Electronics, Communications and Networks
A.J. Tallón-Ballesteros et al. (Eds.)

doi:10.3233/FAIA231174

Outlet Recognition Based on Deep Learning and Sensors for Large-Scale Cable Laying

Mengbo CHEN [a], Weiwei WANG [b] and Junhao GENG [a,1]
[a] *Northwestern Polytechnical University, Xi'an, China*
[b] *Xi'an C&C Software Technology Co., Ltd., Xi'an, China*

Abstract. The outlet recognition for large-scale cable laying has problems such as difficulty obtaining prior information, similarity in local recognition information, and low recognition accuracy, making it difficult for most outlet recognition based on real prior information for large-scale cable laying to proceed smoothly. In response to the above issues, this paper proposes an outlet recognition method based on deep learning and sensors for large-scale cable laying. A system of outlet recognition based on deep learning and sensors for large-scale cable laying is formed based on the self-labeling of outlet images, data augmentation of outlet features, and relative pose constraints. The feasibility and effectiveness of this technology have been verified by recognizing outlets at different locations in a large-scale cable laying scene. This technology can recognize and locate outlets, effectively solving the problem of misrecognition at different positions.

Keywords. Self-labeling of outlet images, data augmentation of outlet features, outlet recognition

1. Introduction

In large-scale cable laying scenes, due to an extensive laying range, a large number of cables, and mixed rigidity and flexibility of assembly parts, cable laying outlets have a large distribution span, a large quantity, and a similarity of local features. It leads to outlet recognition for large-scale cable laying based on traditional methods being unable to effectively solve problems such as difficulty obtaining prior information and misrecognition. The target recognition method based on deep learning gradually enters people's vision due to its strong generalization ability, low background interference, and high recognition rate.

Although deep learning driven by a three-dimensional model can solve the problem of difficult acquisition of prior information for a single target, it still cannot solve the problem of target misrecognition caused by the similarity of local features in large-scale scenes. Therefore, new technical research is needed to address the above issue.

Sensors have been widely used in augmented reality due to their multimodal, flexible, and real-time characteristics. Zhang et al. [1] combined UWB (ultra-wideband) positioning with augmented reality, using localization algorithms to locate actors and

[1] Corresponding Author: Jun-Hao Geng, School of Mechanical Engineering, Northwestern Polytechnical University, Xi'an 710000, China; E-mail: gengjunhao@nwpu.edu.cn.

bring a higher sense of immersion to spatial augmented reality; Wang [2] combined iBeacon indoor location with augmented reality to detect user location and recommend suitable exhibits for more efficient navigation; Ma et al. [3] used multi-sensor fusion location technology to track camera positions and achieve an enhanced display of indoor maps. In addition to being applied to augmented reality, sensors are combined with intelligent algorithms for reliable target recognition. Li et al. [4] combined a hybrid tactile sensor with a robot and analyzed the tactile information feedback from the sensor through machine learning to achieve object recognition; Jeon et al. [5] combined visual sensors with convolutional neural networks to achieve classification of retail products, solving the problem of similarity between classes; Chen et al. [6] implemented heterogeneous fusion of LiDAR and visual sensors, combined with deep learning, to achieve obstacle detection and distance measurement in front of vehicles.

In addition, researchers solved the problem of target misidentification by removing interference and feature augmentation. Tao et al. [7] proposed an adaptive interference removal framework that addresses the interference of noisy frames and background clutter in video sequences on person re-identification; Tao et al. [8] proposed a pixel-level supervised neural network that considers pixel-level supervised information to solve the problem of high false alarm rates in smoke recognition; Tao [9] proposed a label-relevance multi-direction interaction network with enhanced deformable convolution to solve the problem of smoke recognition missed alarms.

