

Coal and Gangue Separating Robot System Based on Computer Vision

Zhiyuan Sun ^{1,*}, Linlin Huang ¹ and Ruiqing Jia ²

¹ School of Mechanical Electronic & Information Engineering, China University of Mining & Technology, Beijing 100083, China; zqt1800402042g@student.cumtb.edu.cn

² Mine Robot Research Center, China University of Mining & Technology, Beijing 100083, China; 108207@cumtb.edu.cn

* Correspondence: 13522080222@163.com

Abstract: In coal production, the raw coal contains a large amount of gangue, which affects the quality of coal and pollutes the environment. Separating coal and gangue can improve coal quality, save energy, and reduce consumption and make rational use of resources. The separated gangue can also be reused. Robots with computer vision technology have become current research hotspots due to simple equipment, are efficient, and create no pollution to the environment. However, the difficulty in identifying coal and gangue is that the difference between coal and gangue is small, and the background and prospects are similar. In addition, due to the irregular shape of gangue, real-time grasping requirements make robot control difficult. This paper presents a coal and gangue separating robot system based on computer vision, proposes a convolutional neural network to extract the classification and location information, and designs a robot multi-objective motion planning algorithm. Through simulation and experimental verification, the accuracy of coal gangue identification reaches 98% under the condition of ensuring real-time performance. The average separating rate reaches 75% on low-, medium-, and high-speed moving conveyor belts, which meets the needs of actual projects. This method has important guiding significance in detection and separation of objects in complex scenes.

Keywords: coal and gangue; YOLO; detection; robot



Citation: Sun, Z.; Huang, L.; Jia, R. Coal and Gangue Separating Robot System Based on Computer Vision. *Sensors* **2021**, *21*, 1349. <https://doi.org/10.3390/s21041349>

Academic Editor: Oscar Reinoso Garcia

Received: 31 December 2020
Accepted: 11 February 2021
Published: 14 February 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In coal production, the raw coal mined directly from coal mines contains a lot of gangue, which contains a lot of heavy metals, and affects the calorific value of coal, causing serious pollution to the environment [1]. Therefore, it is essential to separate coal and gangue to make them play their respective roles [2,3]. As an important part of the coal industry, coal gangue separation can improve coal quality, save energy, and reduce consumption and thus rationally use resources [4]. At present, coal and gangue separation includes moving sieve jigging method, heavy medium separation, dual-energy γ -ray separation, X-ray, and image recognition method. The working principle of moving sieve jigging method in gangue sorting is as follows: sieve mechanism moves up and down in an aqueous medium by some kind of power-driven machine, raw coal moves with the media water, and then, coal and gangue forms stratified layers according to their different densities [5]. The principle of heavy-medium gangue separation is to use a liquid with a density between coal and gangue as the medium. If the density of coal is lower than the medium, it will float up, and the gangue will sink if the density is higher than that of the medium to achieve the purpose of separation [6]. However, moving sieve jigging method and heavy medium separation are mechanical methods and therefore, can only deal with gangue within a specific size range with huge equipment, high cost, and high energy consumption. Mechanical methods also cause environmental pollution in the process. The dual-energy γ -ray gangue sorting is composed of a low energy source and a middle

energy source, which form a radiation excitation field. When two rays pass through the raw coal at the same time, different attenuation amplitudes are generated, because the coal and gangue have different radiation absorption capabilities. The density is calculated by the attenuation amplitude to judge the classification [7]. Radioactive separating uses radioactive sources, which have strict requirements for radiation protection [8]. In addition, manual separation must cooperate with some above methods in the follow-up. Manual separating is widely used because it can be applied to every link of the sorting process, not only to identify but also to separate coal and gangue; although workers continue to perform mechanical repetitive work on the production line, the sorting efficiency does not even reach 50%, because there are many missed and wrong sorts.

In order to liberate workers from such dirty, dangerous, and dull (3D) work, a wave of research on industrial robots has been set off. The changes in external information are perceived through vision and fed back to the robot controller, so its application fields will be more extensive [9,10]. Shang proposes a Delta parallel coal gangue online sorting robot based on image recognition and an improved comprehensive calibration method, which avoids the influence of robot installation error on grasping accuracy [11]. Li proposes an image processing-based method and further develops a positioning and identification system [12]. It is worth mentioning that there is no robot for dynamic sorting of multi-target scenarios in previous work [13]. Vision technology can expand the scope of application of robotics, and the interaction between the two can improve the level of industrial production and promote social change [14]. In terms of coal and gangue recognition, people began to use traditional and deep learning methods to extract the characteristics of coal and gangue [15–17]. Traditional methods of manually extracting the gray, texture, and edge contour of coal and gangue have achieved certain results. Lepisto introduced a rock texture classification method, based on the textural and spectral features of the rock. Reddy separated gangue from coal based on histogram thresholding [18]. Zhang used energy, contrast, correlation, and entropy as the feature vector to realize the automatic identification of coal and gangue according to the features with large differences in texture feature values of coal and rock, and the entropy value had the best identification effect [19]. The above coal and gangue separating methods mainly rely on manual extraction of feature values with dependence on experience and knowledge, which are subjective and limited [20]. With the continuous development of deep learning in image recognition, more and more algorithms have begun to be applied to coal and gangue recognition [21,22]. The deep learning algorithm can avoid manual extraction of features and the deep learning algorithm relies on a large amount of data for training [23]. The deeper the convolution level, the more buried information can be found. Alfarzaei built a new model called (CGR-CNN) based on convolutional neural network (CNN) using thermal images as standard images for coal and gangue recognition [24]. Pu employed a convolutional neural network (CNN) to recognize coal and gangue images and help segregate coal and gangue [25]. However, the studies only identified the type of picture, and did not give the location of the object, so subsequent sorting was not possible. Target detection can be seen as a combination of image classification and positioning. This scene is more suitable for target detection of coal and gangue. Wang automatically extracts image features of coal and gangue to obtain feature maps, and then extracts candidate regions of coal or gangue through RPN (Region Proposal Network) [26]. Li designed CG-RPN to generate the target area of candidate objects, which can detect multiple coals and gangue in an irregular arrangement for simple background, and a relatively large distance between the target [9]. Many studies in the laboratory have achieved good accuracy in image recognition for only a single object (gangue or coal) in the image. In fact, the number and class of targets in the picture are uncertain, which makes the precise location of coal and gangue more complicated than general classification tasks. Meanwhile, the object (coal and gangue) is highly similar to the background (black belt), which greatly increases the difficulty of detection, and the detection of dynamic targets will also affect the accuracy. Even if the target is detected, due to the irregular shape of the objects, it is very difficult to extract and capture features [27,28].

