Soshenko, S. V. Bolshedvorskii, O. Rubinas, V. N. Sorokin, A. N. Molyaninov, V. V. Vorobyov, and A. V. Akimov, ''Nuclear spin gyroscope based on the nitrogen vacancy center in diamond,'' *Phys. Rev. Lett.*, vol. 126, no. 19, May 2021, Art. no. 197702.

[34] A. Hernandez-Vicente, D. Hernando, J. Marin-Puyalto, G. Vicente-Rodriguez, N. Garatachea, E. Pueyo, and R. Bailón, ''Validity of the polar H7 heart rate sensor for heart rate variability analysis during exercise in different age, body composition and fitness level groups,'' *Sensors*, vol. 21, no. 3, p. 902, Jan. 2021.

[35] Z. Niu, G. Zhong, and H. Yu, ''A review on the attention mechanism of deep learning,'' *Neurocomputing*, vol. 452, pp. 48–62, Sep. 2021.

[36] C. Janiesch, P. Zschech, and K. Heinrich, ''Machine learning and deep learning,'' *Electron. Mark.*, vol. 31, no. 3, pp. 685–695, Apr. 2021.

[37] E. Fenil, G. Manogaran, G. N. Vivekananda, T. Thanjaivadivel, S. Jeeva, and A. J. C. N. Ahilan, ''Real time violence detection framework for football stadium comprising of big data analysis and deep learning through bidirectional LSTM,'' *Comput. Netw.*, vol. 151, pp. 191–200, Mar. 2019.

[38] M. Stoeve, D. Schuldhaus, A. Gamp, C. Zwick, and B. M. Eskofier, ''From the laboratory to the field: IMU-based shot and pass detection in football training and game scenarios using deep learning,'' *Sensors*, vol. 21, no. 9, p. 3071, Apr. 2021.

[39] M. A. Rahman, ''A deep learning framework for football match prediction,'' *Social Netw. Appl. Sci.*, vol. 2, no. 2, p. 165, Feb. 2020.

[40] P. Thakkar and M. Shah, ''An assessment of football through the lens of data science,'' *Ann. Data Sci.*, vol. 8, no. 4, pp. 823–836, Mar. 2021.

[41] D. Zhou, G. Chen, and F. Xu, ''Application of deep learning technology in strength training of football players and field line detection of football robots,'' *Frontiers Neurorobotics*, vol. 16, Jun. 2022, Art. no. 867028.

[42] M. Muzammal, R. Talat, A. H. Sodhro, and S. Pirbhulal, ''A multi-sensor data fusion enabled ensemble approach for medical data from body sensor networks,'' *Inf. Fusion*, vol. 53, pp. 155–164, Jan. 2020.

[43] R. R. Jha, G. Jaswal, D. Gupta, S. Saini, and A. Nigam, ''PixISegNet: Pixel-level iris segmentation network using convolutional encoder–decoder with stacked hourglass bottleneck,'' *IET Biometrics*, vol. 9, no. 1, pp. 11–24, Nov. 2019.

[44] X. Zou, S. Xu, X. Chen, L. Yan, and Y. Han, ''Breaking the von Neumann bottleneck: Architecture-level processing-in-memory technology,'' *Sci. China Inf. Sci.*, vol. 64, no. 6, Apr. 2021, Art. no. 160404.

[45] S. Chen, Y. Song, J. Su, Y. Fang, L. Shen, Z. Mi, and B. Su, ''Segmentation of field grape bunches via an improved pyramid scene parsing network,'' *Int. J. Agricult. Biol. Eng.*, vol. 14, no. 6, pp. 185–194, 2021.

[46] B. Zhang, T. Hong, R. Xiong, and S. A. Chepinskiy, ''A terrain segmentation method based on pyramid scene parsing-mobile network for outdoor robots,'' *Int. J. Adv. Robotic Syst.*, vol. 18, no. 5, Oct. 2021, Art. no. 172988142110486.

[47] Y. Ouzar, D. Djeldjli, F. Bousefsaf, and C. Maaoui, ''X-iPPGNet: A novel one stage deep learning architecture based on depthwise separable convolutions for video-based pulse rate estimation,'' *Comput. Biol. Med.*, vol. 154, Mar. 2023, Art. no. 106592.

[48] R. Wang, J. Li, S. An, H. Hao, W. Liu, and L. Li, ''Densely connected convolutional networks for vibration based structural damage identification,'' *Eng. Struct.*, vol. 245, Oct. 2021, Art. no. 112871.

[49] R. Zhang, W. Yang, Z. Peng, P. Wei, X. Wang, and L. Lin, ''Progressively diffused networks for semantic visual parsing,'' *Pattern Recognit.*, vol. 90, pp. 78–86, Jun. 2019.

[50] Z. Wu, G. Lin, and J. Cai, ''Keypoint based weakly supervised human parsing,'' *Image Vis. Comput.*, vol. 91, Nov. 2019, Art. no. 103801.

[51] R. Zhang, J. Chen, L. Feng, S. Li, W. Yang, and D. Guo, ''A refined pyramid scene parsing network for polarimetric SAR image semantic segmentation in agricultural areas,'' *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.

[52] S. Zhang, X. Cao, G.-J. Qi, Z. Song, and J. Zhou, ''AIParsing: Anchor-free instance-level human parsing,'' *IEEE Trans. Image Process.*, vol. 31, pp. 5599–5612, 2022.

[53] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, ''DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs,'' *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 4, pp. 834–848, Apr. 2018.

[54] X. Yao, Q. Cao, X. Feng, G. Cheng, and J. Han, ''Scale-aware detailed matching for few-shot aerial image semantic segmentation,'' *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5611711.

[55] J. He, S. Erfani, X. Ma, J. Bailey, Y. Chi, and X. S. Hua, ''α-IoU: A family of power intersection over union losses for bounding box regression,'' in *Proc. NIPS*, vol. 34, 2021, pp. 20230–20242.

[56] Z. Chen, R. Wu, Y. Lin, C. Li, S. Chen, Z. Yuan, S. Chen, and X. Zou, ''Plant disease recognition model based on improved YOLOv5,'' *Agronomy*, vol. 12, no. 2, p. 365, Jan. 2022.

[57] P. Mascagni, A. Vardazaryan, D. Alapatt, T. Urade, T. Emre, C. Fiorillo, P. Pessaux, D. Mutter, J. Marescaux, G. Costamagna, B. Dallemagne, and N. Padoy, ''Artificial intelligence for surgical safety: Automatic assessment of the critical view of safety in laparoscopic cholecystectomy using deep learning,'' *Ann. Surg.*, vol. 275, no. 5, pp. 955–961, May 2022.

[58] X. Huang, L. Zhuo, H. Zhang, X. Li, and J. Zhang, ''Lw-TISNet: Light-weight convolutional neural network incorporating attention mechanism and multiple supervision strategy for tongue image segmentation,'' *Sens. Imag.*, vol. 23, no. 1, p. 6, Jan. 2022.

[59] Y. Zhang, Y. Zhu, X. Liu, Y. Lu, C. Liu, X. Zhou, and W. Fan, ''In-field tobacco leaf maturity detection with an enhanced MobileNetV1: Incorporating a feature pyramid network and attention mechanism,'' *Sensors*, vol. 23, no. 13, p. 5964, Jun. 2023.

[60] L. Ma, Q. Yu, H. Yu, and J. Zhang, ''Maize leaf disease identification based on YOLOv5n algorithm incorporating attention mechanism,'' *Agronomy*, vol. 13, no. 2, p. 521, Feb. 2023.