BİLİMSEL MADENCİLİK DERGİSİ
SCIENTIFIC MINING JOURNAL
TMMOB Maden Mühendisleri Odası Yayını / The Publication of the Chamber of Mining Engineers of Turkey
www.mining.org.tr

Research Article

Intelligent Supervision System of Ore Pass Using Improved YOLO v3

Baoshun Liu[a,*], Yanyu Song[b], Yongjing Ye[a], Zijing Zhang[a]

[a] School of Civil and Resource Engineering, University of Science and Technology Beijing, Beijing, 100083, China
[b] China ENFI Engineering Corporation Changsha Branch, China

Received: 24 August 2023 • Accepted: 14 November 2023

A B S T R A C T

To prevent safety accidents caused by equipment and personnel entering the ore pass by mistake, it is necessary to arrange ore pass control personnel underground the mine. The control personnel in the ore pass are extremely vulnerable to dust and noise during the unloading process of mining vehicles, and there is an urgent need for an intelligent access control system to reduce safety accidents. This article establishes a training set for object detection of the YOLO v3 model based on images of mining vehicles in underground monitoring videos. Through optimizing the training process and algorithm of the YOLO v3 model, and adopting a dual camera collaborative discrimination method, the influence of brightness on the recognition results when mining vehicles are turned on is overcome. In this way, the mining vehicles entering and exiting the pass operation area are accurately identified from the underground monitoring videos. Then, an intelligent access control system for controlling the orepass door is developed based on the Jetson Nano embedded program. The research results show that the average accuracy of the system in identifying target vehicles is greater than 95%.It can rotate the ore pass door by 90 ° in 3 seconds, achieving intelligent control of the mine orepass and promoting the construction of smart mines.

Keywords: *Yolov3, ore pass, smart access control, object detection, recognition of underground mine vehicle.*

Introduction

As a key link in the development and production transportation of mineral resources, mine pass transportation and its system cannot only save transportation equipment, but also improve the convenience of mine production management. However, the transportation of mine passes faces problems such as low underground visibility and multiple blind spots, which can lead to accidents caused by collisions between mobile equipment and personnel. Among them, substations, ore passes, and explosive magazines are areas that require risk control (Imam et al., 2023). To reduce safety risks, some mines have designated personnel on duty in the above-mentioned areas. Although this method is simple and feasible, with the extension of underground tunnels and the widespread application of large-scale mechanized equipment in the underground, the areas that need to be monitored are also constantly expanding. If we continue to use this method, it will significantly increase labor costs but also not comply with the concept of green and intelligent mining construction. On the other hand, the narrow underground space, dust, and noise can also harm the health of personnel on duty (Madahana et al., 2020). Faced with this issue, Guo and Li (2011) used positioning technologies such as RFID, GIS, and ZigBee wireless communication to control the movement of underground personnel in designated areas, which effectively reduced the safety and health risks of underground personnel. Yang et al. (2016) established the Gaussian Mixture Model (GMM) and combined it with surveillance videos to identify underground personnel, providing an early warning mechanism for the safety of underground personnel. They have made beneficial explorations to protect the safety of underground personnel on duty. However, to fundamentally solve the safety problem of mine pass transportation, it is necessary to continue to conduct in-depth research towards unmanned and intelligent directions.

In recent years, the continuous development and maturity of artificial intelligence have injected new vitality into the development of various industries. Object detection is the localization and classification of a variable number of targets in an image, ultimately obtaining the categories and positions of multiple targets in the image. It is one of the important branches in computer image vision. Target detection algorithms can be divided into selec-

* Corresponding author: liubaoshun@ustb.edu.cn • https://orcid.org/

https://doi.org/10.30797/madencilik.1349081

tion-based algorithm models such as R-CNN (Girshick et al., 2014), fast R-CNN (Ren et al., 2017), and regression-based algorithm models such as SSD (Liu et al., 2016) and YOLO (Redmon and Farhadi, 2016).The object detection model represented by YOLO (You Only Look Once) has been widely applied in facial recognition, autonomous driving, agricultural production, and medical image analysis. After Yolov1 was first proposed by Redmon and Farhadi (2016), Kumar et al. (2022) proposed ETL-YOLO v4 model, which added Mosaic and CutMix data enhancement functions during training, significantly improving the recognition accuracy of face wearing masks during the COVID-19 period. Kang et al. (2024) developed a YOLO-FA model for vehicle detection in autonomous driving using the A type-1 fuzzy attention (T1FA) mechanism, which effectively improved the accuracy of vehicle detection in rainy and nighttime scenarios with high uncertainty. Tian et al. (2023) proposed an MD-YOLO model for detecting small target pests and successfully deployed it in pest warning software, promoting the improvement of pest warning mechanisms in agricultural production. Baccouche et al. (2022) combined the deep learning model with YOLO algorithm for early mammograms diagnosis, which significantly improved the diagnostic accuracy of early breast cancer. The above researches prove the excellent performance of image detection algorithms represented by YOLO and can effectively solve relevant practical problems, which also provides the possibility for their application in mine pass transportation.