The strategy of sorting as many objects as possible in the shortest possible time requires reasonable multi-target motion planning.

In this article, we propose a dynamic target detection method called CG-YOLO to identify the classification and location of coal and gangue. In this method, the target boxes in the images are clustered to form anchor boxes with different scales through the K-Means algorithm. Additionally, the improved Darknet network suitable for coal and gangue detection is designed for training images. It is confirmed that CG-YOLO have both real-time and accuracy in detection. Then, we design a six-axis robotic arm system, which is composed of the six-axis robotic arm body, control system, host computer, and stabilized power supply. In order to meet the separating strategy, motion planning is carried out according to the result of target detection, including the position, size, and confidence of coal and gangue. After completing the trajectory planning, the grasping task is assigned to the corresponding robotic arm controller. After the robotic arm acquires the task, it dynamically monitors the target according to the acquired task. When the target enters the robotic arm workspace, the robotic arm is driven to complete the gangue grasping. We build models of image acquisition devices, objects to be identified, conveyor belts, and robotic arms in simulation software CoppeliaSim, and through remote API, the system completes the identification and sorting process in the simulation environment. Then in real experiment, the robot is tested in high, medium, and low speeds targets. The main contributions of the work in this paper are the following: (1) propose a dynamic target detection method and establish a data set of coal and gangue, (2) propose a movement planning method in line with the coal gangue sorting strategy, and (3) control the robot to complete the sorting task.

The paper is structured as follows: Section 2 proposes the method of coal and gangue detection and motion planning of robot. Simulation and experiment and results are monitored and discussed in Section 3. Section 4 discusses the possibility and prospects of the application of artificial intelligence technology in the coal field, and Section 5 concludes this work.

2. The Proposed Method

In this section, we present the construction method of the data set, network and training of the target detection model, and the robot's motion planning strategy for multi-target tasks.

2.1. Sample and Data

The number and richness of samples have a significant impact on training results [26]. The images were taken by Delsa's G3-GM10-M2590 industrial Camera and VS1614-10 m lens. The samples included coal, gangue, and mix (both coal and gangue). Considering the objects at different speeds, images with the conveyor speed at 0–100 mm/s were collected. Images were collected at three typical speeds: low (10 mm/s), medium (50 mm/s), and high (100 mm/s). The sample set contained 2020 pictures. The distribution is shown in Table 1. It is worth mentioning that each picture contained multiple objects.

Table 1. Distribution of dataset.

	Coal	Gangue	Mix
Low speed	100	200	520
Medium speed	100	200	300
High speed	100	200	300

The collected images were manually annotated. There were two predefined classes, coal and gangue, and objects less than three-fourth were ignored when labeling. Each image corresponded to its annotation file, which contained the location and class information. Among the dataset, there were 5034 gangue objects and 3125 coal objects in total.

There were two types of objects in the established data set, which had big differences within the class. The objects in the input images were clustered by K-Means clustering algorithm; then, six anchors were acquired, which were (23, 41), (53, 30), (42, 82), (90, 83), (102, 115), and (164, 158). These anchor frames could achieve a larger IOU with the target on larger scale and smaller scale. For a target that was too small, i.e., cinder, it was ignored in the annotate process, so it was also not considered in the anchor frame clustering process.

2.2. Model Construction and Training

2.2.1. Network

CG-YOLO is based on the deep learning framework Darknet [29–31], which is used to extract features. The backbone network contains a total of 75 layers, of which 53 are convolutional layers and the rest are res layers. The characteristic interaction layer of YOLO network is 75–105 layers. In order to detect objects of size, the input image was down-sampled five times to get different scale. The result of the fifth down-sampling was outputted, and then, this result was up-sampled to integrate the result of the fourth down-sampling as the output of another scale. When the input image was 416×416 , the outputs were 13×13 and 26×26 through CNN. The network structure of CG-YOLO is shown in Figure 1.

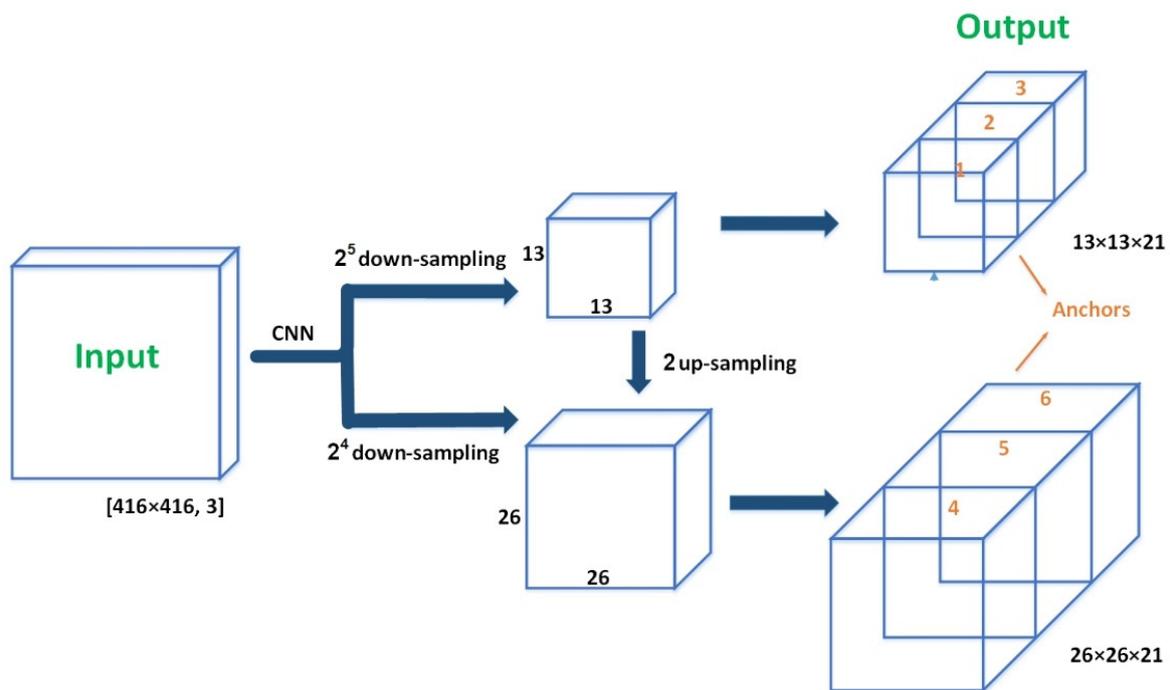


Figure 1. The network structure of CG-YOLO.

