XIN LI[1], SHUANG WANG[2], LEI HE[3], QISHENG LUO[4]

A study on the influence of particle size on the identification accuracy of coal and gangue

Introduction

Coal is one of the primary energy sources in China, and it is also a basic energy resource to guarantee the development of our economy (Li et al. 2019). Coal and gangue separation can not only improve the quality of coal (Sun et al. 2022) and reduce transportation costs but can also enable the use of gangue for underground filling and making gangue bricks.

✉ Corresponding Author: XinLi; e-mail: lixin990122@126.com

[1] School of Mechanical Engineering, Anhui University of Science and Technology, Huainan, China;
State Key Laboratory of Mining Response and Disaster Prevention and Control in Deep Coal Mines,
Anhui University of Science and Technology, Huainan, China; ORCID iD; 0000-0003-4334-3532;
e-mail: lixin990122@126.com

[2] School of Mechanical Engineering, Anhui University of Science and Technology, Huainan, China;
State Key Laboratory of Mining Response and Disaster Prevention and Control in Deep Coal Mines,
Anhui University of Science and Technology, Huainan, China; e-mail: shuangw094@126.com

[3] School of Mechanical Engineering, Anhui University of Science and Technology, Huainan, China;
State Key Laboratory of Mining Response and Disaster Prevention and Control in Deep Coal Mines,
Anhui University of Science and Technology, Huainan, China; e-mail: helei_jxzl@126.com

[4] School of Mechanical Engineering, Anhui University of Science and Technology, Huainan, China;
State Key Laboratory of Mining Response and Disaster Prevention and Control in Deep Coal Mines,
Anhui University of Science and Technology, Huainan, China; e-mail: luoqs1025@163.com

Traditional methods of gangue selection have low efficiency, high water consumption, environmental pollution, and high labor intensity (Hu et al. 2022). With the development of artificial intelligence, intelligent sorting has gradually replaced traditional gangue selection. Intelligent sorting has high efficiency, low labor intensity, environmental friendliness, reduced staff, and improved efficiency, which promotes the development of intelligent mines (Kazanin et al. 2021; Varnavskiy et al. 2022). At present, there are two main ways of intelligent sorting: 1. Image recognition technology based on visible light in which a CCD industrial camera is used to image raw coal on the conveyor belt; 2. Radiation-based image recognition technology is used to reflect raw coal on the conveyor belt with γ and X rays. γ and X-rays can be identified by different degrees of radiation absorption or attenuation in coal and gangue (Zhang and Liu 2018; Robben et al. 2020), but this kind of equipment is radioactive (Sun et al. 2021). CCD industrial cameras use no radiation and low cost, so they are widely used.

In recent years, domestic and foreign scholars have done a lot of research on the extraction of coal gangue features and the influence of the external environment on the coal gangue identification rate.

In terms of the feature extraction of coal and gangue, Reddy, K.G.R. et al. used the traditional threshold segmentation method to study the gray characteristics of coal and coal gangue images and proposed a histogram threshold-based model for the separation of coal gangue from coal (Reddy and Tripathy 2013). Snehamoy, C. et al. extracted color, morphology, and texture features from the images and classified the rock types using an SVM model (Snehamoy et al. 2013). Alpana et al. analyzed the color and texture characteristics of coal and proposed an accurate, timely, and cost-effective automated characterization system based on machine learning (Alpana et al. 2020). Tripathy, D.P. et al. improved the recognition accuracy by combining color and texture features (Tripathy and Reddy 2017). The above scholars only studied the gray-scale texture features of coal gangue and did not integrate the surface gradient information of coal gangue with gray-scale information.

In terms of exploring the influence of external factors on the recognition of coal and gangue, Zhang et al. analyzed the influence on recognition efficiency under different water temperature and liquid intervention (Zhang et al. 2021). Eshaq, R.M.A. et al. verified that the infrared image is superior to the visible image under conditions of different temperatures, different illumination, and different areas of coal and gangue (Eshaq et al. 2020). Li et al. studied changes in the gray level and texture features of coal and gangue images under different surface water contents (Li et al. 2022). Wang et al. studied the influence of surface moisture of coal and gangue on the feature extraction and recognition accuracy of coal and gangue and proposed a coal and gangue identification method based on dielectric features and geometric constraints (Wang et al. 2021). The above scholars only studied the influence of the external environment on the features and recognition accuracy of coal and gangue but ignored the influence of the particle size change of coal and gangue itself.

On the one hand, the above kinds of literature only extract the gray-scale texture features of coal and gangue, ignoring the different effects of gradient information fluctuation.

On the other hand, the influence of coal and gangue particle size variation on the recognition accuracy is not considered. Therefore, in this paper, the features of different particle sizes are extracted from the gray level, gray level co-occurrence matrix, and gray level gradient co-occurrence matrix, and the optimal feature combination of each particle-size range is obtained. The influence of particle-size variation on recognition accuracy was studied under the same external conditions. Through experiments, it is found that when the training model has a single particle-size range, the identification accuracy of each particle-size range will be low. By broadening the training model of the particle-size range, the identification accuracy of coal and gangue in each particle-size range is improved. The method proposed in this paper can effectively improve the performance of the identification of coal and gangue sorting equipment with a single particle-size range, which has a guiding role in practical applications.

1. Experimental platform and image acquisition

The coal and gangue samples in this experiment are from Zhulinshan Mine, Jincheng City, Shanxi Province, which mainly produces anthracite with a particle size distribution of 0~100 mm. In this paper, the particle size range was divided into 20 mm intervals. Coal and gangue are divided into five particle size ranges by vernier caliper measurement, which are 0~20 mm, 20~40 mm,40~60 mm, 60~80 mm, and 80~100 mm. Partial coal and gangue samples are shown in Figure 1.

Fig. 1. Partial elements of coal and gangue

Rys. 1. Fragmenty węgla i skały płonnej

The experimental set up includes a feeder, belt conveyor, Hikvision CCD industrial camera, blue background plate, and LED light.