In the current underground location positioning system of mines, intelligent access control systems are used for controlling risk areas. The system can automatically identify the moving targets entering the risk area in real-time. Once unauthorized personnel or equipment enter, the system will automatically control the access control switch to ensure strict control of high-risk areas. This system cannot only meet the basic requirements of entry and exit management, but also integrates functions such as early warning and alarm, which can be used to timely remind operators or management personnel to pay attention to potential risks and take necessary response measures. In addition, the data statistics and analysis function in the system can provide rich and detailed operation data for management personnel, which is helpful for optimizing mining production strategies and decision-making. However, the underground lighting in mines is dim and the dust concentration is high. If the images collected in the mine are directly used for object detection without preprocessing, there is a problem of low image recognition rate. Inefficient image recognition may not only miss out on real risk events, but may also generate false alarms, affecting the normal operation of mining production.

This article proposes a network training method and algorithm for improving the benchmark YOLO v3 model to address the aforementioned issues in the identification of underground mining cars. By using a dual camera collaborative discrimination method, the mining cars in the pass operation area are identified and the pass door is automatically controlled to improve the safety of the pass operation, providing a new tool for the stable, efficient, and safe operation of the mine pass transportation system.

1. Current situation and existing problems in the control of mine pass

The second mining area is the largest under a certain group, and the development method of the mining area is vertical shaft inclined slope level roadway. The diameter of the ore pass in the mining area is generally 4m, with a depth of 100m-150m.There have been many accidents of construction personnel and equipment falling into the ore pass. The mining area adopts methods such as ore pass doors, covers or special personnel to ensure the safety of personnel and equipment in the operation area. The ore pass doors are generally closed (Fig.1), and that used frequently are manned. When the loader unloads ore (waste rock) in and out of the ore pass, the watchman opens or closes the door. Some less commonly used ore passes require the scraper driver to open or close the door. The cover is opened (Fig.2) or closed (Fig.3) by the hoist using an electric remote control by the watchman or the loader driver.

The shortcomings of the above-mentioned manual control method are: (1) The dust concentration is large during unloading ore or rock, which damages the health of the personnel on duty; (2) It is inconvenient to count personnel and vehicles, and it is easy to have regulatory errors and omissions, and difficult to trace the whole process of the accident once an accident occurs. Considering the underground working environment, the Yolov3 algorithm with fast monitoring speed, low background false detection rate and strong versatility is selected. Computer vision technology is used to automatically identify underground mine vehicle entering and exiting the ore pass. An intelligent ore pass access control system composed of Jetson Nano embedded platform, automatic door opener, and electromagnetic lock is built.

Fig.1 Ore pass door in the mining area

Fig.2 Opened ore pass cover

Fig.3 Closed ore pass cover

2. Technical plan for underground ore pass door

2.1 The principle of YOLO v3

YOLO v3 is the third version of the YOLO model, proposed by Redmon and Farhadi in 2018. It uses Darknet-53 as its backbone network, while using Batch Normalization to accelerate training speed, and increase residual connections to solve gradient vanishing. It also uses cross entropy loss function and dynamic weight decay algorithm to optimize the training effect of the neural network. Compared with YOLO v1 and YOLO v2, YOLO v3 achieves higher detection accuracy while maintaining speed. YOLO v3 uses 9 predefined anchor boxes, divided into 3 groups of 3 in each group. Each group of anchor boxes corresponds to a feature map of a scale to adapt to targets of different sizes, which significantly improves the detection ability for small objects. In previous versions of YOLO, each object could only be assigned to one category, while YOLO v3 allowed an object to have multiple categories simultaneously, achieving multi-label classification. In addition, the loss function of YOLO v3 comprehensively considers the errors of coordinate prediction, target confidence, and category prediction, further improving its detection accuracy.

YOLO v3 not only inherits the real-time and efficient performance of YOLO series models, but also further optimizes accuracy and stability. Through deep network structure, three scale detection, and new technological improvements, YOLO v3 has been widely applied and recognized in object detection.