The detection object is found by the anchor boxes, and then, the position relative to the anchor box (t_x , t_y , b_h , and b_w) of each prediction box, where center coordinates t_x and t_y , box height b_h and width b_w , a target confidence level P_{obj} , and the class of object of the anchor are returned. The information C contained in each cell is as follows:

$$C = 3 \times (5 + N), \quad (1)$$

in which N is the number of classes. Each cell of CG-YOLO returns $C = 21$.

However, the anchor box does not match the object in the image, so it needs to be adjusted to become a prediction box, the process is shown in Figure 2. First, the initial position of the anchor box (c_x , c_y) is the position of the cell in 13×13 or 26×26 , which is adjusted by offset according to the returned (t_x , t_y). Then, the width and height of the

anchor (p_w, p_h) are adjusted according to the returned (t_w, t_h) . Finally, a prediction box $(b_x, b_y, b_w, \text{ and } b_h)$ is formed, and the formula is as follows:

$$\begin{aligned} b_x &= c_x + \sigma(t_x) \\ b_y &= c_y + \sigma(t_y) \\ b_w &= p_w e^{t_w} \\ b_h &= p_h e^{t_h} \end{aligned} \quad (2)$$

in Equation (2): $\sigma(x) = \frac{e^x}{1+e^x}$.

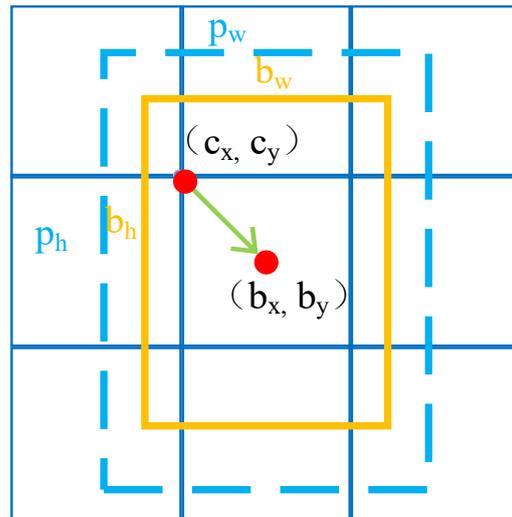


Figure 2. Position offset and width and height stretch.

2.2.2. Training Process

The learning goals of training are the offsets of $t_x, t_y, b_h,$ and $b_w,$ adjusting the prediction through the supervision information of the ground truth box. The ground truth box $(gt_x, gt_y, gt_w,$ and $gt_h)$ is artificially labeled, in which (gt_x, gt_y) is the actual center position and $(gt_w$ and $gt_h)$ is the width and height of the box, respectively. The cell where the center of object is located is responsible for detecting this object, and the anchor with the largest IoU matches this object, as shown in Figure 3. The calculation of supervision information $(t_x^*, t_y^*, t_w^*,$ and $t_h^*)$ is shown in the following formula:

$$\begin{aligned} \sigma(t_x^*) &= gt_x - c_x \\ \sigma(t_y^*) &= gt_y - c_y \\ t_w^* &= \log\left(\frac{gt_w}{p_w}\right) \\ t_h^* &= \log\left(\frac{gt_h}{p_h}\right) \end{aligned} \quad (3)$$

For the calculation of the loss, overall loss function is calculated into three parts. First classify the prediction frame, with the real frame falls on, without the real frame falls on, and the frame that does not match but with relatively large IoU. The first one is a positive example, we calculate the location loss, objectness loss, and classification loss. The second is a negative example, return the objectness of the prediction box expected to 0. The third

is a non-positive and non-negative example, and its loss function is not calculated. The overall loss function is as follows:

$$\begin{aligned}
 \text{Loss} = & \sum_{i=0}^{S^2} \sum_{j=0}^B \exists_{i,j}^{obj} [(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2] + \sum_{i=0}^{S^2} \sum_{j=0}^B \exists_{i,j}^{obj} [(w_i - \hat{w}_i)^2 + (h_i - \hat{h}_i)^2] \\
 & + \sum_{i=0}^{S^2} \sum_{j=0}^B \exists_{i,j}^{obj} (C_i - \hat{C}_i)^2 + \sum_{i=0}^{S^2} \sum_{j=0}^B \exists_{i,j}^{nobj} [(C_i - \hat{C}_i)^2] \\
 & + \sum_{i=0}^{S^2} \exists_i^{obj} \sum_{c \in \text{classes}} (p_i(c) - \hat{p}_i(c))^2
 \end{aligned} \quad (4)$$

where (x_i, y_i, w_i, h_i) is the prediction box, $(\hat{x}_i, \hat{y}_i, \hat{w}_i, \hat{h}_i)$ is the real box, the first line judges that the box j in the grid i is responsible for detecting this object and calculating the positioning loss, and the second line is the confidence loss, when there is a target in the grid, it returns to 1, and when it does not exist, it returns to 0. The last line calculates the classification loss.