The resolution of the HikVision CCD industrial camera is 2448×2048, and the number of frames collected is 35 fps. The industrial camera is connected to a computer through USB 3.0 to capture and store images. Coal and gangue are transported by belt conveyor and discharged into a blue background plate industrial camera for real-time image collection. Through the LED light supplement light, the light intensity is constant at 1600 lux, and the conveyor-belt speed is 0.3 m/s.
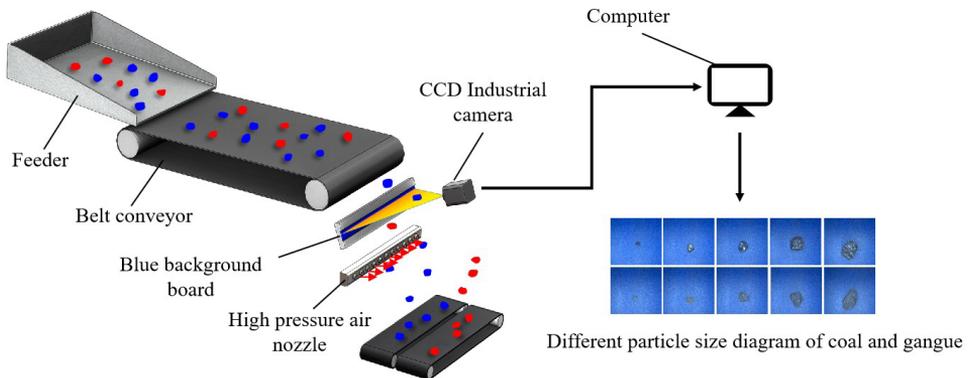
Fig. 2. Experimental platform and coal and gangue of different particle sizes

Rys. 2. Platforma doświadczalna oraz węgiel i skała płonna o różnej wielkości cząstek

In each particle size range, 200 pieces of coal and 200 pieces of gangue are analyzed. A total of 2,000 pieces of coal and gangue in the five groups with particle size range are analyzed. A total of 1,500 pieces of coal and gangue are taken as the training set, and the other 500 pieces are taken as the test set. The collected images of some different particle sizes are shown in Figure 2.

2. Image processing

2.1. Image segmentation

In the process of image acquisition, uneven illumination occurs in the background due to non-parallel light irradiation, as shown in Figure 3.

As shown in Figure 4, OTSU global threshold segmentation can find a segmentation threshold by the maximum between-class variance. However, if the illumination is not
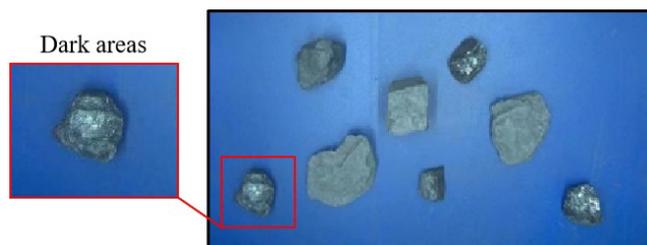

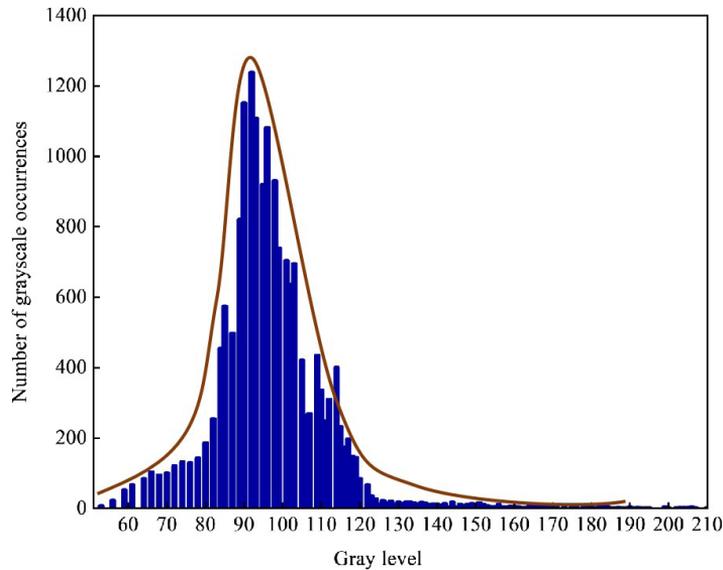
Fig. 3. Dark area

Rys. 3. Ciemny obszar

Fig. 4. Histogram of gray distribution

Rys. 4. Histogram rozkładu szarości

uniform in the image, it is easy to make the gray distribution histogram of coal and gangue images in the dark area without forming a double peak, failing traditional OTSU segmentation.

To solve the problem of uneven illumination, this paper proposes a block OTSU threshold segmentation method. In this method, the image is divided into multiple image blocks, and each image block is segmented by OTSU threshold segmentation. The uneven illumination is weakened in each image block, and the background and target gray histogram have obvious double peaks. The acquired image is divided into 5×5 blocks, and the OTSU method is used to calculate the segmentation threshold for each image block. Due to some image blocks only containing a background, it is necessary to calculate the average gray difference formula (1) between classes of each image block to judge whether there is a target in the image block and to perform OTSU threshold segmentation. If there is a target, OTSU threshold segmentation is performed. If there is no target, all of the pixels of the segmented image is set to zero.

$$\Delta d = \left| d_1\left(T\right) - d_2\left(T\right) \right| \tag{1}$$

Δd –   average gray difference between classes,
T –   segmentation threshold,
$d_1(T)$ –   average gray value of the background image when the threshold is T,
$d_2(T)$ –   average gray value of the target image when the threshold is T.

As shown in Figure 5, Δd of the objects is greater than 20, and Δd of the background is less than 20. In this paper, the average gray difference between classes is selected as 20 to distinguish the two image blocks with pure backgrounds and objects. Finally, the image block with the target is binarized.

By comparing the Canny (Kalbasi et al. 2020), LOG (Mallick et al. 2014; Ansari et al. 2017), OTSU global threshold (Zou et al. 2020) segmentation methods with the block OTSU threshold segmentation method, it can be seen from Figure 6 that Canny and LOG

Fig. 5. Information diagram of image block segmentation

Rys. 5. Diagram informacyjny segmentacji bloków obrazu
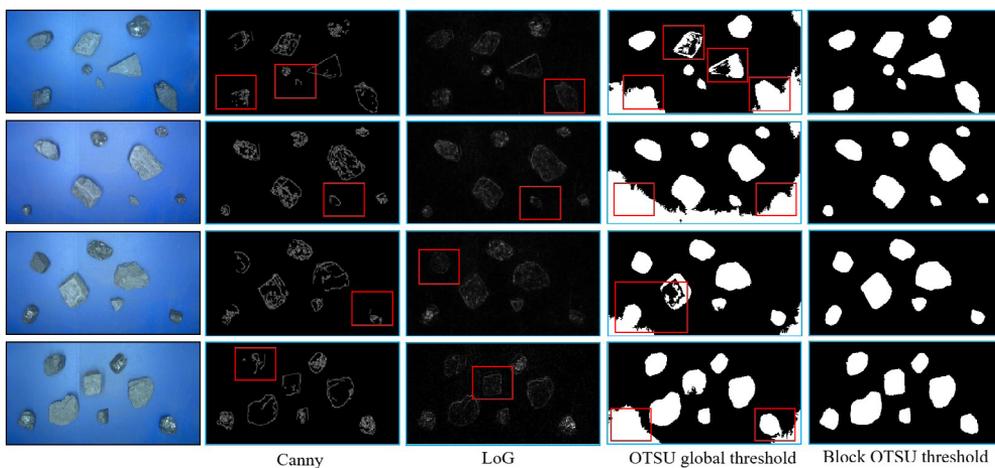