In summary, there are rich research achievements in applying sensors in augmented reality and deep learning, but they mainly focus on indoor location and navigation, robot object recognition, visual object classification, and other fields. There is still a lack of research on outlet recognition for large-scale cable laying based on deep learning and sensors to address challenges such as difficulties and inaccuracies in obtaining prior information and recognition. In response to the above issues, this paper combines deep learning and sensor technology to research outlet recognition for large-scale cable laying, forming an outlet recognition method based on deep learning and sensors for large-scale cable laying. The main contributions of this paper are as follows.

1) A self-labeling method for outlet images based on virtual agents is proposed. Use a virtual agent to determine the three-dimensional position of the outlet and complete the self-labeling of outlet images. Virtual prior information is provided for outlet recognition, solving the problem of difficulty in obtaining outlets' positions and datasets.

2) A data augmentation method for outlet features based on cable random offset is proposed. This method dynamically changes the cable model by randomly offsetting the discrete points of the cable, introducing flexible feature changes for the outlet features. It compensates for the defect of no change in the local flexible features of the outlet.

3) An outlet recognition method based on relative pose constraints is proposed. This method uses UWB and IMU (inertial measurement unit) to achieve camera pose estimation and model registration. The relative pose between the outlet and the camera is obtained based on the registration information to distinguish the outlets at different positions. It can effectively solve the problem of outlet misrecognition caused by the similarity of local features.

It has been applied and verified through the pre-cabling process of high-altitude cable racks with a recognition accuracy of 97.76%, an average recall rate of 97.82%, an average precision of 97.91%, and an F-Score of 97.86%. It will provide technical support for recognizing outlets for large-scale cable laying in AR guidance.

The rest of this paper is organized as follows. Section 2 introduces the proposed method in detail. Section 3 conducts the application validation. Finally, Section 4 concludes this paper.

2. Research Method

This paper takes the large-scale cable laying model as a prior information source and uses augmented reality registration, deep learning, and other technologies to address the problem of misrecognition caused by the local similarity of the outlet. It researches outlet recognition method based on deep learning and sensors. The overall process is shown in Figure 1.

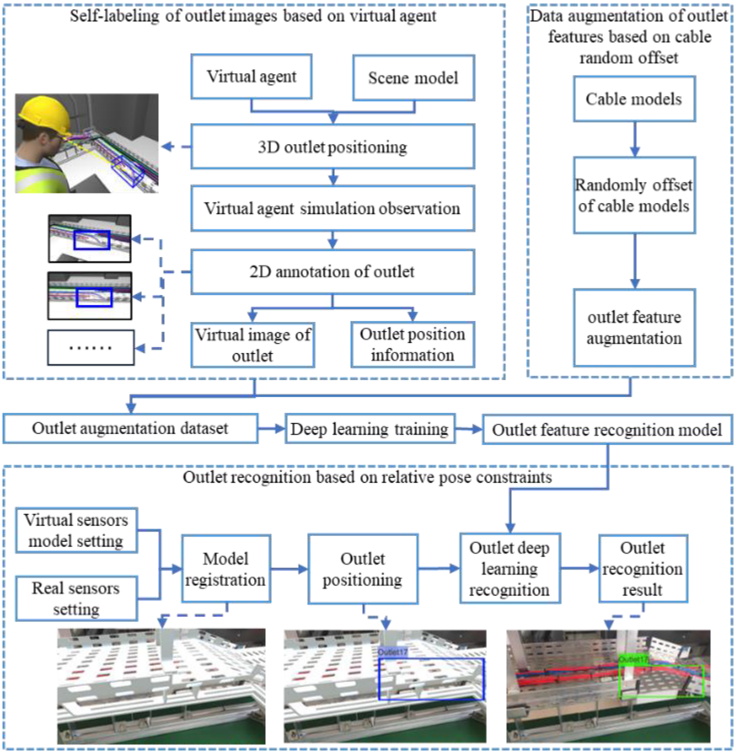

Figure 1. The overall process of outlet recognition method.