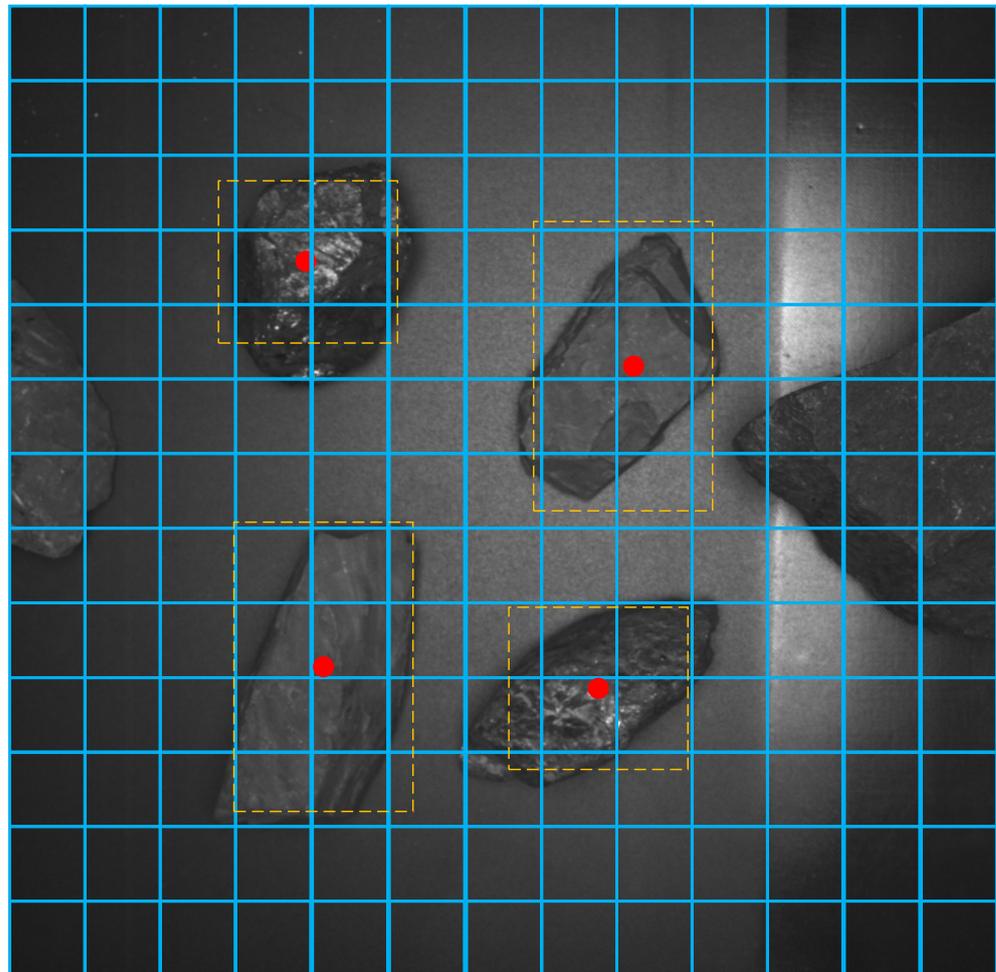


Figure 3. Matching and annotation of real frame.

2.3. Multi-Objective Motion Planning of Robot

After obtaining the location information of coal and gangue, if they are sent to the robot controller at the same time, it will cause confusion. However, if sorting in random order, the efficiency is difficult to be guaranteed. According to the CG-YOLO detection results, the targets are sorted by confidence, and the distance between the targets and the end of the robot should be calculated, respectively. The distances between the robot and the targets are different, and the particle size is also different. In addition, the confidence

levels returned are not the same. According to the sorting strategy, reasonable motion planning of the robotic arm is designed by separating as much gangue as possible in the shortest possible time. Assuming the robot end coordinates $(x_r, y_r, \text{ and } z_r)$, the distance between object n with the robot end coordinates is d_n , and the formula is as follows:

$$d_n = \sqrt{(x_r - x_o)^2 + (y_r - y_o)^2 + (z_r - z_o)^2} \quad (5)$$

In order to comprehensively consider the situation of each target, we propose a model of multi-objective motion planning of robot. In this model, the sorting weight guides the robot sorting. For the sorting weight, the distance between the target and the robot, the confidence of the target, and the size of the target will be included. The calculation of sorting weight is as follows:

$$W_n = d_n + n \times P_{objn} \times w_n \times h_n \quad (6)$$

In the formula, n is used to adjust the proportion of prediction box. When the value of n is larger, the sorting weight is more sensitive to prediction box. For example, there are three targets in the image, but the time is only enough for the robot to perform one movement, as shown in Figure 4.

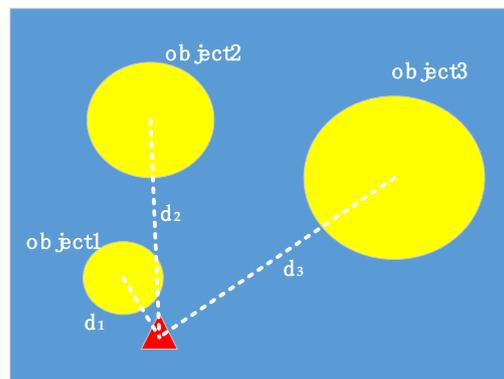


Figure 4. Distribution of objects and robots.

In Figure 4, the three yellow ellipses are the objects, the red triangle is the initial robot position. The ideal situation is to sort out the largest object (object 3), but if the distance of the largest target is relatively long and the robot trajectory is long, the time for the robot to move to object 3 may not be enough. If only consider sorting the nearest target, it is possible to pick the small one (object 1) and miss the big ones. However, by this model, the optimal solution can be acquired. The value of n should be related to the position and the execution speed of the robot and the speed of the conveyor. In conclusion, sorting weight is an important indicator of robot sorting, which determines the order of sorting, and therefore affects the efficiency of sorting.

3. Simulation and Experiment

3.1. Simulation

It is very complicated to build and debug the robot system, because the research and development cost is high and the cycle is long. When there are any problems in the basic links, the algorithm cannot be realized. In order to solve this problem, the entire sorting system is modeled in the robot simulator software CoppeliaSim, and the workspace of the robot is calculated, and the layout of the camera, conveyor belt, and robotic arm are adjusted through simulation [32].

The system is composed of target, image collector, conveyor belt, image recognition system, mechanical arm, and robot control system, as shown in Figure 5. In CoppeliaSim, randomly shaped coal and gangue are simplified into spherical shapes, distinguished by

different gray levels. Under this simplified model, the location of the target is identified. By setting the range of motion of each joint of the robot, three-dimensional workspace of it can be simulated. The objects are sorted in the workspace, but other parts, such as the camera, need to be placed outside the workspace to ensure safety. In a virtual environment, it is easier and intuitive to realize the adjustment of the target distribution, the speed of the transmission belt, and the movement of the hardware in the system in the space, which has guiding significance for the experiment.