Fig. 6. Comparison of different segmentation methods

Rys. 6. Porównanie różnych metod segmentacji

cannot segment the edge of each piece of coal and gangue well, and it is almost impossible to segment small targets. The OTSU global threshold segmentation method is affected by uneven illumination, and the large area background is segmented into the foreground target. The block OTSU threshold segmentation method adopted in this paper has the best effect.

As shown in Figure 7, channel B is extracted from the collected RGB image (a) to obtain picture (b). Median filtering is used to denoise image (b), and the denoised image is segmented by the OTSU threshold to obtain a binarized image (c). Based on the binarization image, the original image background is removed by mask processing, and the object image (d) is obtained without the background. After obtaining the target without the background, feature extraction can be carried out on a single target.

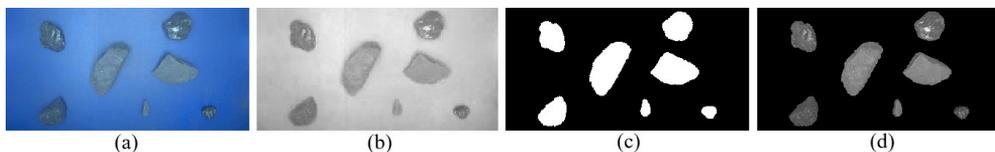(a)   (b)   (c)   (d)

Fig. 7. Foreground image extraction
a – RGB image, b – B-channel image, c – binarized image, d – image without background target

Rys. 7. Ekstrakcja obrazu pierwszego planu
a – obraz RGB, b – obraz z kanałem B, c – obraz zbinaryzowany, d – obraz bez docelowego tła

3. Feature extraction

3.1. Texture features

3.1.1. Gray gradient co-occurrence matrix

A gray level gradient co-occurrence matrix (GGCM) is a combination of the gray level of the image and gradient information (Hong 1984; Luo et al. 2022). The image gray level value can describe the specific content of the image, and the image gradient value can provide image contour edge information. Therefore, the gray gradient co-occurrence matrix can display the feature information of the image texture more comprehensively.

The elements $h(x, y)$ in the gray gradient co-occurrence matrix H are defined by the total number of pixels with a gray value of x and a gradient value of y contained in the normalized gray image $F(i, j)$ and the normalized gradient image $G(i, j)$. The elements in the co-occurrence matrix H are obtained by the statistics $x = F(i, j)$ and $y = G(i, j)$ in the normalized gray image $F(i, j)$ and gradient image $G(i, j)$. To reduce the computational burden, the gray level, gradient, and gray level gradient co-occurrence matrix should be normalized.

The normalized gray image is defined as:

$$F(i, j) = \frac{f(i, j) \cdot (L_{gy} - 1)}{f_{\max}} + 1 \tag{2}$$

↳ $f(i, j)$ – gray value of the original image at point (i, j),
  $F(i, j)$ – gray value of the transformed image at point (i, j),
  $f_{\max}$ – maximum gray value of the original image,
  L_{gy} – maximum gray value after normalization.

In this paper, $L_{gy} = 16$ (Luo et al. 2022).
The normalized gradient image is defined as:

$$G(i, j) = \frac{g(i, j) \cdot (L_{gt} - 1)}{g_{\max}} + 1 \tag{3}$$

↳ $g(i, j)$ – gradient value of the original image at point (i, j),
  $G(i, j)$ – gradient value of the transformed image at point (i, j),
  $g_{\max}$ – maximum gradient value of the original image,
  L_{gt} – normalized maximum gradient value.

In this paper, $L_{gt} = 16$ (Luo et al. 2022), $g(i, j)$ is obtained by the average sum of the gradient method of the Sobel operator in horizontal and vertical directions, as shown in the equation:

$$g(i, j) = \sqrt{g_x^2 + g_y^2} \tag{4}$$

$$g_x = f(i+1, j) - f(i, j) \tag{5}$$

$$g_y = f(i+1, j) - f(i, j) \tag{6}$$

The logarithms of pixels containing $x = F(i, j)$ and $y = G(i, j)$ in the normalized gray matrix $F(i, j)$ and the normalized gradient matrix $G(i, j)$ were counted.
The normalized gray gradient co-occurrence matrix is:

$$h'(x, y) = \frac{h(x, y)}{\displaystyle\sum_{x=0}^{L_{gy}} \sum_{y=0}^{L_{gt}} h(x, y)} \tag{7}$$

The gray gradient co-occurrence matrix contains small gradient advantages (T1), big gradient advantages (T2), gray level distribution inhomogeneity (T3), gradient distribution inhomogeneity (T4), energy (T5), average gray level (T6), average gradient (T7), gray mean square error (T8), gradient mean square error (T9), relevant (T10), gray entropy (T11), gradient entropy (T12), mixing entropy (T13), inertia (T14) and deficit moment (T15). A total of fifteen features were extracted.

3.1.2. Gray level co-occurrence matrix

Use of a gray level co-occurrence matrix (GLCM) is a common method to describe the texture which reflects the texture condition through the spatial correlation features of the gray level in a certain region. The gray level co-occurrence matrix can reflect the comprehensive information of the image gray level in the direction, adjacent interval, and change amplitude. Haralick et al. proposed a total of fourteen texture features in the gray-level co-existence matrix (Haralick et al. 1973). In this paper, four features including energy, entropy, contrast, and correlation were selected as texture features (Li et al. 2022), and the specific formula is shown in Table 1.

