A Novel Particle Size Detection System Based on RGB-laser Fusion Segmentation with Feature Dual-recalibration for Blast Furnace Materials

Jinshi Liu, *Graduate Student Member,IEEE*, Zhaohui Jiang, Weihua Gui, *Member,IEEE*, and Zhiwen Chen, *Member,IEEE*

***Abstract*—Particle size detection (PSD) is used to obtain the particle size distribution of materials in the blast furnace charging process, which is significant for optimizing the gas flow distribution and ensuring stable production. However, due to the complex surface texture of the materials and the uneven illumination of the production environment, existing methods have difficulty obtaining the particle size distribution efficiently. This paper proposes an end-to-end PSD system based on image segmentation to obtain the particle size distribution online with high accuracy. First, to further enhance the expression of edge features and reduce the interference of complex textures, an RGB-laser particle segmentation network (RLPNet) is developed to obtain high-precision segmentation images by camera-LiDAR sensor fusion. Moreover, to improve the fusion of RGB and laser features, a feature dual-recalibration (FDR) module was designed and embedded in RLPNet, consisting of independent recalibration and joint recalibration with T-convolution. Finally, to reduce the error caused by missing edge particle pixels, an edge-recognition-based particle size calculation strategy (ERP) is presented. Experimental results demonstrate that the proposed method performs well on the constructed dataset and in industrial applications. With the segmentation accuracy of RLPNet reaching 64.19%, the similarity between the particle size distribution predicted with ERP and the actual distribution reaches 79.19%.**

***Index Terms*—Image segmentation, Sensor fusion, Feature Recalibration, Particle size detection, Convolutional neural networks.**

I. INTRODUCTION

BLAST furnaces are some of the most critical pieces of production equipment in the ironmaking industry, of which the internal reaction stability determines the quality of ironmaking [1]–[3]. The particle size distribution of blast furnace (BF) materials is a key performance indicator in the charging system, and its measurement accuracy and continuity are essential for improving the gas permeability and internal reaction stability in BFs, as shown in Fig. 1. Therefore, the efficient and accurate particle size detection (PSD) of BF materials is of great significance for optimizing the gas flow distribution and guaranteeing stable production [4]–[6].

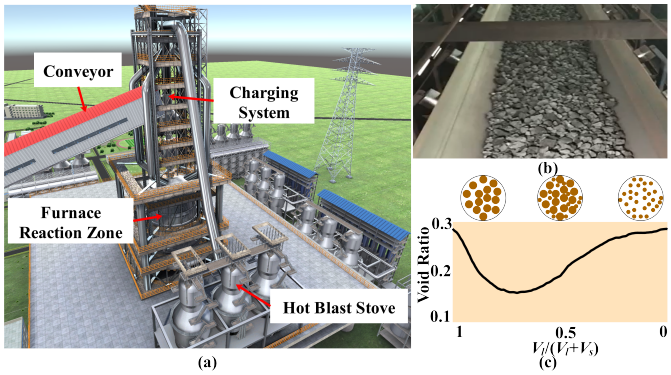


Fig. 1. Blast furnace equipment images from an ironmaking plant: (a)virtual 3D ironworks; (b)blast furnace charging conveyor. And (c)curve of the permeability of the gas stream plotted against the void ratio, which depends on the ratio of small particles (V_s) to large particles (V_l), i.e., the particle size distribution [7], [8]. When the content of large particles is close to 100%, the void ratio depends almost entirely on the voids between large particles, and the void ratio is larger at this time. When the content of small particles is close to 100%, the void ratio almost entirely depends on the voids between small particles. Although the voids between small particles are smaller than those between large particles, the number of small particles is greater, so the void ratio with fixed space is still larger. When the content of large particles is close to 70%, small particles just fill the voids of large particles, so the void ratio at this time is the smallest.

After preliminary screening, the BF materials are transported to the BF top through a conveyor belt. Due to mixing and losses during transportation, the particle size of the materials entering the BF changed compared to that in preliminary screening. In addition, since preliminary screening is carried out through 3-6 layers of fixed-size sieves, the particle size distribution obtained by the preliminary screening is less classified and has low accuracy. The above problems make it difficult to accurately evaluate the porosity between BF material stacks. The gas permeability in the furnace may deteriorate for this, and the contact area between the gas flow and the material surface will be affected, reducing the sufficiency of the reduction reaction [9]. Therefore, only by online high-precision detection of the particle size distribution of BF materials on the conveyor can the ingredient and charging system be adjusted according to the specific distribution to

This work was supported in part by the National Major Scientific Research Equipment of China (Grant No. 61927803). (Corresponding author: Zhaohui Jiang).

Jinshi Liu, Zhaohui Jiang, Weihua Gui and Zhiwen Chen are with the School of Automation, Central South University, Changsha 410083, China. (e-mail: ljs11528@csu.edu.cn; jzh0903@csu.edu.cn; gwh@csu.edu.cn; zhiwen.chen@csu.edu.cn.)

