International Journal of Heat and Technology
Vol. 37, No. 1, March, 2019, pp. 345-350
Journal homepage: http://iieta.org/Journals/IJHT

# Robotic Slag Offloading and Process Improvement of Magnesium Smelting in Pidgeon Process with Faster Region-based Convolutional Neural Network

Jin Hua[1,2*], Lile He[1], Keding Yan[2], Min Wang[3]

[1] College of Mechanical Engineering, Xi`an University of Architecture and Technology, Xi'an 710055, China
[2] College of Electronic Information Engineering, Xi'an Technological University, Xi'an 710021, China
[3] College of Electronic Engineering, Xi'an Petroleum University, Xi'an 710065, China

Corresponding Author Email: huahua_dz@aliyun.com

https://doi.org/10.18280/ijht.370141

**Received:** 12 October 2018
**Accepted:** 20 December 2018

***Keywords:***
*robotic slag offloading, positioning, Pidgeon process, magnesium smelting, faster region-based convolutional neural network (Faster R-CNN)*

## ABSTRACT

This paper attempts to replace the traditional manual slag offloading of magnesium smelting in Pidgeon process with robotic slag removal. Specifically, the high-temperature infrared dot matrix was used to measure the slag positions indirectly; the faster region-based convolutional neural network (Faster R-CNN) was trained with thermal image of the reduction jar as the dataset; the isothermal image of the reduction jar was plotted based on the slag centers, and adopted to detect the opening direction of the jar and the slag positions. The indirect measurement results show that the actual internal temperature of the jar can be detected accurately through repeated experiments, with an error of less than 10 °C. Finally, the proposed method was verified through a case study on 1,000 images. The results show that our model can correctly identify more than 90% of crude magnesium in the actual jar.

## 1. INTRODUCTION

Magnesium is a lightweight green material widely used in automobile, electronic communication, medical service and aerospace. It can be synthesized with one or more metal/nonmetal into alloys of various special properties, breaking through the limitations of common metal materials in many fields. At present, 98% of magnesium in China is produced by the low-efficient, energy-intensive Pidgeon process. The traditional Pidgeon process has a long reduction cycle that requires many reduction jars, and causes serious pollution to the environment. These defects push up the production cost and hinder the development of magnesium industry in China.

Figure 1 is a photo taken at a magnesium factory in Shenmu County, northwestern China’s Shaanxi Province. Traditionally, the magnesium is smelted in the Pidgeon process through following steps: blend the magnesium-containing calcined dolomite evenly with the reducing agent (ferrosilicon powder) and the mineralizer (fluorite powder) at a certain mix ratio; press the mixture into pellets; load a fixed amount of the pellets into the test tube in the reduction jar. The following takes place in an airtight regenerative reduction furnace, where the reduction jar is located (Figure 2). The equipment is water-cooled, and then vacuumed. After that the reduction jar is heated with lump coal or gas. In this way, magnesium can be reduced from the calcined dolomite in the form of vapor, and released to the vacuum zone of the jar through the voids in the pellets. Then, the magnesium vapor will enter the crystallizer through a small hole in the heat shield. After water cooling, crude magnesium can be obtained as solid crystalline.

To sum up, the magnesium-containing calcined dolomite gives off magnesium vapor under high temperature (1,150~1,250 °C) and vacuum (1.33~133 Pa), which later crystallizes and condenses to obtain crude magnesium. The reaction can be described as follows:

$$2(MgO\bullet CaO)+Si(Fe)=2Mg(g)+2CaO\bullet SiO_2+(Fe) \qquad (1)$$

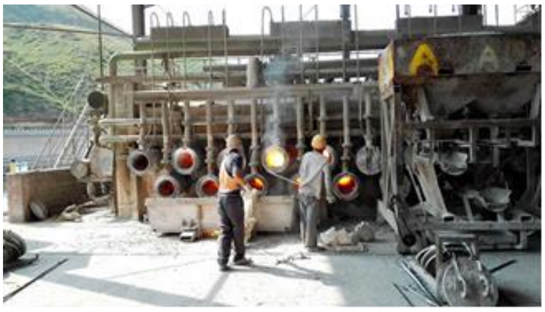

**Figure 1.** A magnesium factory in Shenmu County

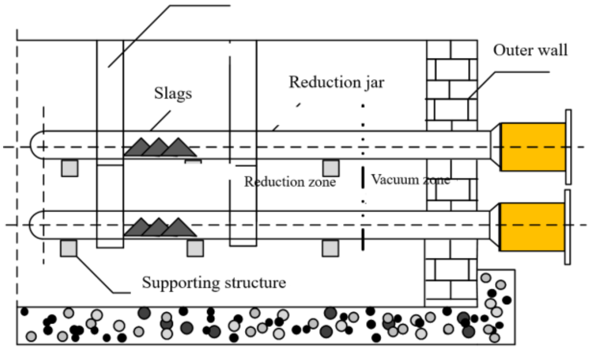

**Figure 2.** Structure of regenerative reduction furnace

As shown in Figure 2, the magnesium smelting produces a lot of slags. It is very laborious and time-consuming to offload the slags manually, not to mention the danger of working at such a high temperature in the furnace. To solve the problem, this paper attempts to introduce multi-legged robot to realize automatic slag offloading in magnesium smelting.

## 2. LITERATURE REVIEW

The slag offloading effect of the robot depends on the robot's near-infrared vision [1-3], i.e. the ability to locate the high-temperature targets (slags). An important way to improve this ability is to couple infrared thermal imaging sensor with high-precision laser/ultrasonic depth sensor [4-6]. Below is a brief review of the relevant studies in China.

Using a binocular infrared thermal imager, Chen et al. measured high-temperature targets in infrared binocular vision, developed a calibration method for infrared binocular positioning, and designed a special infrared calibration target [7]. This positioning method requires sensitive sensors, high cost and meticulous calibration.

Liu et al. dynamically positioned targets with an infrared pyroelectric sensor in the following manner [8]: Considering its passive detection feature, the pyroelectric sensor was mounted on a rotating platform; when a target entered the blind zone, the hotspot signal received by the sensor produced a step change, signifying the presence of the target; then, a pyroelectric sensor network was built and the acquisition area was calibrated to locate the target. However, the experimental environment is too complex to build easily and the positioning is both inaccurate and time-consuming.

Pan et al. integrated the charge-coupled device (CCD) camera with infrared filter to create an infrared stereovision, put forward a discrete traversal visual target search algorithm, and actively traversed the search space for the visual targets through discrete control of the rotation and tilting of the two cameras [9]. There are two problems with their approach: only the position of the cameras is adjustable in the experimental platform, and the algorithm is very complex.

## 3. METHODOLOGY

### 3.1 Process improvement

The magnesium smelting in Pidgeon process was improved in many aspects. The improved method is compared with the traditional approach in Figure 3. It can be seen that the steps from dolomite calcination to the crystallization of crude magnesium remain unchanged in the improved method. The main difference between the two methods lies in the step of slag offloading. In the traditional process, the slags left in the reduction zone at the end of the furnace are offloaded by workers with long rods This approach is purely empirical, and often unable to clean up all the slags. The residual slags will hinder the reaction in the next cycle. What is worse, the hot working environment may endanger the health of the workers.

In the improved process, the slags are automatically removed by a robot, which indirectly measures the position of slags using a high-temperature infrared dot matrix [10, 11]. Specifically, the robot shots a thermal image of the reduction jar, taking it as the training dataset of the machine learning network, and draws an isothermal diagram for slag center recognition. On this basis, the faster region-based convolutional neural network (faster R-CNN) was adopted to detect the opening direction of the reduction jar, the position of crude magnesium, and monitor the temperature field. Through the arms, the robot can detect the depth of a desired target point, and determine whether the object at that point is crude magnesium or slag.