In this paper, the distance $d = 1$ is selected to extract features from the four directions of 0°, 45°, 90° and 135° (Zhang et al. 2017), and the means and standard deviations of the

Table 1.   GLCM feature calculation formula

Tabela 1.   Formuła obliczania cech GLCM

Feature Name	Computational Formula
Energy	$A = \sum\limits_{i=0}^{Z-1} \sum\limits_{j=0}^{Z-1} p^2(i,j,d,\theta)$
Entropy	$E = -\sum\limits_{i=0}^{Z-1} \sum\limits_{j=0}^{Z-1} p(i,j,d,\theta)\log_2 p(i,j,d,\theta)$
Contrast	$C = \sum\limits_{n-0}^{Z-1} n^2 \left[\sum\limits_{i=0}^{Z-1}\sum\limits_{j=0}^{Z-1} p^2(i,j,d,\theta) \right]$
Correlation	$R = \dfrac{\sum\limits_{i=0}^{Z-1} \sum\limits_{j=0}^{Z-1} ijp^2(i,j,d,\theta) - \mu_1\mu_2}{\sigma_1\sigma_2}$

Where Z is the gray level series, $P(i, j, d, \theta)$ represents the number of pixel pairs in direction θ with adjacent distance d; μ_1 is the mean value of variable i; μ_2 is the mean value of variable j; σ_1 is the variance of variable i; σ_2 is the variance of variable j.

four directions are taken as the basis for analysis and recognition. The mean energy (T16), the standard deviation of energy (T17), the mean entropy (T18), the standard deviation of entropy (T19), the mean contrast (T20), the standard deviation of contrast (T21), the mean correlation (T22) and the standard deviation of correlation (T23).

3.2. Gray level features

The gray level feature is the most intuitive reflection of the difference between coal and gangue. The surface of coal is black, and there is a reflection phenomenon of mirror body tissue after lighting, and the gray level is low. The surface of gangue is gray white, not reflective, the gray level is higher. In this paper, the mean (T24) and variance (T25) of the gray level of coal and gangue are selected as the gray level features (He et al. 2023). The gray mean m and gray variance σ are described as:

$$m = \frac{1}{W} \sum_{i=1}^{N} \sum_{j=1}^{M} q(i,j) \tag{8}$$

$$\sigma = \sqrt{\frac{1}{W} \sum_{i=1}^{N} \sum_{j=1}^{M} (q(i,j) - m)^2} \tag{9}$$

Where W is the number of non-zero image pixels, N and M are the length and width of the image, $q(i,j)$ is the gray value of image pixels.

4. Feature selection and classifier

4.1. Relief feature selection

Because image feature extraction takes a lot of time and the importance of each feature to classification is different, the less important features should be eliminated. Relief feature selection assigns different weights to features according to the relevant statistics of each feature, which can effectively eliminate redundant features. Compared with traditional filtering feature-selection methods, such as Information gain (IG), PCA dimensionality reduction, the relief feature selection method has a linear increase in running time with the increase of sampling times and the number of original features (He et al. 2022; Dou et al. 2019). Therefore, the relief method has high operational efficiency and accurate results (Urbanowicz et al. 2018).

5. Experimental analysis

5.1. Experimental parameter configuration

The training and testing in this study were run on a computer with an AMD Ryzen 5 5600X 6-Core CPU and an Nvidia GeForce RTX3060 GPU. MatlabR2020b was installed and used as the programming language to implement this study.

5.2. Classifier selection

Five groups of coal and gangue images with different particle sizes were formed into a 1500×25 matrix dataset, and the dataset was randomly sorted. The machine-learning classifier types were compared and analyzed by using the five-fold cross-validation method. Common classifier models include SVM with different kernel functions, K-nearest neighbor (KNN), Decision TREES and Naive Bayes.

To ensure the stability of recognition accuracy of each classifier model, the average value of each classifier model was repeated five times. During the experiment, the recognition
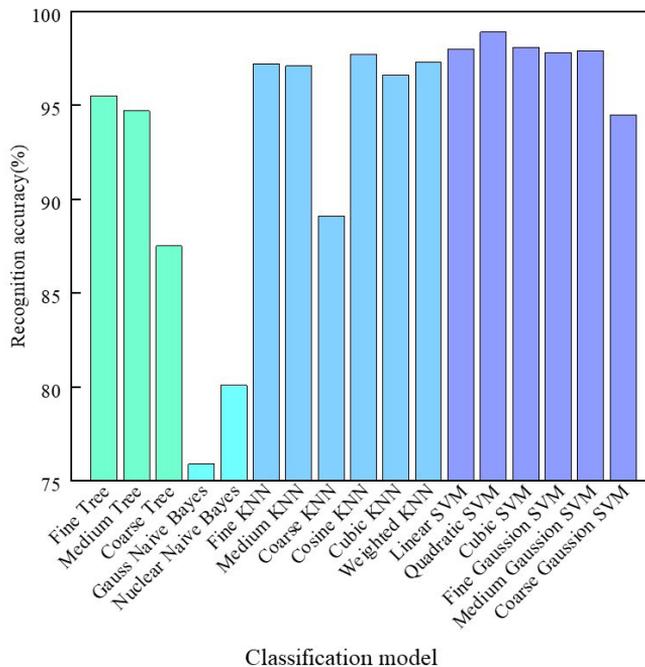

Fig. 8. Recognition accuracy of each classifier

Rys. 8. Dokładność rozpoznawania każdego klasyfikatora

accuracy of each classifier model was counted and the bar chart was drawn for comparative analysis. The results are shown in Figure 8.

According to Figure 8, the recognition accuracy of the SVM classifier is higher than that of other classifier models and the quadratic SVM has the highest recognition accuracy among SVM classifiers. Therefore, quadratic SVM will be used in this paper as a classifier for coal and gangue identification in each particle-size range.
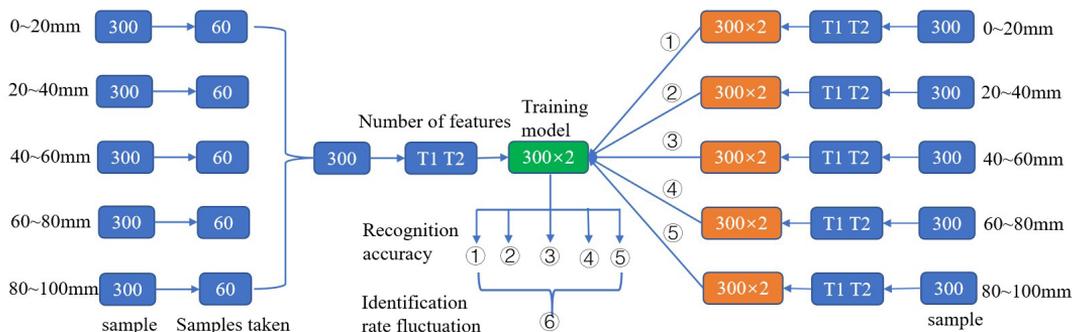