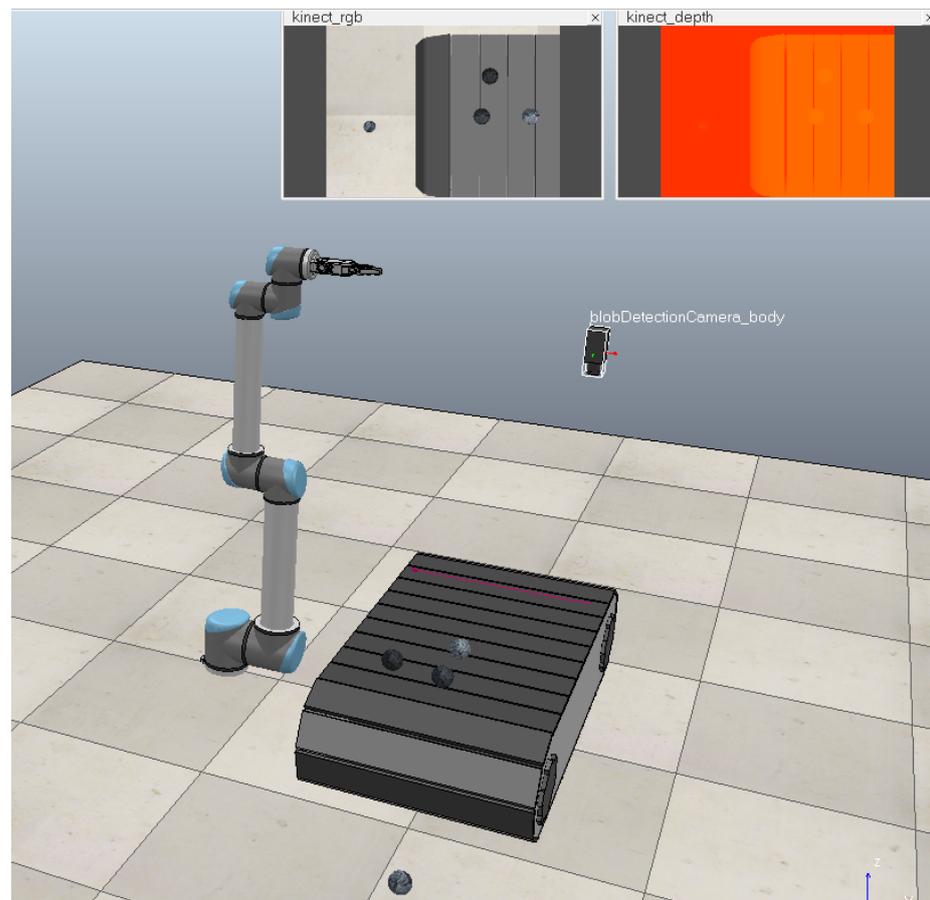


Figure 5. Simulation of coal gangue separating robot system.

In the simulation, due to the fast response speed of the robot, the communication time is negligible. When the conveyor speed is lower than 70 mm/s and the target number is less than or equal to 3, all objects sorting can be achieved.

3.2. Experiment and Result

Some tricky problems in the experiment were solved through simulation, including the position and height matching of conveyor belt, camera, robot body, etc.

In the target detection experiment, we divide the data set into a training set and a validation set at a ratio of 8:2. The training set contains 1616 images, including 3623 gangue objects and 2312 coal objects. After 4000 rounds of training, the loss is reduced to an acceptable range and the model is obtained, as shown in Figure 6. The trained model is tested by the valid images, and the AP of coal is 98.05%, the AP of gangue is 98.27%, and the mAP of this model is 98.16%. On the basis of already meeting the high recognition accuracy rate, the recognition speed meets the real-time requirements.

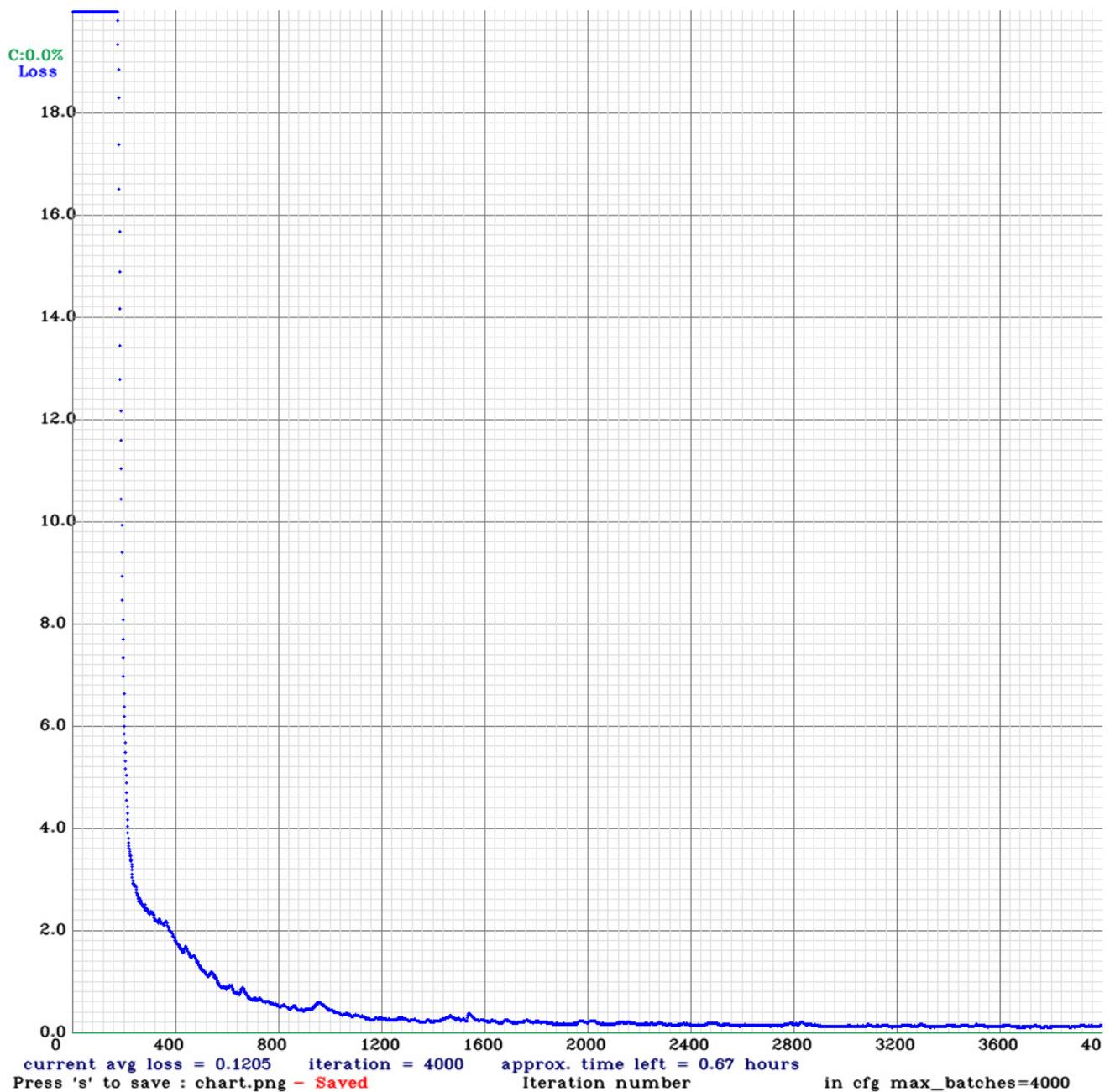


Figure 6. Target detection model training process.