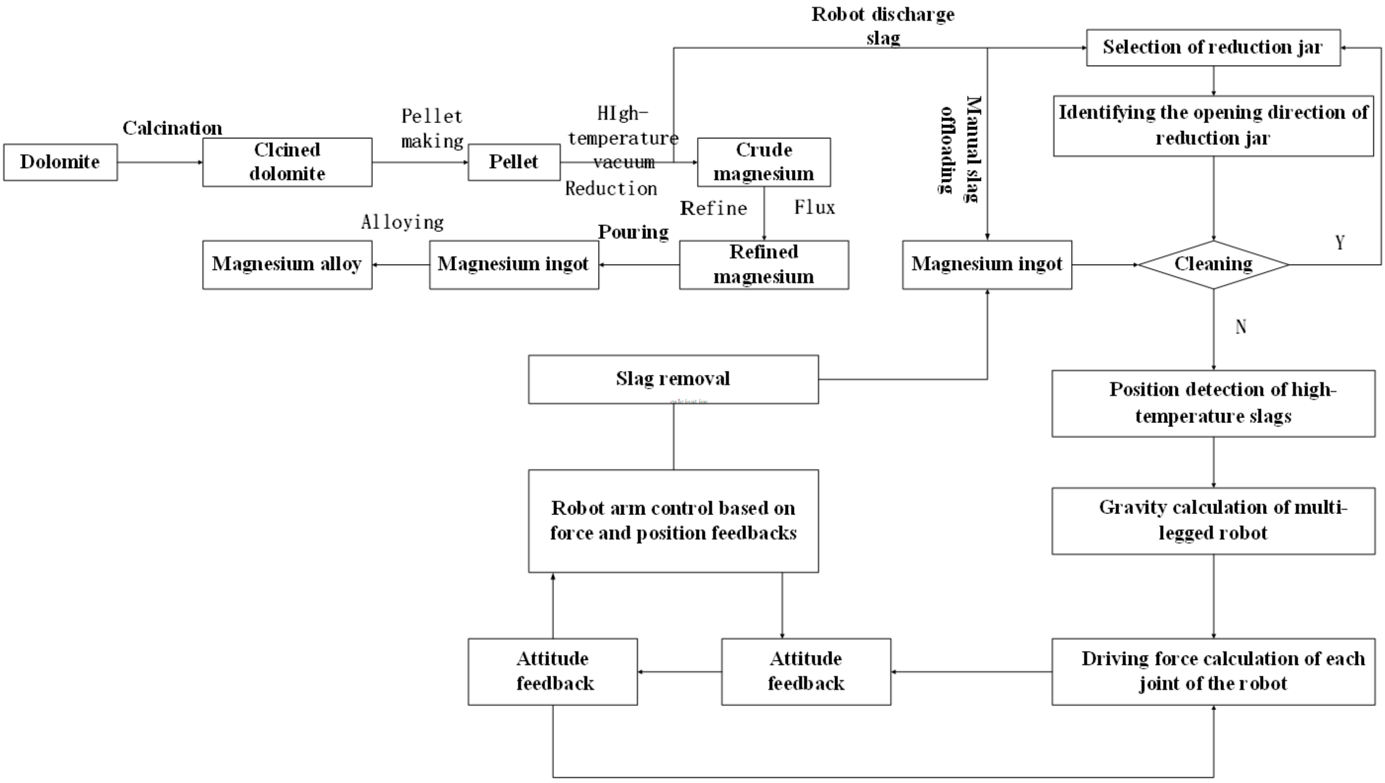
Robot discharge slag
Selection of reduction jar
Calcination
Pellet making
HIgh-temperature vacuum Reduction
Manual slag offloading
Identifying the opening direction of reduction jar
Dolomite
Clcined dolomite
Pellet
Crude magnesium
Refine
Flux
Alloying
Pouring
Y
Magnesium alloy
Magnesium ingot
Refined magnesium
Magnesium ingot
Cleaning
N
Slag removal
Position detection of high-temperature slags
Robot arm control based on force and position feedbacks
Gravity calculation of multi-legged robot
Attitude feedback
Attitude feedback
Driving force calculation of each joint of the robot

**Figure 3.** Flow chart of traditional and improved magnesium smelting in Pidgeon process

### 3.2 Faster R-CNN construction

As mentioned above, the faster R-CNN was selected to support the robotic slag offloading. The faster R-CNN adopted for our research uses the region proposal network (RPN) to acquire the candidate anchor boxes [12-16]. The layers of the faster R-CNN are detailed in Figure 4.

(1) Fully-convolutional network (FCN) layer

The FCN can handle input images of various sizes. This layer collects the deconvolution features of each input image, forming a feature map. In addition, the FCN ensures that the size of the output image is the same as that of the input image, such that the spatial information of the input image is preserved for accurate classification. The FCN provides a feature map for subsequent layers such as RPN. In other words, our fast R-CNN only needs to extract image features once, reducing the computation time of the system.