Fig. 9. Flow chart of recognition rate fluctuation when the number of features is two

Rys. 9. Schemat przepływowy fluktuacji współczynnika rozpoznawania, gdy liczba cech wynosi dwa
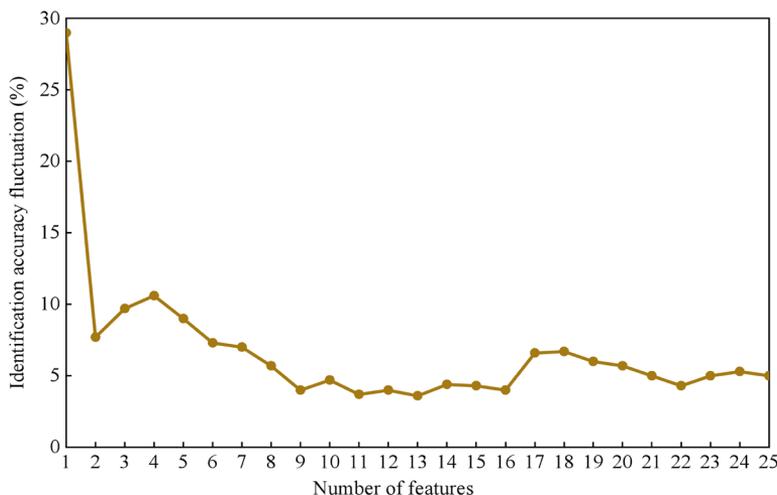

Fig. 10. Fluctuation of recognition accuracy

Rys. 10. Fluktuacja dokładności rozpoznawania

5.3. The influence of particle-size change on recognition rate

From 0~20 mm, 20~40 mm, 40~60 mm, 60~80 mm, and 80~100 mm, each particle size range of 60 data selected, a total of 300, the number of features increased from 1 to 25, each additional feature formed a training model, a total of 25 models. The 300 data with particle sizes ranging from 0 to 20 mm were increased from 1 to 25, and each additional feature formed a test set. The same was true for other particle-size ranges. Each model is trained and tested to obtain the recognition accuracy of five particle-size ranges (①, ②, ③, ④, ⑤), and the maximum difference of the five recognition accuracies (⑥) is calculated, which is the fluctuation of the recognition rate. Figure 9 shows the flow chart of recognition rate fluctuation when the number of features is two. Figure 10 shows the fluctuation of the recognition rate.

As can be seen from Figure 10, when the number of feature inputs is 11, the fluctuation is still 3.7% and shows a large fluctuation. It can be seen that the change in particle size has an impact on recognition accuracy.

5.4. Feature selection

5.4.1. Characteristic weights of coal and gangue with different particle sizes

The twenty-five gray level and texture features of coal and gangue images in each particle size range are extracted. Set 1 as the label of coal and 2 as the label of gangue. Some feature extraction results are shown in Table 2.

Table 2.   Partial feature extraction results

Tabela 2.   Wyniki częściowej ekstrakcji cech

Feature Label	T1	T2	T16	T18	T21	T23	T24	T25
1	0.9765	0.0662	0.9354	0.00014	0.2281	0.0031	75.88	192.48
1	0.9651	0.1319	0.9167	0.00017	0.3238	0.0049	99.96	450.57
1	0.9630	0.1121	0.9048	0.00018	0.3395	0.0051	82.53	344.02
1	0.9640	0.1106	0.9035	0.00021	0.3664	0.0057	95.21	504.75
2	0.9665	0.0787	0.9060	0.00018	0.2981	0.0037	105.52	515.88
2	0.9792	0.0563	0.9387	0.00019	0.2201	0.0032	104.90	337.47
2	0.9760	0.0579	0.9300	0.00015	0.2376	0.0030	107.76	403.77
2	0.9742	0.0648	0.9241	0.00019	0.2648	0.0036	113.69	487.63

In MatLab R2020b, the weight of coal and gangue features of each particle-size range is calculated by relief feature selection, and the weight curve of each feature is drawn, as is shown in Figure 11.

As can be seen from Figure 11, the arrangement of feature weights with different particle sizes is different, and the arrangement order is shown in Table 3.
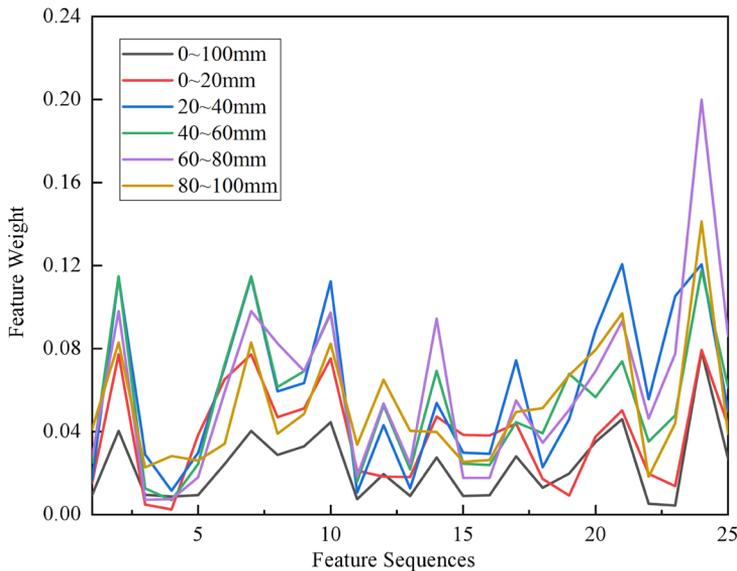