The trained model is tested by the valid images (the validation set in the data set), and the AP of coal is 98.05%, and the AP of gangue is 98.27%, and the mAP of this model is 98.16%. On the basis of already meeting the high recognition accuracy rate, the recognition speed meets the real-time requirements.

In real-time coal and gangue sorting robot experiment based on target detection, we take the captured picture 2048×2048 scaling to the input size of the convolutional network 416×416 according to the aspect ratio as input. The frame coordinates obtained through the above network are $(t_x, t_y, b_h, \text{ and } b_w)$, i.e., the bounding box is relative to the location and size of the feature map and the confidence level. After Non-Maximum Suppression (NMS), predictions with high confidence are left. At the same time, the maximum number of prediction boxes is set to 20. The return information of the prediction box is enlarged to the original picture 2048×2048 , and the target in the picture is finally detected as shown in

Figure 7, and each is returned separately. The confidence levels of the target, the position, and the width and height of the prediction box are as shown in Table 2.

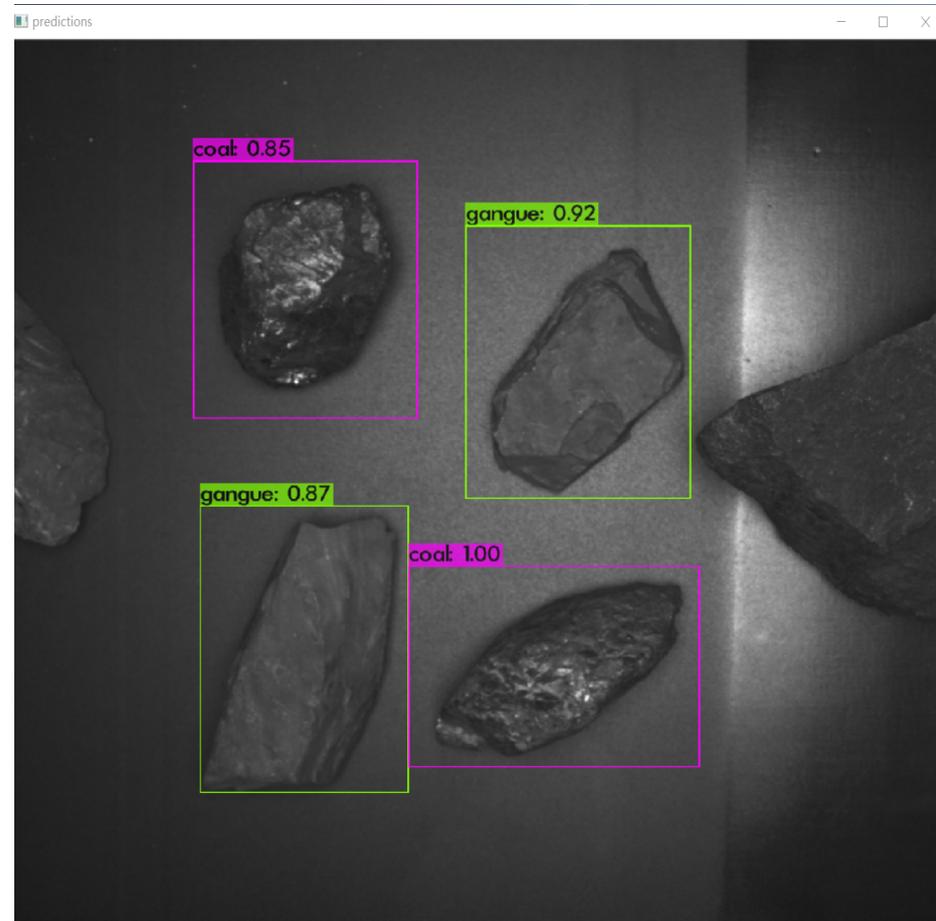


Figure 7. Detection of coal and gangue.

Table 2. Return value of coal and gangue detection.

Objects	Confidence	Horizontal Coordinate	Longitudinal Coordinates	Width	Height
gangue	0.924509	619.060181	265.919861	923.960083	654.644470
gangue	0.865843	260.949890	664.420837	542.986572	1072.973511
coal	0.854325	251.618530	173.845474	554.370850	539.189697
coal	0.998821	542.005005	750.845947	935.131104	1036.864746

Four targets are identified in the image, two coal pieces and two gangues. The scheme of sorting gangue out of the mixture is considered. The gangue identification results are sent to the multi-objective motion planning model for robot motion planning; then, the planned sorting sequence is sent to the robot controller through the serial port to control the robot, as shown in Figure 8. After many experiments, it is found that the number and position of targets in the field of view affect the separating efficiency. When the conveyor belt is at low (10 mm/s), medium (50 mm/s), and high (100 mm/s) speeds, the separating rate of robot using the motion planning algorithm proposed in this paper is shown in Table 3 below.