2.2 Modeling process

2.2.1 Building a dataset for underground mining vehicles

Local features such as license plates and global features such as the outline and color of the vehicle can be used for visual recognition of mining vehicles. The license plates of underground mining vehicles are easily affected by factors such as oil pollution, dust, and sudden changes in light, resulting in the inability to accurately extract license plate information, thereby affecting the recognition effect. Due to the current lack of underground mine vehicle datasets, it is necessary to construct the datasets of four underground mine vehicles, including scraper, 20t tunnel transport vehicle, 25t A-type mine vehicle, and 25t B-type mine vehicle. The video of the underground network surveillance camera in the second mining area is collected. The images of four types of trucks under different lighting environments, different shooting conditions and different driving speeds are obtained by framing after the video is edited. After data enhancement such as rotation, cropping, brightness adjustment, and distortion of these images, 2224 photos of each of the four types of underground mine vehicle are obtained. Then the LabelImg software is used to manually label the underground mine vehicle features on each vehicle's image to generate a underground mine vehicle dataset. To facilitate the model training and the judgment logic of opening and closing doors, each type of underground mine vehicle is subdivided into eight categories according to the method marked in Table 1.The label information includes the vehicle class name and coordinates in the image, and stores it in the format of a standard VOC dataset. Figure 4 is the underground mine vehicle dataset after completing the front label of the-vehicle.

Fig.4 Underground mine vehicles dataset after labeling

Table 1. Underground mine vehicle labeled in the dataset

Vehicle category	Detection classification	Label name
Loader	Head	LHD_1
	Trail	LHD_2
20t tunnel transport vehicles	Head	20t Truck_1
	Trail	20t Truck_2
25t A-type minevehicle	Head	25t Truck_1
	Trail	25t Truck_2
25t B-type minevehicle	Head	Steyr Truck_1
	Trail	Steyr Truck_2

2.2.2 Improved benchmark Yolov3 algorithm and training strategy

The detection effect of using the benchmark Yolov3 algorithm model to construct an intelligent unmanned control system for the ore pass does not meet expectations. Occasionally, underground mining vehicles may not be recognized or classified incorrectly, mainly due to the following reasons:

(1) The image data of the underground vehicle dataset constructed in the second mining area is insufficient. It is relatively cumbersome to screen out vehicle videos in a large number of video data. The images of underground mine vehicle are mainly taken from the webcam video of the 1170 Maocang rest chamber and the Maocang ore pass. The collected underground mine vehicle images are relatively limited, which reduces the generalization of underground mine vehicle features, thereby affecting the effect of target recognition.

(2) The benchmark Yolov3 algorithm is difficult to adapt to the complex underground operating environment. The recognition effect of underground mining vehicles is optimal when driving smoothly and parking. However, when the vehicle is close to a certain range of the underground surveillance camera, due to the weak underground light, the program is easy to fail to obtain the underground mine vehicle or the underground mine vehicle is blurred under the strong light of the headlights, thus affecting the detection accuracy of vehicle identification.

To build an efficient and reliable intelligent unmanned control system for ore pass, it is necessary to improve the training strategy and optimize the algorithm of the model to better adapt to the underground environment according to the above problems.

1) Spatial pyramid pooling structure

Convolutional neural network is mainly composed of convolutional layer and fully connected layer, where the parameters of the fully connected layer are related to the input size. Only using a fixed input size can the fully connected layer parameters in the training be determined. Therefore, the Spatial Pyramid Pool (SPP) structure is proposed. Any scale feature input into the structure, after its pooling kernel processing, can output features of the same scale, thus achieving mutual fusion of features of different scales. This article inserts an SPP module into the Convolutional Set of the Yolov3 algorithm model, and the modified Convolutional Set module is shown in Fig. 5.The SPP module consists of four parallel branches, except for the first branch that outputs the results directly from the convolutional set: the second branch is the maximum pooling of 5×5 pooling kernels; the third branch is the maximum pooling of pooling kernels of 9×9, and the fourth branch is the maximum pooling of pooling kernels of 13×13.That is, after the SPP module processing, the original input channel will be expanded by 4 times. For example, the input of the convolution set in the SPP module is 16×16×512, and after Concatenate fuses the channels, the output is expanded to 16×16×2048.