Fig. 11. The weight of each feature with different particle sizes

Rys. 11. Masa każdej cechy przy różnych rozmiarach cząstek

Table 3.   Weight sizes of different particle sizes

Tabela 3.   Wielkości wagowe cząstek o różnych rozmiarach

Particle-size range	Weight Size Order
0~20 mm	T24 > T2 > T7 > T10 >T6 > T9 > T21 > T14 > T8 > T17 > T25 > T15 > T5 > > T16 > T20 > T11 > T22 > T12 > T13 > T18 > T1 > T23 > T19 > T3 > T4
20~40 mm	T21 > T24 > T2 > T7 > T10 > T23 > T20 > T17 > T6 > T9 > T8 > T22 > T14 > > T25 > T19 > T12 > T15 > T5 > T16 > T3 > T18 > T1 > T13 > T4 > T11
40~60 mm	T24 > T2 > T7 > T10 > T21 > T6 > T14 > T9 > T19 > T8 > T25 > T20 > T12 > > T23 > T17 > T18 > T22 > T1 > T15 > T16 > T5 > T13 > T11 > T3 > T4
60~80 mm	T24 > T2 > T7 > T10 > T14 > T21 > T25 > T8 > T23 > T9 > T20 > T6 > T17 > > T12 > T19 > T22 > T18 > T1 > T13 > T11 > T5 > T15 > T16 > T4 > T3
80~100 mm	T24 > T21 > T2 > T7 > T10 > T20 > T19 > T12 > T18 > T17 > T9 > T23 > T1 > > T13 > T14 > T8 > T25 > T6 > T11 > T4 > T16 > T5 > T15 > T3 > T22

5.4.2. The optimal characteristic combination of coal
and gangue with different particle sizes

According to 5.4.1, the feature weights of coal and gangue with different particle sizes are in a different order. The features of coal and gangue with each particle-size range are increased from large to small, forming 300×1, 300×2 ... for 300×25 matrices, the 50% fold cross-validation method was used, and the recognition accuracy of the number of features in each particle-size range from 1 to 25 was counted. To reduce the fluctuation of operating results and ensure the stability of experimental verification, the above process was repeated five times for each particle-size range, and the final results were taken as the optimal average value, as shown in Figure 12.
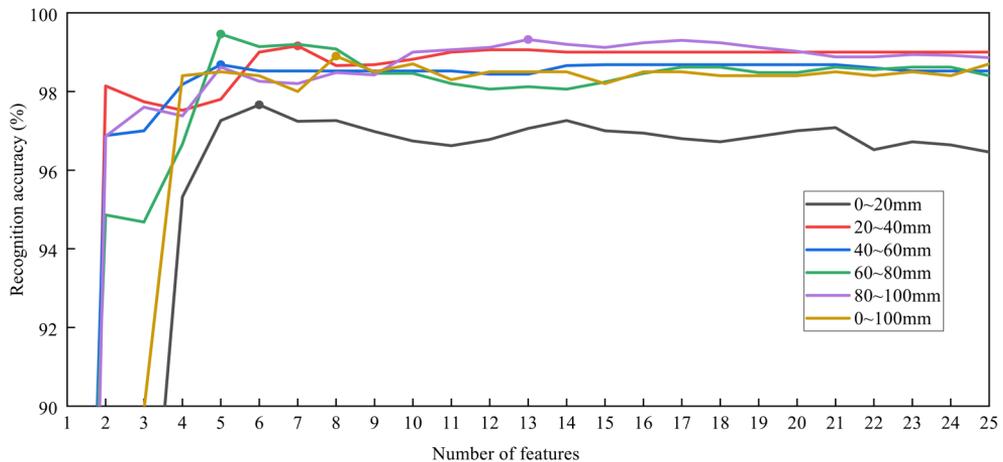

Fig. 12. Optimal combination of features with different particle sizes

Rys. 12. Optymalna kombinacja cech przy różnych wielkościach cząstek

As can be seen from Figure 12, the optimal characteristic combination of coal and gangue is different due to the change in particle size. When the recognition accuracy reaches the optimum, the continuous increase of features will lead to a decrease in recognition accuracy. Therefore, the optimal number of features from 0 mm to 20 mm is 6, which are T24, T2, T7, T10, T6 and T9. The optimal number of features from 20 mm to 40 mm is 7, which are T21, T24, T2, T7, T10, T23 and T20. The optimal number of features from 40 to 60 mm is 5, which are T24, T2, T7, T10 and T21. The optimal number of features from 60 mm to 80 mm is 5, which are T24, T2, T7, T10 and T14. The optimal number of features from 80 mm to 100 mm is 13, which are T24, T21, T2, T7, T10, T20, T19, T12, T18, T17, T9, T23 and T1.

5.5. Particle-size identification experiment

The coal gangue with different particle sizes was identified by particle sizes through setting experiments, as shown in Table 4. The average recognition rate is 97.2%.

To improve the recognition accuracy of different particle sizes, experiments were set to broaden the particle size range of the training set to 0~100 mm, as shown in Table 5. The experimental results show that the recognition accuracy of each particle size range is improved and the average recognition accuracy reaches 99.2%.

Therefore, increasing the training particle-size range of the training set can effectively improve the recognition accuracy of coal and gangue with different particle sizes, and the maximum increase is 4%. If the particle-size range of the training set is the same as that of the test set, the recognition accuracy of the test set is low.

Table 4.   Particle size identification accuracy

Tabela 4.   Dokładność identyfikacji wielkości cząstek

Training Set	Test Set	Recognition accuracy
300×6 (0~20 mm)	100×6 (0~20 mm)	97%
300×7 (20~40 mm)	100×7 (20~40 mm)	97%
300×5 (40~60 mm)	100×5 (40~60 mm)	96%
300×5 (60~80 mm)	100×5 (60~80 mm)	98%
300×13 (80~100 mm)	100×13 (80~100 mm)	98%

Table 5.   Particle size identification after broadening the training set particle-size range

Tabela 5.   Identyfikacja wielkości cząstek po rozszerzeniu zakresu wielkości cząstek w zbiorze uczącym

Training Set	Test Set	Recognition accuracy
1500×8 (0~100 mm)	100×8 (0~20 mm)	99%
1500×8 (0~100 mm)	100×8 (20~40 mm)	100%
1500×8 (0~100 mm)	100×8 (40~60 mm)	100%
1500×8 (0~100 mm)	100×8 (60~80 mm)	98%
1500×8 (0~100 mm)	100×8 (80~100 mm)	99%

6. Experimental verification

To verify the reliability of the above experiments by changing the extent of broadening the particle-size range of the training set , the 0~100 mm particle-size range was divided into the 0~40 mm particle size range and the 40~100 mm particle size range, and the training set with 600×8 and 900×8 matrices was composed. The 100×8 test sets of 0~20 mm, 20~40 mm, 40~60 mm, 60~80 mm and 80~100mm were classified, and the accuracy of training and recognition in a single particle size range was compared with Table 4. This is shown in Figure 13.
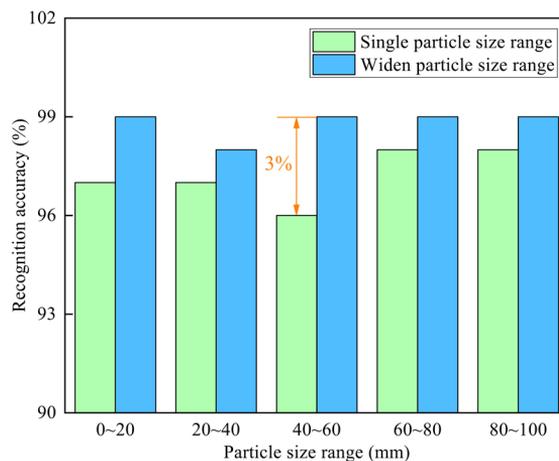

Fig. 13. Experimental verification diagram

Rys. 13. Schemat weryfikacji eksperymentalnej

As can be seen from Figure 13, the recognition accuracy after broadening the particle size range of the training set is higher than that of a single particle size range, and the maximum increase is 3%. Therefore, broadening the particle size range of the training set is an effective method to improve the recognition accuracy of a single particle size range.