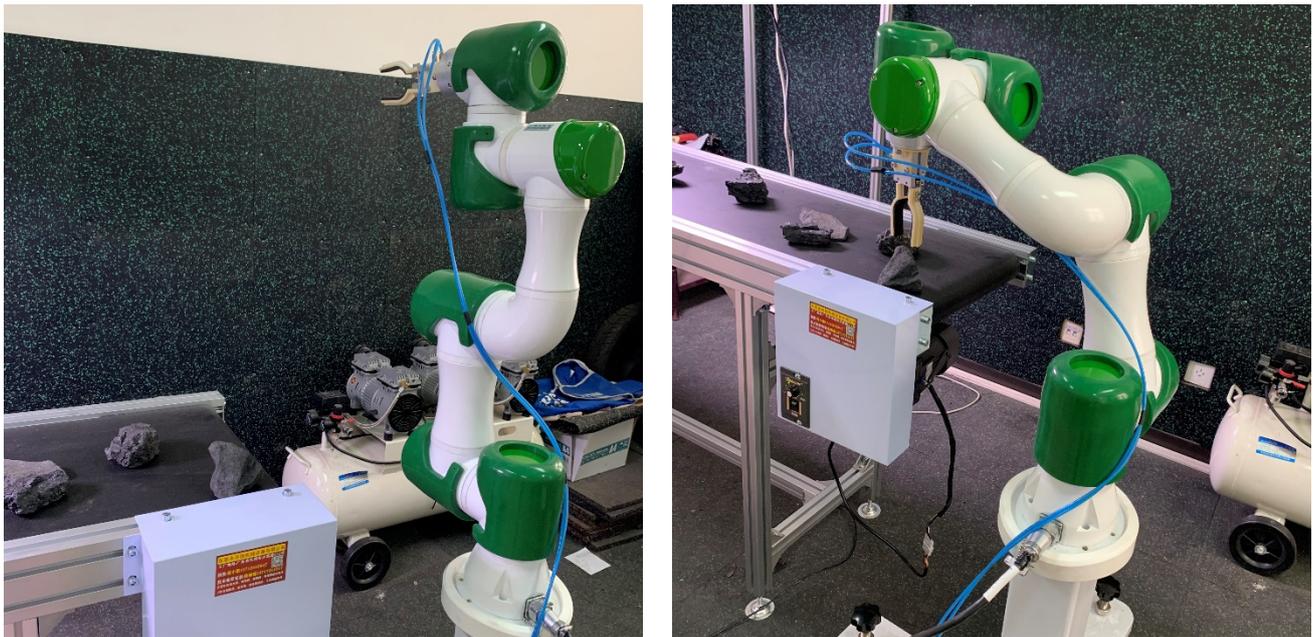


Figure 8. Robot sorting experiment based on multi-objective motion planning model.

Table 3. Comparison of sorting rate under motion planning.

Number of Objects	Low Speed	Medium Speed	High Speed
≤2 (with plan)	100%	100%	100%
≤2	100%	100%	100%
3–4 (with plan)	100%	65%	35%
3–4	100%	45%	20%
>4 (with plan)	72%	35%	35%
>4	60%	20%	20%

According to the table, when moving at a low speed and when there are a few sorting targets, whether to plan has little effect on the sorting effect. However, when the conveyor belt moves faster, reasonable planning can ensure that the most cost effective gangue is picked out in a limited time.

4. Discussion

In coal production, the separating of coal and gangue is a typical repetitive heavy manual labor. The application of artificial intelligence technology to replace manual labor can liberate workers from harsh working conditions and improve work efficiency. However, in this process, the identification of coal and gangue and the sorting of robots are both difficult. Many studies limit the recognition to the two-class classification of coal and gangue, but classification alone cannot achieve sorting. We proposed the application of computer vision technology to identify coal and gangue. In order to ensure real-time performance, we first train the dataset established by us through YOLO v3. The target detection effect obtained was not ideal. We aimed at the characteristics of coal and gangue and proposed CG-YOLO recognition algorithm; then, CG-YOLO has achieved a recognition accuracy of more than 98%. CG-YOLO has reference significance for target detection such as medical images with similar foreground and background and large differences within the class. Most of the research on coal and gangue sorting focuses on identification problems, but the follow-up still relies on manual separation. As an essential part of coal pre-combustion treatment, it is not enough to identify coal and gangue. A system of hand and eye coordination composed of vision technology and robotics technology can be formed to break the limitations of the application of a single technology, solve

problems in industrial production, and promote industry progress. The entire system is a comprehensive application of artificial intelligence technology and has strong portability in various coal mines. Through simple training, even workers who do not understand algorithms can monitor the operation of the system on the upper computer, and therefore workers do not need to work in dangerous and harsh environment, which can reduce casualties. At the same time, the robot applied to the specific scene can work 24 h a day, and the mistakes will be reduced, the work efficiency can be improved, and more benefits could be created for the enterprise.

5. Conclusions

In this paper, we presented a complete separating system, not just image recognition like most studies. A target detection algorithm that can recognize and locate coal and gangue is proposed, and this method can achieve high recognition accuracy. According to the detection results, we designed a robot motion planning model, so that the robot can sort out as many gangues as possible in a limited time. In this case, the sorting efficiency is improved. Finally, we do the robot grasping experiment, and get an average grasping rate of 75%, improving the accuracy and efficiency of coal and gangue separating. Although the types of coal and gangue in different regions are not the same and the composition and appearance are slightly different, the method proposed in this article has strong flexibility and versatility. With simple equipment, being environmentally friendly, and nonrequirement of extensive reconstruction of the workshop, this method could be transplanted to other coal mines to solve production problems by the parameter adjustment. Using vision-based robot systems to replace workers in coal production with harsh, dangerous, and repetitive tasks can improve sorting efficiency and ensure economic benefits. In view of the many casualties and safety hazards caused by coal mine production, the development of the coal industry in the future will inevitably develop in the direction of artificial intelligence and robots. The coal gangue robot sorting system based on computer vision proposed in this paper is an application of artificial intelligence and robot technology. The realization of unmanned or less humanized underground operations has important practical significance and is a frontier attempt.

In the next work, we will expand the dataset further, so that the model has better accuracy and robustness. For robots, we will improve the control algorithm from the perspective of dynamics and strive for robot execution speed.

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

Funding: This research was funded by the National Natural Science Foundation of China under Grants 51375481.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Dai, S.; Zheng, X.; Wang, X.; Finkelman, R.B.; Jiang, Y.; Ren, D.; Yan, X.; Zhou, Y. Stone coal in China: A review. *Int. Geol. Rev.* **2018**, *60*, 736–753. [[CrossRef](#)]
2. Chen, J.; Liu, G.; Kang, Y.; Wu, B.; Sun, R.; Zhou, C.; Wu, D. Coal utilization in China: Environmental impacts and human health. *Environ. Geochem. Health* **2014**, *36*, 735–753. [[CrossRef](#)] [[PubMed](#)]
3. Li, N.; Han, B.Q. Chinese research into utilisation of coal waste in ceramics, refractories and cements. *Adv. Appl. Ceram.* **2006**, *105*, 64–68. [[CrossRef](#)]