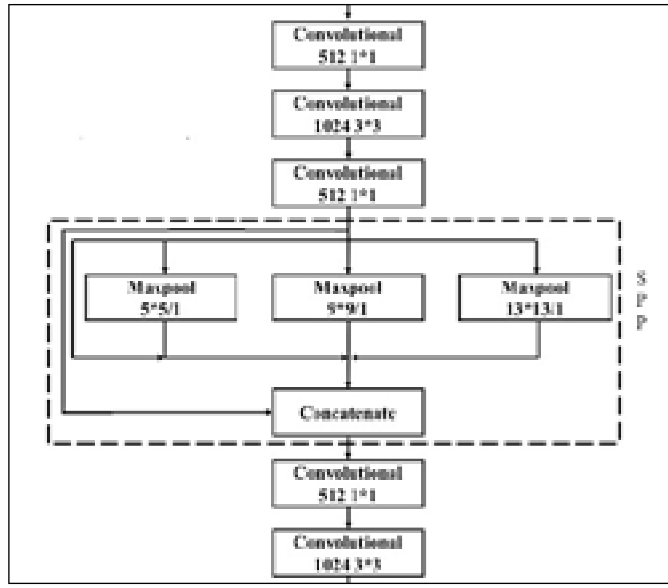

Fig. 5 Insert the SPP module in Yolov3

2) Use K-means++ clustering

Yolov3 uses the K-means clustering algorithm to calculate the anchor box. The anchor box is obtained by cluster regression based on the size data of the detection target during training using the following steps:

① The algorithm model randomly selects several points in the dataset as the cluster center.

② The Euclidean distance formula (1) is used to calculate the distance from all data points to the coordinates of each cluster center separately. Then regression calculations are performed based on the calculated distance, and the data point categories are divided by the number of cluster centers.

$$d(x,y) = \sqrt{(x_1-y_1)^2+(x_2-y_2)^2+\ldots+(x_n-y_n)^2} = \sqrt{\sum_{i=1}^{n}(x_i-y_i)^2} \quad (1)$$

③ The distance from the clustered data point to the center of each cluster is recalculated. The mean of the data is then calculated, and the cluster centers of each class are re-divided according to this mean.

④Repeat steps ②and ③until the resulting central positions of each type of cluster remain constant or the sum of squared errors is minimized.

It can be seen from the K-means clustering algorithm that the core of the algorithm is to select the initial clustering center. When the cluster center selection is not suitable (such as the distribution of each cluster center point is concentrated), the continuously calculated results may cause the algorithm model to fall into the local optimum, so that the expected clustering results cannot be obtained. To avoid this situation, the K-means++ algorithm is used to optimize all cluster centers randomly selected at one time in K-means, and the steps are as follows:

① A random point is selected in the dataset and set as the initial cluster center point.

② The shortest distance of each data point from the coordinates of the above cluster center point is calculated, and the calculation results are stored in an array. That is, the array represents the shortest distance from each data point to the center of all clusters.

③ All the elements in the above array are added to obtain a distance sum. Randomly taking a value within the total distance range to iterate through all elements in the array, a cluster center point is obtained by calculation.

④ The shortest distance between each data point and the coordinates of the two cluster center points is calculated, and the smallest one is selected and stored in the array.

⑤ Repeat steps ② and ③ to calculate the desired cluster center point.

⑥ The cluster center point obtained above is used to replace the initial cluster center randomly selected by the K-means clustering algorithm to perform cluster calculation and reduce the influence of the initial point on the algorithm. At the same time, a new distance formula (2) is defined to replace the original Euclidean distance calculation formula. The distance () calculated using this formula is inversely proportional to the interaction ratio. The larger the image interaction ratio (IOU), and the smaller the distance, which is more likely to be classified into the same category.

$$d(box, centroid) = 1 - IOU(box, centroid) \tag{2}$$

3) UsingMosaic data augmentation

Mosaic data augmentation is a method to enhance the breadth of deep learning data and improve the stability of models. It allows the model to learn to use an unlimited amount of data without introducing new real data. At the same time, the markup of the dataset and the comprehensibility of semantic information are increased, which can improve the generalization ability of the model. The Mosaic data enhancement method is to merge multiple different types of images, randomly select one of the images as the main reference, and then randomly stitch the remaining images near the main image after image transformation operations such as cropping, rotating, and brightness adjustment to form a larger Mosaic image.

4) Activation function

The activation function used by Yolov3 is Leaky ReLU, as shown in Eq.3.

$$Leaky\,ReLU(x) = \begin{Bmatrix} x & x > 0 \\ \alpha x & x \le 0 \end{Bmatrix} \tag{3}$$

To further expand the scope of functions, accelerate the training speed of the algorithm and enhance the robustness of the model, this paper uses the Mish activation function (Equation 4) to replace the Leaky ReLU function.

$$Mish(x) = x tanh(ln(1+e^{x})) \tag{4}$$

The Mish function maintains better network gradients than the Leaky ReLU function. The Mish curve is smooth, and its training results have optimal generalization, and are easy to update a large number of neurons during training. Although the Mish function is more complex and will occupy more computing resources, it will obtain better calculation results. The mean Average Precision (mAP) of the ReLU activation function and the Mish activation function are 91.62% and 92.83%, respectively. It can be seen that the Mish activation function has stronger performance, with an improvement of approximately 1.3%.