Conclusions

The difference in the optimal feature combination of different particle sizes was studied. The fluctuation of recognition accuracy caused by the change in particle size was studied. In addition, methods to improve the accuracy of coal and gangue identification in a single particle-size range are proposed. The specific conclusions are as follows:

1. The optimal feature combination of coal and gangue with different particle-size ranges is different. Change of particle size is a key factor affecting the recognition rate.

The influence of particle-size change of coal gangue should be considered in the recognition technology based on visible light technology. The fluctuation of recognition accuracy due to the change of particle size gradually decreases with the increase of the number of features.

2. The size range of the training set is broadened and the maximum improvement of the recognition accuracy of particle size is 4%, and the average recognition accuracy is 99.2%. Therefore, for equipment sorting coal and gangue in a certain single particle-size range, materials that exceed the upper and lower limits of the particle-size range treated by the equipment should be introduced to train the identification model in order to improve the identification performance of the equipment.

REFERENCES

Alpna. and Chand, S. 2020. An intelligent technique for the characterization of coal microscopic images using ensemble learning. *Journal of Intelligent & Fuzzy Systems* 38(5), pp. 6257–6267, DOI: 10.3233/JIFS-179707.

Ansari et al. 2017 – Ansari, M., Diksha, K. and Manish, D. 2017. A comprehensive analysis of image edge detection techniques. *International Journal of Multimedia and Ubiquitous Engineering* 12(11), pp. 1–12, DOI: 10.14257/ijmue.2017.12.11.01.

Dou et al. 2019 – Dou, D., Wu, W., Yang, J. and Zhang, Y. 2019. Classification of coal and gangue under multiple surface conditions via machine vision and relief-SVM. *Powder Technology* 356, pp. 1024–1028, DOI: 10.1016/j.powtec.2019.09.007.

Eshaq et al. 2020 – Eshaq, R., Hu, E., Li, M. and Alfarzaeai, M.S. 2020. Separation between coal and gangue based on infrared radiation and visual extraction of the YCbCr color space. *Ieee Access* 8, pp. 55204–55220, DOI: 10.1109/ACCESS.2020.2981534.

Haralick et al. 1973 – Haralick, R.M., Shanmugan, K. and Dinstein, I. 1973. Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics* SMC-3(6), pp. 610–621.

He et al. 2022 – He, L., Wang, S., Guo, Y., Hu, K. and Cheng, G. 2022. Shape selection recognition and scattering distribution prediction of adhesion targets in multi-scale dual-energy X-ray images of coal and gangue. *International Journal of Coal Preparation and Utilization*. DOI: 10.1080/19392699.2022.2122453.

He et al. 2023 – He, L., Wang, S., Guo, Y., Hu, K., Cheng, G. and Wang, X. 2022. Study of raw coal identification method by dual-energy x-ray and dual-view visible light imaging. *International Journal of Coal Preparation and Utilization* 43(2), DOI: 10.1080/19392699.2022.2051013.

Hong, J. 1984. Gray Level-Gradient Cooccurrence Matrix Texture Analysis Method. *Acta Automatica Sinica* 10(1), pp. 22–25 (*in Chinese*).

Hu et al. 2022 – Hu, F., Zhou, M., Yan, P., Liang, Z. and Li, M. 2022. A Bayesian optimal convolutional neural network approach for classification of coal and gangue with multispectral imaging. *Optics and Lasers in Engineering* 156, DOI: 10.1016/j.optlaseng.2022.107081.

Kalbasi, M. and Nikmehr, H. 2020. Noise-Robust, Reconfigurable Canny Edge Detection and its Hardware Realization. *IEEE ACCESS* 99, pp. 39934–39945, DOI: 10.1109/ACCESS.2020.2976860.

Kazanin et al. 2021 – Kazanin, O., Sidorenko, A. and Drebenstedt, C. 2021. Intensive underground mining techno-logies: Challenges and prospects for the coal mines in Russia. *Acta Montanistica Slovaca* 26(1), pp. 60–69, DOI: 10.46544/AMS.v26i1.05.

Li, J. and Wang, J. 2019. Comprehensive utilization and environmental risks of coal and gangue: A review. *Journal of Cleaner Production* 239, DOI: 10.1016/j.jclepro.2019.117946.

Li et al. 2022 – Li, M., He, X., Duan, Y. and Yang, M. 2022. Experimental study on the influence of external fac-tors on image features of coal and gangue. *International Journal of Coal Preparation and Utilization* 42(9), pp. 2770–2787, DOI: 10.1080/19392699.2021.1901692.

Luo et al. 2022 – Luo, Q., Wang, S., Li, X. and He, L. 2022. Recognition of coal and gangue based on multi-dimen-sional gray gradient feature fusion. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects* 44(3), pp. 8060–8076, DOI: 10.1080/15567036.2022.2119309.

Mallick et al. 2014 – Mallick, A., Roy, S., Chaudhuri, S. and Roy, S. 2014. Optimization of Laplace of Gaussian (LoG) filter for enhanced edge detection: A new approach. *In Proceedings of the 2014 International Conferen-ce on Control, Instrumentation, Energy and Communication (CIEC)*, pp. 658–661.

Reddy, K. and Tripathy, D. 2013. Separation of gangue from coal based on histogram thresholding. *International Journal of Technology Enhancements and Emerging Engineering Research* 1(4), pp. 31–34.

Robben et al. 2020 – Robben, C., Condori, P., Pinto, A., Machaca, R. and Takala, A. 2020. X-ray-transmission based ore sorting at the San Rafael tin mine. *Minerals Engineering* 145, DOI: 10.1016/j.mineng.2019.105870.