First, based on the virtual agent's position and line of sight, locate the outlet's three-dimensional(3D) position and label the outlet in 3D space. Based on ergonomics, simulate virtual agents to mimic human movement and actions in real-world scenes. Based on the virtual agent's simulation results, record the outlet's virtual images and position information and achieve multi-view images self-labeling of the outlet. Second, during the self-labeling process, considering the flexible characteristics of the cable, the 3D cable model in the labeled area of the outlet is randomly offset locally to introduce flexible feature changes for the outlet features for data augmentation. Then, multiple sensor data are used to achieve virtual and real fusion, perform 3D model registration, calculate the camera's pose in the scene, and analyze the pose relationship between the camera and the outlet model to obtain the outlet that can be recognized. Finally, using the feature recognition model corresponding to the outlet, the input image is processed in real-time through deep learning to recognize and locate the outlet in the corresponding image of the real scene.

2.1. Self-labeling of Outlet Images Based on Virtual Agent

The large number and wide distribution of outlets make obtaining available virtual images and location information of outlets in a 3D model difficult. A virtual agent can simulate the changes in human posture and perspective in a real-world scene and observe virtual outlets. Therefore, this paper uses a virtual agent to locate the position of the outlet and automatically obtain the virtual outlet images and their annotation information.

In order to locate the outlet and obtain the position $P_{ol}(x_{ol}, y_{ol}, z_{ol})$ of the outlet in the model coordinate system O_M, this paper uses the position seen by the virtual agent's line of sight as the position of the outlet. As shown in Figure 2 (a), the left-hand coordinate system O_{vh} for the virtual agent observation is denoted by taking the midpoint of the line connecting the virtual agent's eyes as the origin O, the front of the head as the positive direction of the Z-axis OZ_{vh}, and the right of the head as the positive direction of the X-axis OX_{vh}. The positive direction of OZ_{vh} is the line of sight OP. Project the origin O along OP towards the model to obtain P_{ol}. Using the large-scale outlet size as a reference, set the default 3D bounding box size and label the outlet in 3D. Denote a left-hand coordinate system as O_{bb} based on the center of the 3D bounding box, where OX_{bb} is consistent with OX_{vh} and OY_{bb} is the normal vector of the model at the position P_{ol}.

After locating and 3D labeling the outlet, obtaining virtual images and annotation information is necessary. This paper combines critical dimensions such as height and arm length of the human body in ergonomics [10], as well as the range of waist bending (within 45°) and head rotation (±45°), to simulate the motion, head rotation, waist bending, and other actions of a virtual agent. It aims to simulate the position and observation situation of a person in a real-world scene and obtain the appropriate position and perspective of the virtual agent observation outlet, including the range of movement (120°), observation distance (1m), and observation angle (20°), as shown in Figure 2 (b-d). The observation distance is jointly controlled by virtual agent movement and waist bending.

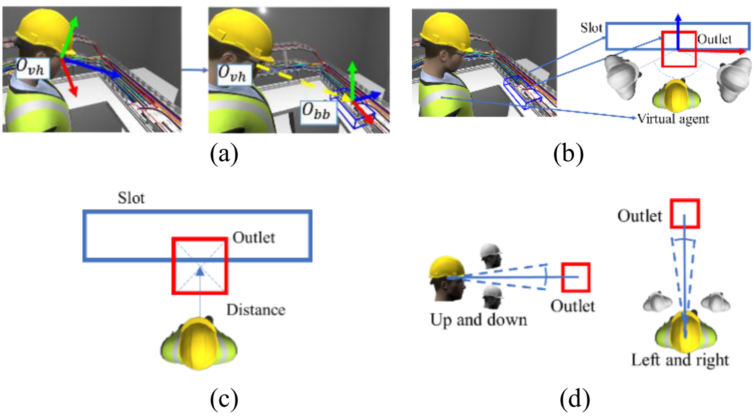

Figure 2. Location of the outlet and virtual agent simulation. (a) Outlet locating. (b) Movement range simulation. (c) Observation distance simulation. (d) Observation angle simulation.