(2) RPN layer

The RPN is a simple FCN consists of a convolutional layer, an activation function and the classification & regression layer. Different from the selective search, the RPN generates k anchor boxes in each sliding window. The width-height ratio of the k anchor points is 1:1, 1:2 and 2:1. The features of these boxes are extracted from the FCN layer. To extract feature submaps (proposals), the Softmax function judges if the candidate box is in the foreground or the background, while the bounding box regressor adjusts the box position. The acquired features lay the basis for classification and bounding box regression.

(3) Region-of-interest (ROI) pooling layer

The ROI pooling normalizes proposals of different sizes into the same size, leading to faster training and testing and more accurate detection. This layer is necessary because the RPN generate proposals of varied sizes and shapes, while the traditional CNN only accepts fixed size inputs and fixed dimension/vector outputs.

(4) Classification layer

In this layer, the object classification, position adjustment and bounding box regression are carried out based on the feature map outputted by the ROI layer.

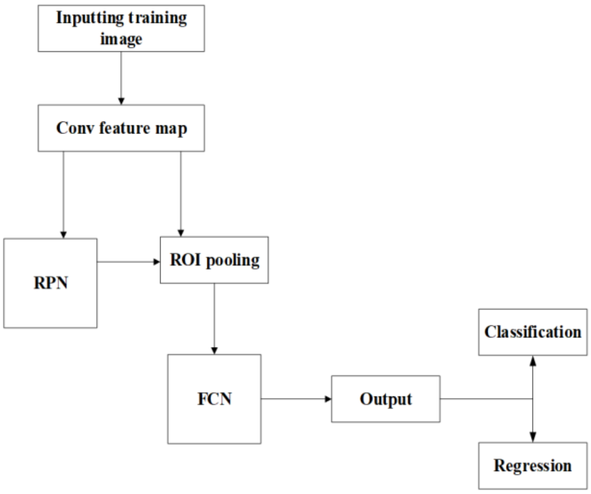

**Figure 4.** Flow chart of the proposed recognition method

## 4. EXPERIMENT

### 4.1 Faster R-CNN structure

The features were extracted from 51×39,512 channels of image data. In our experiment, nine candidate anchor boxes are selected, whose scales are {$128^2$, $256^2$, $512^2$} respectively. The anchor boxes were selected in the RPN layer. The parameters of this layer are listed in Table 1. In the classification layer, each output position (cls_score) refers to the probability for the nine anchor boxes to fall in the foreground or background when they are outputted through the 256 channels. In the regression layer, each output position (bbox_pred) describes the scales of the nine boxes with 4 translation scaling parameters.

**Table 1.** Parameters of the RPN layer

| Layer | Bottom | Top | Kernel size | Pad | Stride | Number | Channel |
|---|---|---|---|---|---|---|---|
| rpn_conv/3*3 | conv5 | rpn/output | 3 | 1 | 1 | 1 | 256 |
| rpn_relu/3*3 | rpn | rpn/output | | | | 1 | 256 |
| rpn_cls_score | rpn/output | rpn/output | 1 | 0 | 1 | 1 | 3 |
| rpn_bbox_pred | rpn/output | rpn_bbox_pred | 1 | 0 | 1 | 1 | 12 |
| cls_score_reshape | rpn_cls_score | rpn_cls_score_reshape | | | | 1 | 2 |

The network parameter $\omega$ was initialized based on the choice of the sharing feature. Both the RPN and fast R-CNN require an original feature extraction network [17]. Hence, the ImageNet Classification Library was used to train the initial parameter. Then, the network was fine-tuned by the specified dataset. For anchor box i, the parameters of the prediction box can be expressed as [18]:

$$t_x=(x-x_a)/w_a,\ t_y=(y-y_a)/h_a \quad (2)$$

$$t_w=log(w/w_a),\ \mathrm{t_h}=log(h/h_a) \quad (3)$$

$$t_x^*=(x^*-x_a)/w_a,\ t_y^*=(y^*-y_a)/h_a \quad (4)$$

$$t_w^*=log(w^*-w_a)/w_a,\ t_h^*=log(h^*-h_a)/h_a \quad (5)$$

where x, y, w and h are the center coordinates predicted by the RPN. For each anchor box, the center coordinates $x_a$, $y_a$, $w_a$ and $h_a$ depend on the size and ratio of the box. Meanwhile, the loss function of the real target box can be defined as:

$$L(\{p_i\},\{t_i\})=1/N_{cls}\sum L_{cls}(p_i,p_i^*)+1/N_{reg}\sum p_i^*L_{reg}(t_i,\mathrm{t_i}^*) \quad (6)$$

where x*,y*, w* and h* are the center coordinates of the real target box; $p_i$ is the probability of recognizing an object at anchor point i; $p*_i$ is the probabilities of the anchor box after traversing the entire image; $t_i$ and $t*_i$ are the positions of the prediction box and the anchor box, respectively; $L_{reg}$ is the Softmax loss of two classes. If the label assigned by the anchor box is true, then $p*_i$ =1. The points corresponding to $p_i$ and $p*_i$ are the four vertices of the candidate anchor box, whose coordinates are $t_i$ ($t_x$, $t_y$, $t_w$, $t_h$) and $t*_i$ ($t*_x$, $t*_y$, $t*_w$, $t*_h$).

### 4.2 Dataset preparation

A total of 100 8×8 infrared array sensors were selected, forming an 80×80 temperature data acquisition array. In each second, the array captured 6,400 pieces of temperature data. The data were processed to generate pseudo-color temperature distribution maps. Then, the temperature data matrix was plotted as a 2D pixel map, where the coordinates are expressed as ($u_i$, $v_i$). The temperature detected at each point is denoted as $t_i$.

Following the above procedure, 6,000 sets of temperature data were collected by the 80×80 array in the said magnesium factory, from the same shooting angle of Figure 1. In this way, 6,000 pseudo-color temperature distribution maps were obtained from array data. Then, the high and sub-high temperature ranges of each map were marked by the ROI framed standard (Figure 5).

The registration results of 5,000 field images and 5,000 pseudo-color distribution maps were selected as the training set of our model, the remaining 1000 sets of temperature data were used as the testing set, and the remaining 1,000 field images were taken as the verification set.

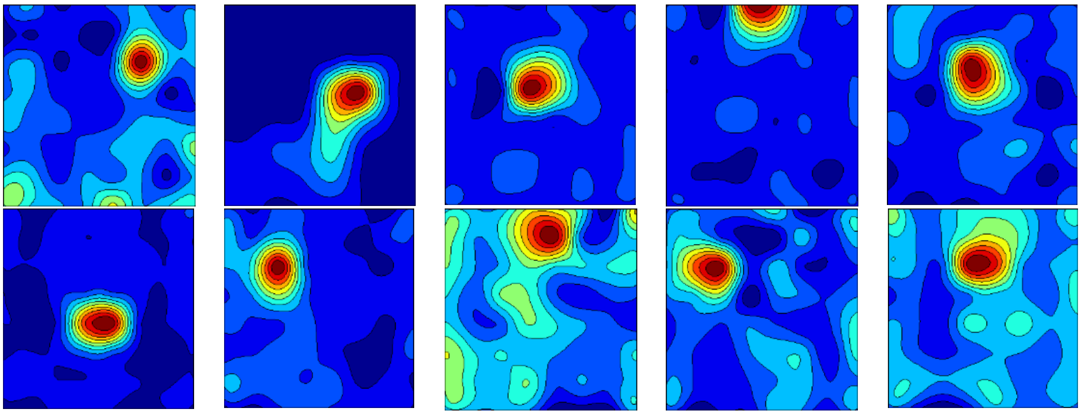

**Figure 5.** Thermal images after isotherm labeling

### 4.3 ROI marking

The ROI of each image was marked as per the requirement of the faster R-CNN. The coordinates of each ROI (excluding the slag position) were determined by software on the registration dataset and saved as an XML file.

### 4.4 Experimental environment

The Caffe deep learning framework was selected for our research. The platform was equipped with NVIDIA GTX1050Ti graphics card and Ubuntu 16.04. Under the machine learning framework of TensorFlow, our model was trained by two Visual Geometry Group (VGG) models (VVG16 and VGG24).

### 4.5 Parameter setting

The solver parameters of the network were set as follows:

*Stage1_fast_rcnn_train.pt*: base_lr *(fast):* 0.001, lr_policy *(fast):* "step", stepsize *(fast):* 30000, display *(fast):* 20, average_loss *(fast):* 100, momentum *(fast):* 0.9, weight_decay *(fast):* 0.0005.