5) Training and optimization of models

Three sets of models: Yolov3-416, Yolov3-tiny-416, and improved Yolov3-416, are compared. Yolov3-tiny is a streamlined version of Yolov3, which is more lightweight and has lower requirements for graphics and memory, making it easier to apply in embedded systems. Its fast detection speed makes it easier to achieve real-time detection. However, Yolov3-tiny is essentially based on Yolov3, which removes one-third of the non-redundant intermediate layers through pruning compression, which can easily cause missed and false positives. Therefore, Yolov3 has an advantage in accuracy. The value 416 after the three sets of model names represents the image input_shape. The Yolov3 algorithm can uniformly transform the size of the original image into an input_shape that can be set independently, and then send the image to the training network after the unified size. Setting different input sizes can obtain model weight files with different detection effects. Generally, the larger the input size, the higher the accuracy, but the higher the requirements for memory and other aspects. The commonly used input size is 416×416.

During the training process, the dataset is distributed according to train set: test set: validation set (val) = 6:2:2.The weight model is set to be saved every 3 iterations, and the validation set loss value of each iteration val_loss is recorded. When there is a failure of 10 consecutive iterations to be less than the minimum value of the recorded val_loss, stop training, and the val_loss minimal weight model is resumed as the final training output model. The image dataset, model parameter settings, and training environment are the same when the three algorithm models are trained. The loss rates of the three model algorithms are shown in Fig. 6, and the number of network layers and mAP are shown in Table 2.

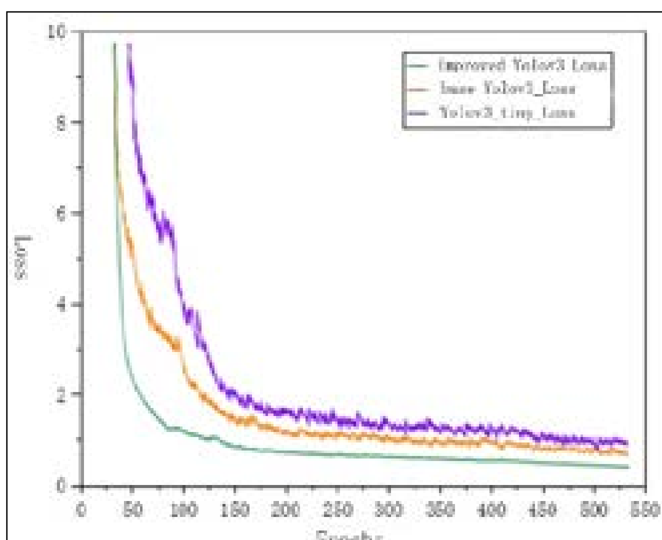

Fig. 6 Loss value change curve of the training process of the three algorithm models

Table 2. Network result parameters and test results of different algorithm models

Algorithmic models	The number of network layers	mAP/%	F1-Score/%	Time/ms
Yolov3-416	106	92.59	93	28.55
Yolov3-tiny-416	23	89.49	90	3.26
Improved Yolov3-416	80	98.88	98	24.27

It can be seen that compared with the benchmark Yolov3, Yolov3-tiny has fewer network layers and the lowest model complexity. It can significantly reduce the detection time, with an average accuracy reduction of 3.1%, and a reduction of 3% in F1 score. According to the training results of benchmark Yolov3, the improved Yolov3 model converges faster and is relatively more stable. After completing 500 iterations, the loss function tends to smooth. The average accuracy and detection time have been improved while the number of network layers has decreased. It can be seen that among the three Yolov3 models, the improved Yolov3 algorithm model can better adapt to the underground operation environment of the second mining area.

3. Design and implementation of intelligent ore pass access control

Considering the practical application conditions in underground mines, an intelligent ore pass access control system based on Jetson Nano embedded development (Fig.7) is designed. The hardware platform of this system mainly consists of three parts: vehicle video capture, Jetson Nano central processing unit, and access control system.

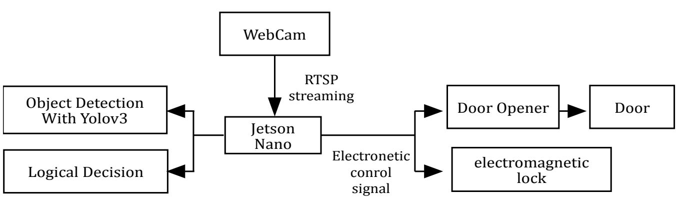

Fig.7 Intelligent ore pass access control system

3.1 Central processing unit

Due to the special underground environment, the use of embedded development platform as the central processing unit of the access control system cannot only meet the integration of software and hardware, achieve the effect of image recognition without networking, but also have the advantages of small size and high stability. It is easy to install near the door, and the system uses Jetson Nano (Fig.8) as the embedded processor.