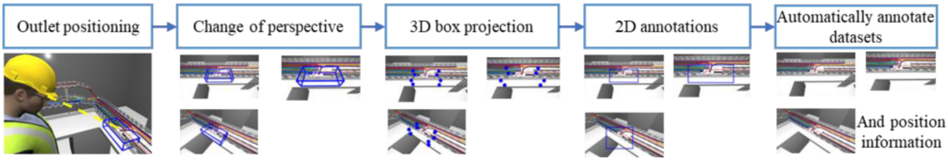

Figure 3. Self-labeling based on virtual agent.

Based on the simulation results of the virtual agent, automatically annotate the outlet image from the virtual agent's perspective, as shown in Figure 3. First, locate the outlet and perform 3D annotation. Second, adjust the posture of the virtual agent based on the simulation results and change the observation perspective. Then, project the eight vertices of the 3D bounding box from the 3D space onto the 2D imaging plane. Next, calculate the maximum bounding rectangle corresponding to eight 2D pixel points to achieve 2D annotation. Finally, record the virtual image of each perspective and the corresponding 2D location information of the outlet, forming a dataset.

2.2. Data Augmentation of Outlet Features Based on Cable Random Offset

Due to the dynamic changes in outlet features during cable laying and the inconsistency between the actual cable laying form and the design form, a dataset based on rigid part features cannot meet the requirements of outlet recognition. Therefore, this paper combines the characteristics of cable flexibility and variable laying positions to expand the data of outlet features.

In 3D simulation, a virtual cable model is obtained by controlling the motion of discrete nodes to drive the display of geometry, with physical attributes as constraints [11]. Therefore, the cable model is randomly offset by controlling the position changes of discrete nodes, introducing flexible features for the outlet. Figure 4 represents a cable local offset based on discrete nodes.

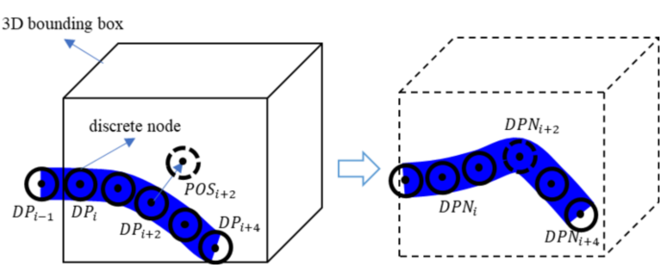

Figure 4. A cable local offset based on discrete nodes.

First, obtain the cable discrete nodes set $DP = \{dp_i(xdp_i, ydp_i, zdp_i)|i \in \{n, \dots, N\}\}$ corresponding to the cable within the 3D bounding box and use the cable end node as the local cable endpoint. Second, generate a random offset vector for each discrete node, where the set of the random offset vector is represented as $POS = \{pos_i(xos_i, yos_i, zos_i)|i \in \{n, \dots, N\}\}$. Then, the random offset vector pos_i is sequentially applied to the node dp_i to obtain the cable node located in the new position dpn_i, as Eq. (1). Next, driven by the physics engine (such as gravity, elasticity, interference), obtain the updated cable discrete nodes set $DPN = \{dpn_i(xdpn_i, ydpn_i, zdpn_i)|i \in \{n, \dots, N\}\}$. Finally, using the DPN as the new DP, a corresponding random offset vector POS is added to the next cable node, and the cable position is updated in the physical engine until all cable nodes are offset and updated, resulting in the final local cable model.

$$dpn_i = dp_i + pos_i \tag{1}$$

During self-labeling at the outlet images, the cable model is randomly shifted locally from the perspective of each virtual observation camera to expand the original dataset. The example of data augmentation is shown in Figure 5.