*Stage1_rpn_train.pt*: base_lr *(fast):* 0.001, lr_policy *(fast):* "step", stepsize *(fast):* 60000, display *(fast):* 20, average_loss *(fast):* 100, momentum *(fast):* 0.9, weight_decay *(fast):* 0.0005.

*Stage2_faster_rcnn_train.pt*: base_lr *(faster):* 0.001, lr_policy *(faster):* "step", stepsize *(faster):*30000, display *(faster):* 20, average_loss *(faster):* 100, momentum *(faster):* 0.9, weight_decay *(faster):* 0.0005.

*Stage2_rpn_train.pt*: base_lr *(faster):* 0.001, lr_policy *(faster):* "step", stepsize *(faster):* 60000, display *(faster):* 20, average_loss *(faster):* 100, momentum *(faster):* 0.9, weight_decay *(faster):* 0.0005.

The partial parameters of the network were set as follows: *data_param_str_num_classes*: 4, *cls_score_num_output*: 4, *bbox_pred_num_output*: 16.

### 4.6 Results analysis

The training results of VGG16 and VGG24 are listed in Table 2 below.

**Table 2.** The training results of the VGG16 and VGG24

| Pascal_voc. Model | High-Temperature Slags | Low-Temperature Slags | Background |
|---|---|---|---|
| VGG16 | 0.9739 | 0.8806 | 0.3604 |
| VGG24 | 0.9782 | 0.8812 | 0.3646 |

As shown in Table 2, the background value of the training detection improved with the increase of the network depth. The recognition rate of high-temperature slags, which have a few training data, was relatively high.

Next, the proposed method was applied to locate high-temperature slags. The results were subjected to grayscale processing and binarization, yielding the feature point coordinates of high-temperature targets on the pixel plane.

**Table 3.** Feature point coordinates of high-temperature targets on the pixel plane

| Target | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| U-axis Coordinates | 807 | 390 | 1159 | 822 | 306 | 1268 |
| V-axis Coordinates | 387 | 385 | 382 | 329 | 295 | 294 |

As shown in Table 3, six groups of data were obtained from the positioning experiment on high-temperature slags on the pixel plane. Then, the global coordinates of the slags were computed [19, 20], and compared with the measured coordinates in actual operation (Table 4).

The comparison between the global coordinates and measured coordinates shows that our method can control the error at about 10mm.

**Table 4.** Comparison between the global coordinates and measured coordinates

| Group | Global Coordinate (mm) | Measured Coordinates (mm) | Error (mm) |
|---|---|---|---|
| 1 | (-4.8, 667.4, 95.0) | (-10.0, 675.0, 95.0) | 9.2 |
| 2 | (165.6, 674.8, 95.0) | (155.0, 675.0, 95.0) | 10.6 |
| 3 | (-146.3, 682.6, 95.0) | (-145.0, 675.0, 95.0) | 7.7 |
| 4 | (-11.1, 902.2, 95.0) | (-10.0, 900.0, 95.0) | 2.5 |
| 5 | (314.6, 1126.2, 95.0) | (310.0, 1125.0, 95.0) | 4.8 |
| 6 | (-292.6, 1135.5, 95.0) | (-290.0, 1125.0, 95.0) | 10.8 |

## 5. CONCLUSIONS

This paper introduces faster R-CNN to identify the slags in the reduction jar of magnesium melting process. The RPN was adopted to generate an efficient and accurate region proposal. In each image, the proposed method locates the difference and range between slag center and slag temperature. The isothermal image of the reduction jar was plotted based on the slag centers, and adopted to detect the opening direction of the jar and the slag positions. The indirect measurement results show that the actual internal temperature of the jar can be detected accurately through repeated experiments, with an error of less than 10°C. Finally, the proposed method was verified through a case study on 1,000 images. The results show that our model can correctly identify more than 90% of crude magnesium in the actual jar.

## ACKNOWLEDGEMENT

This paper was supported by Shaanxi provincial science and Technology Department's industrial field funding (GY2017-027); National sonar Key Laboratory Open Fund (6142109KF2018) National Fund Youth Project (11804263).