Snehamoy, C. 2013. Vision-based rock-type classification of limestone using multi-class support vector machine. *Applied intelligence* 39(1), pp. 14–27, DOI: 10.1007/s10489-012-0391-7.

Sun et al. 2021 – Sun, Z., Huang, L. and Jia, R. 2021. Coal and gangue separating robot system based on computer vision. Sensors 21(4), DOI: 10.3390/s21041349.

Sun et al. 2022 – Sun, Z., Lu, W., Xuan, P., Li, H., Zhang, S., Niu, S. and Jia, R. 2019. Separation of gangue from coal based on supplementary texture by morphology. *International Journal of Coal Preparation and Utilization* 42(3), pp. 221–237, DOI: 10.1080/19392699.2019.1590346.

Tripathy, D. and Reddy, K. 2017. Novel Methods for Separation of Gangue from Limestone and Coal using Mul-tispectral and Joint Color-Texture Features. *Journal of The Institution of Engineers (India): Series D* 98, pp. 109–117, DOI: 10.1007/s40033-015-0106-4.

Urbanowicz et al. 2018 – Urbanowicz, R., Meeker, M., La Cava, W., Olson, R. and Moore, J. 2018. Relief-based feature selection: Introduction and review. *Journal of biomedical informatics* 85, pp. 189–203, DOI: 10.1016/j.jbi.2018.07.014.

Wang et al. 2021 – Wang, X., Wang, S., Guo, Y., Hu, K. and Wang, W. 2021. Recognition of coal and gangue based on dielectric characteristics and geometric constraints under multi factors. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, DOI: 10.1080/15567036.2021.1968546.

Varnavskiy et al. 2022 – Varnavskiy, K., Nepsha, F. and Chen, Q. 2022. The assessment of functional efficiency of technological structure for the coal mine working face–an application of IIIE. *Journal of Industrial Informa-tion Integration* 26, DOI: 10.1016/j.jii.2021.100262.

Zhang, N. and Liu, C. 2018. Radiation characteristics of natural gamma-ray from coal and gangue for recognition in top coal caving. *Scientific Reports* 8, DOI: 10.1038/s41598-017-18625-y.

Zhang et al. 2017 – Zhang, X., Cui, J., Wang, W. and Lin, C. 2017. A study for texture feature extraction of high-re-solution satellite images based on a direction measure and gray level co-occurrence matrix fusion algorithm. *Sensors* 17(7), DOI: 10.3390/s17071474.

Zhang et al. 2021 – Zhang, J., He, G. and Yang, S. 2021. Controlling water temperature for efficient coal/gangue recognition. *Materials Today Chemistry* 22, DOI: 10.1016/j.mtchem.2021.100587.

Zou et al. 2020 – Zou, X., Tan, W., Huang, X., Nan, S., Bai, Y. and Fu, X. 2020. Imaging quality enhancement in binary ghost imaging using the Otsu algorithm. *Journal of Optics* 22(9), DOI: 10.1088/2040-8986/aba22e.

A STUDY ON THE INFLUENCE OF PARTICLE SIZE ON THE
IDENTIFICATION ACCURACY OF COAL AND GANGUE

Keywords

particle size, gray feature, texture feature, support vector machine,
coal and gangue identification

Abstract

In order to explore the impact of coal and gangue particle size changes on recognition accuracy and to improve the single particle size of coal and gangue identification accuracy of sorting equipment, this study established a database of different particle sizes of coal and gangue through image gray and texture feature extraction, using a relief feature selection algorithm to compare different particle size of coal and gangue optimal features of the combination, and to identify the points and particle size of coal and gangue. The results show that the optimal features and number of coal and gangue are different with different particle sizes. Based on visible-light coal and gangue separation technology, the change of coal and gangue particle size cause fluctuations in the recognition accuracy, and the fluctuation of recognition accuracy will gradually decrease with increases in the number of features. In the process of particle size classification, if the training model has a single particle size range, the recognition accuracy of each particle size range is low, with the highest recognition accuracy being 98% and the average recognition rate being only 97.2%. The method proposed in this paper can effectively improve the recognition accuracy of each particle size range. The maximum recognition accuracy is 100%, the maximum increase is 4%, and the average recognition accuracy is 99.2%. Therefore, this method has a high practical application value for the separation of coal and gangue with single particle size.

BADANIE WPŁYWU WIELKOŚCI CZĄSTEK NA DOKŁADNOŚĆ IDENTYFIKACJI
WĘGLA I SKAŁY PŁONNEJ

Słowa kluczowe

wielkość cząstek, cecha szarości, cecha tekstury, maszyna wektorów pomocniczych,
identyfikacja węgla i skały płonnej

Streszczenie

W celu zbadania wpływu zmian wielkości cząstek węgla i skały płonnej na dokładność rozpoznawania oraz poprawienia dokładności identyfikacji pojedynczych cząstek węgla i skały płonnej przez urządzenia sortujące, w ramach tej pracy utworzono bazę danych różnych rozmiarów cząstek węgla i skały płonnej za pomocą obrazów szarych i ekstrakcję cech tekstury przy użyciu algorytmu wyboru cech reliefowych w celu porównania różnych rozmiarów cząstek węgla i skały płonnej przy

optymalnych cechach kombinacji oraz identyfikacji punktów i wielkości cząstek węgla i skały płonnej. Wyniki pokazują, że optymalne liczby cech węgla i skały płonnej są różne dla różnych rozmiarów cząstek. W oparciu o technologię separacji węgla i skały płonnej w świetle widzialnym, zmiana wielkości cząstek węgla i skały płonnej powoduje fluktuacje dokładności rozpoznawania, a te z kolei będą stopniowo zmniejszać się wraz ze wzrostem liczby cech. W procesie klasyfikacji wielkości cząstek, jeśli model uczący ma jeden zakres wielkości cząstek, dokładność rozpoznawania każdego zakresu wielkości cząstek jest niska, przy czym najwyższa dokładność rozpoznawania wynosi 98%, a średni wskaźnik rozpoznawania wynosi tylko 97,2%. Metoda zaproponowana w tym artykule może skutecznie poprawić dokładność rozpoznawania każdego zakresu wielkości cząstek. Maksymalna dokładność rozpoznawania wynosi 100%, maksymalny wzrost to 4%, a średnia dokładność rozpoznawania to 99,2%. Dlatego ta metoda ma dużą praktyczną wartość użytkową do oddzielania węgla i skały płonnej według rozmiaru pojedynczej cząstki.