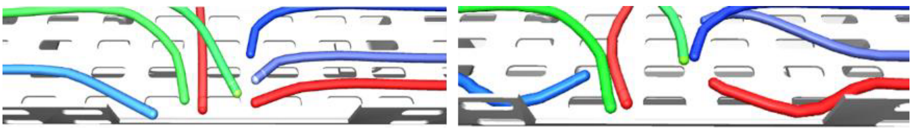

Figure 5. Example of dataset augmentation of outlet features. (a) Original outlet features from a certain perspective. (b) Outlet features after random offsets of cables.

2.3. Outlet Recognition Based on Relative Pose Constraints

The large-scale outlet has the characteristic of local similarity in features, which leads to the incorrect recognition of the outlet. This paper distinguishes the outlet through the relative pose constraints between the camera and the outlet to solve the problem of misrecognition.

UWB and IMU are gradually being applied to model registration in the AR field [12], in addition to indoor positioning. This paper uses UWB, IMU, and Camera for sensor fusion to achieve camera and model pose estimation. The UWB base station is arranged in real space based on the virtual design position, and the UWB receiver and IMU are fixed on the camera. The camera position is obtained based on the UWB positioning algorithm [13], and the camera attitude is obtained based on the IMU attitude. Estimate the camera pose to achieve model registration. Afterward, an outlet that can be recognized is obtained based on the pose constraints between the outlet and the camera. The corresponding outlet recognition result is obtained through deep learning recognition, as shown in Figure 6.

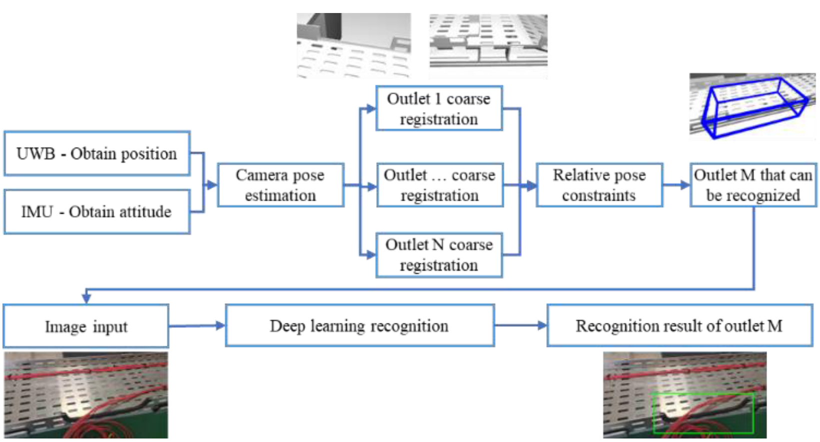

Figure 6. Recognition process of outlet based on relative pose constraints.

Due to model registration errors caused by sensor errors, the projection result of outlet registration cannot be directly used as the recognition result. However, the wide distribution of outlets for large-scale cable laying makes them unique within a specific range so that the recognition result can be determined by the relative pose of the camera and the outlet. First, obtain the current outlet set $OL = \{ol_i(x_i, y_i, z_i) | i \in \{1, \dots, N\}\}$ to be recognized based on the design information, including its corresponding 3D bounding box information. Second, based on the result of virtual and real fusion, calculate the position $P_r(x_r, y_r, z_r)$ and attitude $R_r(x_r, y_r, z_r)$ of the virtual camera. Then, calculate the distance set $PO = \{po_i | dis(P_r, ol_i), i \in \{1, \dots, N\}\}$ between the camera and the current outlet OL where $dis(\,)$ represents the distance between two points, the vector set $PL = \{\overrightarrow{pl_i} | ol_i - P_r, i \in \{1, \dots, N\}\}$ of the camera pointing towards the outlet, and the angle set $RO = \{ro_i | eul(z(R_r), pl_i), i \in \{1, \dots, N\}\}$ between the camera orientation and PL, where $z(\,)$ represents the z-axis orientation vector, $eul(\,)$ represents the angle between vectors. Next, when the distance po_i and angle ro_i satisfy the Eq. (2), an outlet ol_i that can be recognized is obtained. Finally, deep learning recognition is performed when ol_i is visable, and the recognition result is the corresponding outlet.