## REFERENCES

[1] Cui YP, Lin YC, Zhang XL. (2005). Study on camera calibration for binocular vision based on neural network. Journal of Optoelectronics Laser 16(9): 1097-1100. https://doi.org/10.1081/CEH-200044273

[2] Yim W, Singh SN. (1996). Predictive end-point trajectory control of elastic manipulators. Journal of Robotic Systems 13(9): 561–569. https://doi.org/10.1002/(sici)1097-4563(199609)13:9<561::aid-rob1>3.0.co;2-k

[3] Škrjanc I, Klančar G. (2017). A comparison of continuous and discrete tracking-error model-based predictive control for mobile robots. Robotics & Autonomous Systems 87: 177-187. https://doi.org/10.1016/j.robot.2016.09.016

[4] Wang JC, Chen WD, Hu SY, Zhang X. (2010). Mobile robot localization in outdoor environments based on Near-infrared vison. Robot 32(1): 97-103. https://doi.org/10.3724/SP.J.1218.2010.00097

[5] Yang RB, Cheng WB, Qian Q, Zhang Q, Qian J, Pan YJ. (2017). Fall detection algorithm based on multi feature extraction in infrared image. Infrared Technology 39(12): 1131-1138.

[6] Zhuang Y, Wang W, Wang K, Xu XD. (2005). Mobile robot indoor simultaneous loealization and map ping using laser range finder and monoeular vision. Acta Automatica Sinica 31(6): 925-933.

[7] Chen YH, Zhao MR, Lin YC, Song L, Cui SM, Tao L. (2012). Study on infrared edge enhancement algorithm for infrared binocular vision. Laser & Infrared 42(5): 526-529. https://doi.org/10.3969/j.issn.1001-5078.2012.05.011

[8] Liu QJ, Yang W, Liu YW. (2014). Research and design of multi-node positioning system based on pyroelectric infrared senso. Computer Measurement & Contrl 22(9): 2947-2948, 2956. https://doi.org/10.3969/j.issn.1671-4598.2014.09.071

[9] Pan M, Meng F, Bao H, Ni WC. (2013). Research on infrared stero visual localization system. Micricintrollers & Embadded Systems 13(3): 7-10.

[10] Redmon J, Divvala S, Girshick R, Farhadi A. (2016). You only look once: Unified, real-time object detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition: 779-788.

[11] Zitnick CL, Dollár P. (2014). Edge boxes: Locating object proposals from edges. 13th European Conference on Computer Vision: 391-405. https://doi.org/10.1007/978-3-319-10602-1_26

[12] Liu P, Zhang H, Eom KB. (2017). Active deep learning for classification of hy-perspectral images. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing: 1-13. https://doi.org/10.1109/JSTARS.2016.2598859

[13] Pan B, Shi Z, Xu X. (2017). R-VCANet: A new deep-learning-based hyperspectral image classification method. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing: 1-20. https://doi.org/10.1109/jstars.2017.2655516

[14] Yang J, Parikh D, Batra D. (2016). Joint unsupervised learning of deep repre-sentations and image clusters. IEEE Conference on Computer Vision and Pattern Recognition (CVPR): 5147-5156. https://doi.org/10.1109/CVPR.2016.556

[15] Klein AG, Messier P, Frost AL, Palzer D, Wood SL. (2016). Deep learning classification of photographic paper based on clustering by domain expert-s. 50th Asilomar Conference on Signals, Systems and Computers. https://doi.org/10.1109/ACSSC.2016.7869011

[16] Eastwood M, Jayne C. (2014). Dual deep neural network approach to matching data in different modes. 2014 International Joint Conference on Neural Networks (IJCNN): 1688-1694.

https://doi.org/10.1109/IJCNN.2014.6889877

[17] Ren S, He K, Girshick R, Sun J. (2015) Faster R-CNN: Towards real-time object detection with region proposal networks. 29th Annual Conference on Neural Information Processing Systems: 91-99. https://doi.org/10.1109/TPAMI.2016.2577031

[18] Girshick R. (2015). Fast r-cnn. Proceedings of the IEEE international conference on computer vision: 1440-1448. https://doi.org/10.1109/ICCV.2015.169

[19] Škrjanc I, Klančar G. (2017). A comparison of continuous and discrete tracking-error model-based predictive control for mobile robots. Robotics & Autonomous Systems 87: 177-187. https://doi.org/10.1016/j.robot.2016.09.016

[20] Bascetta L, Rocco P. (2010). Revising the robust-control design for rigid robot manipulators. IEEE Transactions on Robotics 26(1): 180-187. https://doi.org/10.1109/TRO.2009.2033957