4. Li, J.; Yang, D.; Du, C. Evaluation of an Underground Separation Device of Coal and Gangue. *Int. J. Coal Prep. Util.* **2013**, *33*, 188–193. [[CrossRef](#)]
5. Zhang, Z.; Fu, X.; Wei, R.; Xu, Z. Fluid kinetics study on jig washer bed dilation and stratification. *J. China Coal Soc.* **2003**, *28*, 193–198.
6. Liu, X. Status quo and developing trend of China's Coal Preparation Technology. *Coal Prep. Technol.* **2018**, *6*, 12–15.
7. Zhang, P.; Kong, L.; Huang, X. Research on an on-line coal & gangue identification model based on dual energy gamma-ray transmission. *Ind. Instrum. Autom.* **2000**, *2*, 53–55.
8. Oikonomopoulos, I.K.; Perraki, M.; Tougiannidis, N.; Perraki, T.; Frey, M.J.; Antoniadis, P.; Ricken, W. A comparative study on structural differences of xylite and matrix lignite lithotypes by means of FT-IR, XRD, SEM and TGA analyses: An example from the Neogene Greek lignite deposits. *Int. J. Coal Geol.* **2013**, *115*, 1–12. [[CrossRef](#)]
9. Li, D.; Zhang, Z.; Xu, Z.; Xu, L.; Meng, G.; Li, Z.; Chen, S. An Image-Based Hierarchical Deep Learning Framework for Coal and Gangue Detection. *IEEE Access* **2019**, *7*, 184686–184699. [[CrossRef](#)]
10. Sun, Z.; Lu, W.; Xuan, P.; Li, H.; Zhang, S.; Niu, S.; Jia, R. Separation of gangue from coal based on supplementary texture by morphology. *Int. J. Coal Prep. Util.* **2019**, 1–17. [[CrossRef](#)]
11. Shang, D.; Wang, Y.; Yang, Z.; Wang, J.; Liu, Y. Study on Comprehensive Calibration and Image Sieving for Coal-Gangue Separation Parallel Robot. *Appl. Sci.* **2020**, *10*, 7059. [[CrossRef](#)]
12. Li, M.; Duan, Y.; He, X.; Yang, M. Image positioning and identification method and system for coal and gangue sorting robot. *Int. J. Coal Prep. Util.* **2020**. [[CrossRef](#)]
13. Gualtieri, M.; ten Pas, A.; Saenko, K.; Platt, R.; IEEE. High precision grasp pose detection in dense clutter. In Proceedings of the 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Daejeon, Korea, 9–14 October 2016; pp. 598–605.
14. Coccia, M.; Watts, J. A theory of the evolution of technology: Technological parasitism and the implications for innovation management. *J. Eng. Technol. Manag.* **2020**, *55*. [[CrossRef](#)]
15. Lv, Z.; Wang, W.; Xu, Z.; Zhang, K.; Lv, H. Cascade network for detection of coal and gangue in the production context. *Powder Technol.* **2021**, *377*, 361–371. [[CrossRef](#)]
16. Lei, S.; Xiao, X.; Zhang, M.; Dai, J. Visual classification method based on CNN for coal-gangue sorting robots. In Proceedings of the 5th International Conference on Automation, Control and Robotics Engineering, CACRE 2020, Dalian, China, 19–20 September 2020; pp. 543–547.
17. Wang, W.; Zhang, C. Separating coal and gangue using three-dimensional laser scanning. *Int. J. Miner. Process.* **2017**, *169*, 79–84. [[CrossRef](#)]
18. Lepisto, L.; Kunttu, I.; Autio, J.; Visa, A. Rock image classification using non-homogenous textures and spectral imaging. In Proceedings of the WSCG '2003: Short Papers: The 11-th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision 2003, Plzen, Czech Republic, 3–7 February 2003; pp. 82–86.
19. Zhang, W.; Wang, Z. Characteristic Analysis and Recognition of Coal and Rock Based on Visual Technology. *Coal Technol.* **2014**, *33*, 272–274.
20. Su, L.; Cao, X.; Ma, H.; Li, Y. Research on Coal Gangue Identification by Using Convolutional Neural Network. In Proceedings of the 2nd IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference, IMCEC 2018, Xi'an, China, 25–27 May 2018; pp. 810–814.
21. McCoy, J.T.; Auret, L. Machine learning applications in minerals processing: A review. *Miner. Eng.* **2019**, *132*, 95–109. [[CrossRef](#)]
22. Hou, W. Identification of Coal and Gangue by Feed-forward Neural Network Based on Data Analysis. *Int. J. Coal Prep. Util.* **2019**, *39*, 33–43. [[CrossRef](#)]
23. Coccia, M. Deep learning technology for improving cancer care in society: New directions in cancer imaging driven by artificial intelligence. *Technol. Soc.* **2020**, *60*. [[CrossRef](#)]
24. Alfarzaei, M.S.; Niu, Q.; Zhao, J.; Eshaq, R.M.A.; Hu, E. Coal/Gangue Recognition Using Convolutional Neural Networks and Thermal Images. *IEEE Access* **2020**, *8*, 76780–76789. [[CrossRef](#)]
25. Pu, Y.; Apel, D.B.; Szmigiel, A.; Chen, J. Image Recognition of Coal and Coal Gangue Using a Convolutional Neural Network and Transfer Learning. *Energies* **2019**, *12*, 1735. [[CrossRef](#)]
26. Wang, P.; Cao, X.; Xia, J.; Wu, X.; Ma, H. Research on multi-manipulator coal and gangue sorting robot system based on machine vision. *Ind. Min. Autom.* **2019**, *45*, 47–53.
27. Surapong, N.; Mitsantisuk, C. Position and Force Control of the SCARA Robot based on Disturbance Observer. In Proceedings of the 2016 International Electrical Engineering Congress, IEECON2016, Chiang Mai, Thailand, 2–4 March 2016; pp. 116–119. [[CrossRef](#)]
28. Zhang, W.; Mei, J.; Ding, Y. Design and Development of a High Speed Sorting System Based on Machine Vision Guiding. In Proceedings of the International Conference on Solid State Devices and Materials Science, Kyoto, Japan, 25–27 September 2012; pp. 1955–1965.
29. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the 29th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 779–788.

-
30. Redmon, J.; Farhadi, A. YOLO9000: Better, faster, stronger. In Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2017), Honolulu, HI, USA, 21–26 July 2017; pp. 6517–6525.
 31. Redmon, J.; Farhadi, A. YOLOv3: An Incremental Improvement. *arXiv* **2018**, arXiv:1804.02767.
 32. Liu, J. Research on Robot Grasping Simulation Training Technology Based on Deep Learning. Master's Thesis, Harbin Institute of Technology, Harbin, China, 2018.