$$\begin{cases} po < 1 \\ ro < 90^o \end{cases} \tag{2}$$

3. Application Verification

Based on the Windows system, a prototype system of outlet recognition based on deep learning and sensors for large-scale cable laying has been developed using the Unity 3D 2020 development platform, Visual Studio 2019 development environment, and C Sharp development language. In addition, we chose YOLOv7 as the deep learning validation model due to its high target detection accuracy and real-time performance. We used Huawei Mate 50 with high performance and a built-in IMU, and the DW1000 UWB system. This section verifies the feasibility and effectiveness of the technology through the recognition of the outlet during the pre-cabling process of high-altitude cable racks, as shown in Figure 7.

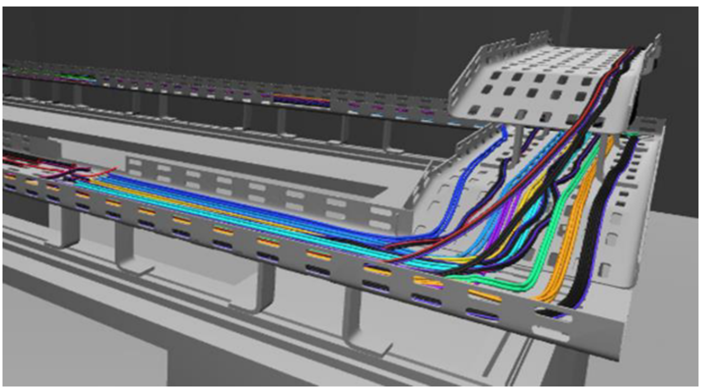

Figure 7. The pre-cabling scene of high-altitude cable racks.

The selected verification scene has a large number and a wide distribution span of outlets, which can meet the characteristics of outlets for large-scale cable laying. Self-label the outlet and refer to the size of the outlet in the model to set the default 3D bounding box with a length of 0.3m, a height of 0.08m, and a width of 0.16 m. During the self-labeling process, data augmentation is carried out, where the maximum offset value of each axis in the random offset vector of the discrete nodes is set to the size of the cable diameter, and 1400 images and data are obtained for each outlet; Divide each outlet images dataset into a training set, validation set, and testing set in a 3:1:1 ratio. With 200 training epochs, the average loss of the training model is 0.01, and the mean average precision on the testing set is 98.57%. Arrange sensors in the scene for model registration, set a confidence level of 0.8, and then perform recognition of the outlet with an accuracy of 97.76%, an average recall rate of 97.82%, an average precision of 97.91%, and an F-Score of 97.86%, as shown in Table 1. Part of the verification results are shown in Figure 8, which shows that the outlets at different positions have local similarities.

Table 1. Evaluation results of outlet recognition

Evaluation metric	Accuracy	Average recall	Average precision	F-Score
Result	97.76%	97.82%	97.91%	97.86%

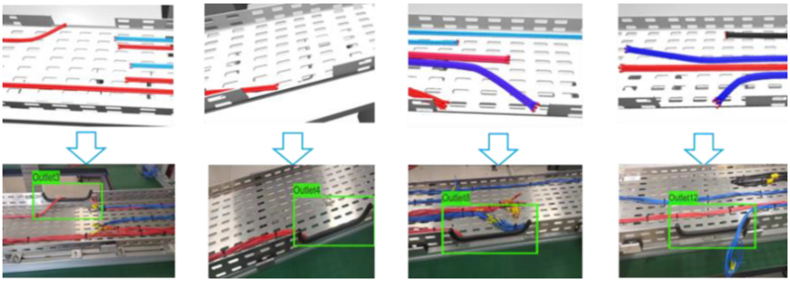

Figure 8. Outlet recognition results.