3.2 Video capture

The Hikvision all-in-one network cable is connected to the corresponding interface of Jetson Nano, and the IP address, subnet mask and gateway where the camera is located are configured. When the vehicle enters the camera's field of view, OpenCV programming is used to read Hikvision's standard RTSP push format to obtain surveillance images.

3. 3 Access control system

Access control systems include automatic door opener, electromagnetic lock and door. The automatic door opener uses the motor to drive the door to rotate, and the opener controls the motor forward or reverse according to the signal sent by the processor to realize the automatic opening and closing of the door. The body of the selected brushless floor opener (as shown in Fig. 9) has two 400mm long swing arms that can drive a door up to 150kg. The control logic of the automatic opener is to short-circuit the GND and N1 interfaces in the motherboard to open the door, and short-connect the GND to N2 to close the door. The GPIO pin of the Jetson Nano can output high and low levels according to the Yolov3 image recognition results. This signal is passed to the relay shown in Fig. 10 that can be triggered by high and low levels as a switch controlled by an automatic gate signal. The relay controls the status of the electromagnet through electrical signals, which in turn controls the switching at the output. To correctly judge the opening and closing state of the door, the opener needs to be equipped with an electromagnetic lock to receive the electronic control signal of the opening and closing door transmitted by the opener, so as to control whether the electromagnetic lock is energized. The electromagnetic lock selected is a 12V-180kg tensile magnetic lock shown in Fig.11, which has a normally open and closed signal feedback function and can be connected to the main control board of the access control system. According to the net section of the roadway, the guardrail net with an external dimension of 4.5m ×4.8m is selected as the door.

Fig.8 Jetson Nano

Fig.9 Door opener

Fig.10 Relay

Fig.11 Electromagnetic lock

3.4 The judgment logic of smart ore pass access control

The default state of the door is set to be closed. After the detection code runs, the screenshot of each frame of the surveillance video is continuously detected. If 5 consecutive pictures of the loader head are recognized, the Jetson Nano outputs a opening signal; if five consecutive images of the scraper tail are identified, the Jetson Nano outputs a closing signal. When the light of the underground mine vehicle is too bright and the blurred recognition rate of the front image is low, the rear of the car is used to assist the recognition.

4. Results and discussion

4.1 Underground implementation results of intelligent ore passdoor

After the installation of Jetson Nano embedded development components, door openers, door locks and other hardware devices, network cables are used to connect the Jetson processor with the camera. When the vehicle enters the camera's field of view, OpenCV is used to return the vehicle image. The Yolov3 algorithm model is called for vehicle recognition classification, and the electronic control signal is output to the access control processor to control the brushless motor and electromagnetic lock. After the system is running, the identification and classification effect of the four types of mining vehicles in the state of constant speed driving and parking is optimal (Fig.12). After installing dual-camera, the recognition accuracy of the four types of underground mine vehicle is 95%-100%. According to the algorithm logic, the target vehicle is identified 5 times in a row, and the program will output the change of the corresponding variable value in the console. It can be seen that the program detects vehicles with rapid response and low latency. Jetson outputs the electronic control signal to the automatic door opener and electromagnetic lock for opening(closing), and the door rotates 90° in about 3s, meeting the needs of the system.

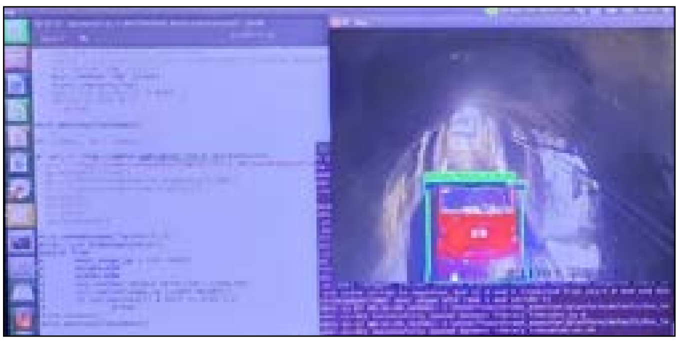

Fig. 12 Target vehicle detection screen

4.2 Problems and solutions encountered during system operation

4.2.1 Error message of memory overflow during Jetson object detection process

When Jetson uses RTSP to receive the live view of the webcam, the console displays an error message of memory overflow. The object detection program is written using OpenCV to call RTSP to pass in video frames, and the core code is as follows:

```
vid =cv2.VideoCapture(url)
success,img = vid.read()
cv2.imshow("object detection",img)
```

img is the variable of the incoming image frame by the camera. After analysis, when OpenCV is used to obtain webcam data based on RTSP protocol, the FFMPEG framework is used to process video stream data by default. However, when h264 is selected based on RTSP call to obtain webcam data format, FFMPEG cannot call libx264 encoder based on GPL license because OpenCV is MIT licensed. The program will alarm and replace it with avc1 by default, resulting in increased processing time. At the same time, due to the limited memory of the embedded system, the CPU used by the program to call the video stream is too long, resulting in memory overflow. The VideoCapture function in OpenCV follows the first-in-first-out principle when reading video stream data in H264 format. Since the program needs to obtain real-time monitoring images, it chooses to use the stack to store the latest video stream data read by VideoCapture. Finally, append and pop functions in the Python list are used to simulate a last-in, first-out stack that passes the latest frames between OpenCV and Yolo. In the case of memory overflow, the amount of stack pressure must be greater than the amount of stack output. The accumulation of a large amount of data on the stack will also cause the program to crash, and it is necessary to set a capacity cap value for the stack. When the amount of parameters stored in the stack reaches the threshold, a preset cleanup program is started; the state of the stack is reset, and memory resources are reclaimed. To ensure that the program does not overflow under long-term operation, the adjusted core code is as follows:

```
cap = cv2.VideoCapture(url)
 while success:
success,img = cap.read()
  if success:
stack.append(img)
   if len(stack) >= top:
    del stack[:]
gc.collect()
yolo = YOLO(**config)
if len(stack)!=0:
value = stack.pop()
image = Image.fromarray(cv2.cvtColor(value,cv2.COLOR_BGR2RGB))
yolo_img = yolo.detect_image(image)
result = cv2.cvtColor(np.asarray(yolo_img), cv2.COLOR_RGB2BGR)
 cv2.imshow("img",result)
```

4.2.2 Too strong vehicle lights will lead to poor recognition

Underground illumination is poor, and the lights are usually turned on for safe drive. When the lights are too bright, the video collected by the surveillance screen will be blurred (as shown in Fig. 13), resulting in the program being unable to accurately identify the target cart or the accuracy of recognition is less than 80%. To this end, a webcam is installed to assist identification in the manner of Fig. 14. When the front camera picture is blurred due to direct headlights, the rear camera can avoid the area with strong lights in front of the vehicle, and provide a clearer video stream data from the rear of the vehicle for final recognition. The camera for auxiliary identification is the same model as the main camera, both are DS-2CD2T25-I5 6mm Hikvision webcams, powered by POE, and pendant. The addition of cameras allows vehicles to be well identified and accurately classified, as shown at the top left of Fig.15.

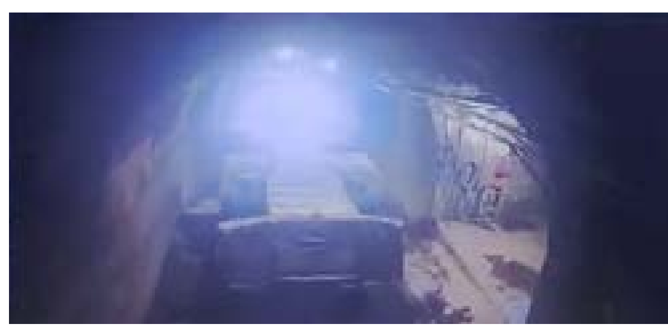

Fig. 13 Video screen of too bring car lights

Fig. 14 Schematic diagram of dual-camera installation

Fig.15 Recognition results of dual-camera with lights on

4.3 Analysis of the implementation effect of intelligent ore pass access control

(1) Enhanced safety protection for underground personnel. Underground miners are unable to pass through the ore pass door, ensuring the separation and control of personnel and vehicles during mining operations, and minimizing the risk of accidents.

(2) Beneficial for integration with mining processes. This intelligent ore pass system serves as a supplement to the mining process, making the control of mining area extraction and sliding systems more integrated. The centralized control room on the surface can obtain the actual situation of mining in different underground mining areas at any time.

(3) System operation and maintenance are relatively easy. The automatic opener, electromagnetic lock, door and camera of the system are located in the tunnel, while the other equipment is located in the rest chamber. Due to the narrow underground space, it is inevitable to cause scratches between the vehicle and the opened door or collision accidents may occur when the system malfunctions and the dooris not opened in a timely manner. Due to the fact that there are usually welders and material reserves in the surface workshop and underground maintenance chambers of mines, the maintenance and replacement of doors are convenient. When the camera malfunctions, the duty personnel can input instructions in the program window to manually control the opening and closing of the access control, without delaying the production of the mine.