The verification results indicate that the outlets at different positions can be effectively recognized. After verification, the technology in this paper is feasible and can meet the requirements of outlet recognition for large-scale cable laying, effectively solving outlet misrecognition caused by local similarity in large-scale scenes.

4. Conclusion and Future Work

This paper researches outlet recognition for large-scale cable laying based on self-labeling of outlet images, data augmentation of outlet features, and relative pose constraints. It solves the problems of outlet recognition based on real prior information, such as the inability to carry out recognition and misrecognition due to a large distribution span, similarity of local features, and changes in outlet features. The outlet recognition method based on deep learning and sensors has been developed and validated in large-scale cable laying scenes. The effectiveness of outlet recognition was verified through the pre-cabling process of high-altitude cable racks, and the goal of outlet recognition for large-scale cable laying was basically achieved. The recognition technology based on deep learning and sensors studied in this paper aims to recognize the outlet of large-scale cable laying. In future work, we will consider researching the recognition of flexible cables in large-scale scenes to complete the full scene recognition for large-scale cable laying.

Acknowledgement

This work was supported by the National Natural Science Foundation of China under Grant No. 52175484 and Shaanxi Qin Chuang Yuan "Scientists + Engineers" Team Development Project.

References

[1] Zhang YX, Shen YR, Zhang WW, Zhu ZQ, Ma PF. Design of an interactive spatial augmented reality system for stage performance based on UWB positioning and wireless triggering technology. Appl Sci-Basel. 2019;9(7):9.
[2] Wang CS. An AR mobile navigation system integrating indoor positioning and content recommendation services. World Wide Web. 2019;22(3):1241-62.
[3] Ma W, Zhang S, Huang JC. Mobile augmented reality based indoor map for improving geo-visualization. PeerJ Comput Sci. 2021;7:25.
[4] Li N, Yin ZH, Zhang WG, Xing CY, Peng TJ, Meng B, et al. A triboelectric-inductive hybrid tactile sensor for highly accurate object recognition. Nano Energy. 2022;96:9.
[5] Jeon JY, Kang SW, Lee HJ, Kim JS. A retail object classification method using multiple cameras for vision-based unmanned kiosks. IEEE Sens J. 2022;22(22):22200-9.
[6] Chen PY, Lin HY, Pai NS, Huang JB. Construction of edge computing platform using 3D LiDAR and camera heterogeneous sensing fusion for front obstacle recognition and distance measurement system. Processes. 2022;10(9):19.
[7] Tao H, Duan Q, An J. An adaptive interference removal framework for video person re-identification. IEEE Transactions on Circuits and Systems for Video Technology. 2023;33(9):5148-59.
[8] Tao H, Duan Q, Lu M, Hu Z. Learning discriminative feature representation with pixel-level supervision for forest smoke recognition. Pattern Recognition. 2023;143:109761.
[9] Tao H. A label-relevance multi-direction interaction network with enhanced deformable convolution for forest smoke recognition. Expert Systems with Applications. 2024;236:121383.
[10] Yin MY, Li JG. A systematic review on digital human models in assembly process planning. Int J Adv Manuf Technol. 2023;125(3-4):1037-59.
[11] Lv NJ, Liu JH, Xia HX, Ma JT, Yang XD. A review of techniques for modeling flexible cables. Comput-Aided Des. 2020;122:16.
[12] Geng J, Chen M, Zhao X, Liu Y, Ma Y. A Markerless AR Guidance Method for Large-Scale Wire and Cable Laying of Electromechanical Products. IEEE Transactions on Industrial Informatics. 2023:1-14.
[13] Che FH, Ahmed QZ, Lazaridis PI, Sureephong P, Alade T. Indoor Positioning System (IPS) Using Ultra-Wide Bandwidth (UWB)-For Industrial Internet of Things (IIoT). Sensors. 2023;23(12):29.