(4) The system is easy to expand. This system is developed based on Jetson and has many reserved interfaces. In the future, new execution modules can be added at any time to achieve the intelligent infrastructure construction of the mine from monitoring center-to-switch-to-doorway-to- intelligent unit, better serving the intelligent construction of the mine.

5. Conclusion

This article constructs an intelligent unmanned control system for the underground special environment, achieving automatic identification of mining vehicles and automatic opening and closing of the pass door, achieving unmanned operation of the pass. At present, the system runs smoothly, with an average accuracy of 95% to 100% for identifying the four types of mining vehicles, and the access control response is fast with almost no delay.

In future research, image recognition algorithms can be used to develop a vehicle loading (waste rock) quantity recognition module based on existing programs, which can count the number of fully loaded and empty vehicles entering the ore pass connecting passage. This provides data support for mines to determine the ore pass material level and ore drawing time. At the same time, it can optimize the mine transportation plan based on this, reduce mine safety risks, and improve mine production efficiency.

References

Baccouche, A., Garcia-Zapirain, B., Zheng, Y., Elmaghraby, A.S. 2022. Early detection and classification of abnormality in prior mammograms using image-to-image translation and YOLO techniques. *Computer Methods and Programs in Biomedicine*, 221, 106884.https://doi.org/10.1016/j.cmpb.2022.106884

Girshick, R., Donahue, J., Darrell, T., Malik, J. 2014. Rich feature hierarchies for accurate object detection and semantic segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 580-587.

Guo, X., Li, Y. 2011. Underground Personnel Positioning System Based on ZigBee. *2011 Fourth International Symposium on Computational Intelligence and Design, Hangzhou, China, IEEE*, 1, 298-300. DOI: 10.1109/ISCID.2011.82

Imam, M., Baïna, K., Tabii, Y., Ressami, E.M., Adlaoui, Y., Benzakour, I., Abdelwahed, E.H. 2023. The Future of Mine Safety: A Comprehensive Review of Anti-Collision Systems Based on Computer Vision in Underground Mines. *Sensors (Basel, Switzerland)*, 23(9), 4294. https://doi.org/10.3390/s23094294

Kang, L., Lu, Z., Meng, L., Gao, Z. 2024. YOLO-FA: Type-1 fuzzy attention based YOLO detector for vehicle detection. *Expert Systems with Applications*, 237(Part B), 121209.https://doi.org/10.1016/j.eswa.2023.121209

Kumar, A., Kalia, A., Kalia, A., 2022. ETL-YOLO v4: A face mask detection algorithm in era of COVID-19 pandemic. *OPTIK*, 259, 169051.https://doi.org/10.1016/j.ijleo.2022.169051

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. 2016. Ssd: single shot multi box detector. *In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds) Computer Vision – ECCV 2016. ECCV 2016. Lecture Notes in Computer Science, Springer, Cham*, 9905, 21-37. https://doi.org/10.1007/978-3-319-46448-0_2

Madahana, M.C.I., Nyandoro, O.T.C., Ekoru, J.E.D. 2020. Intelligent comprehensive Occupational health monitoring system for mine workers. *IFAC-PapersOnLine*, 53(2), 16494-16499. https://doi.org/10.1016/j.ifacol.2020.12.751

Redmon, J., Farhadi, A. 2016. YOLO9000: Better, faster, stronger. *arXiv preprint arXiv:1612.08242*. https://doi.org/10.48550/arXiv.1612.08242

Redmon, J., Farhadi, A. 2018. YOLOv3:an incremental improvement. *arXiv preprint arXiv:1804.02767*. https://doi.org/10.48550/arXiv.1804.02767

Ren, S.Q, He, K., Girshick, R., Sun, J. 2017. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Translations on Pattern Analysis and Machine Intelligence*, 39(6), 1137-1149. DOI: 10.1109/TPAMI.2016.2577031

Tian,Y., Wang, S., Li, E., Yang, G., Liang, Z., Tan, M. 2023. MD-YOLO: Multi-scale Dense YOLO for small target pest detection. *Computers and Electronics in Agriculture*, 213, 108233.https://doi.org/10.1016/j.compag.2023.108233

Yang,C.Y., Li, C., Su, J.C., Wang, X.Q., He, Y.R. 2016. Research on video-based system of activity recognition for coal mine safety surveillance. *Coal Engineering*, 48(4), 111-113+117.