Zhaohui Jiang, Weihua Gui and Zhiwen Chen are with Pengcheng laboratory, Shenzhen 518055, China.

ensure product quality. The existing PSD methods are mainly physical detection methods, instrumental detection methods, and image segmentation methods.

Physical detection methods. Most ironworks use physical detection methods, using the physical properties of material particles such as gravity and density to detect particle size by screening, sedimentation, and other methods. This method is easy to implement but has low accuracy and can only be detected offline.

Instrument detection methods. Instrument detection methods mainly include light scattering, laser phase Doppler, and aerodynamics. The light scattering method calculates the particle size by measuring the size of the scattered light signal emitted by the particles irradiated with light, and its measurement range is 0.1-25 mm [10]. This method provides fast, repeatable, and realistic measurements but requires periodic calibration and offline sampling. The laser phase Doppler method uses Mie scattering theory to measure particle size, and its measurement range is 1.0E-06−1.0E-03 mm [11]. This method avoids the influence of temperature and pressure but can only measure at a single point and requires offline sampling. The aerodynamic method mainly uses the characteristics of particles of different sizes to generate different accelerations in the accelerated airflow to calculate the particle size, and its measurement range is 0.37-20 mm [12]. This method can avoid the interference caused by factors such as the particle refractive index and density, but it is susceptible to the interference of particle overlap and requires offline sampling. Most of these methods require offline sampling, and their measurement range cannot completely cover the particle size range of BF materials. Therefore, it is difficult for these methods to satisfy high-efficiency and high-precision detection requirements.

Image segmentation methods. Due to the advantages of online, fast, and noncontact image processing technology, a set of PSD methods with image segmentation as the core has been gradually formed. However, due to objective factors such as complex surface texture, dense accumulation, and uneven illumination of BF materials, it is difficult for traditional image segmentation methods to use low-level features to effectively distinguish edges from noises. With the wide application of deep learning algorithms in image segmentation tasks, mining high-dimensional features can improve the accuracy of image segmentation.

In 2015, a Fully Convolutional Network (FCN) was proposed to extract image features through the encoder and then gradually merge the high-level features at the top of the encoder and the low-level features at the bottom through the decoder to form high-quality semantic segmentation outcomes [13]. Later, some segmentation networks, such as UNet, SegNet, Deeplab, and GCNs, were successively proposed and achieved good outcomes [14]–[17]. These deep learning networks based on an encoder-decoder structure extract deep-level features, which can improve the image edge recognition ability to a certain extent compared with low-level features. Recently, some scholars have applied these segmentation algorithms to the industrial field and have achieved good application outcomes. Wang et al. proposed a new VGGNet for feature extraction of coal dust images to obtain accurate segmentation images [18]. Singh et al. used UNet to segment river ice images to analyze ice particles [19]. In addition, Liu et al. proposed a novel industrial chip parameter identification method based on cascaded region segmentation for surface-mount equipment [20].

However, these RGB-based segmentation algorithms are still limited in the feature expression of complex objects. The practical information extracted from RGB in complex objects is limited, so it is still insufficient to achieve high precision even if the deep-level features are extracted. Therefore, the complex textural characteristics of BF materials pose a tremendous challenge for high-precision segmentation. Due to objective factors such as uneven illumination, the tiny grooves on surfaces with a certain depth are almost indistinguishable from edges and background in pixel value, resulting in over-segmentation [21]. The fundamental reason for the interference of the groove texture on edge recognition is that the RGB image, a projection on the particle surface, ignores the depth difference between the groove textures and the particle edges. Therefore, to overcome these difficult problems, this paper proposes three contributions:

1) An end-to-end online PSD system based on camera-LiDAR fusion and particle image segmentation is proposed.
2) A feature dual-recalibration (FDR) module consisting of independent recalibration and joint recalibration is proposed to recalibrate the feature weights of two cross-modal data in the self-channel and whole integrated level so that the model can learn more valuable information during fusion.
3) To reduce the error caused by missing edge particle pixels, an edge-recognition-based particle size calculation strategy (ERP) is proposed, and we complete ablation experiments on a field dataset constructed by joint calibration and annotation.

II. SYSTEM DESIGN AND DATA ACQUISITION

A. Overall System Design

The proposed system for the PSD of BF materials based on image segmentation consists of three parts: an RGB-laser data acquisition and processing system, an RGB-laser particle segmentation network (RLPNet), and a particle size calculation method, as shown in Fig. 2.

B. RGB-laser data acquisition and processing system

Hardware equipment and installation. Compared with industrial cameras, single-line laser scanners, high-precision ranging sensors, can obtain high-precision depth information. However, the depth map obtained has low resolution and a large number of information voids, which makes it lack the ability to express details. For LiDAR, it is not easy to participate in training as independent input information. Therefore, these two sensors must complement each other to achieve better outcomes. The RGB-laser data acquisition and processing system consists of an industrial camera, a single-line laser scanner, a light-filling device, sliding brackets,

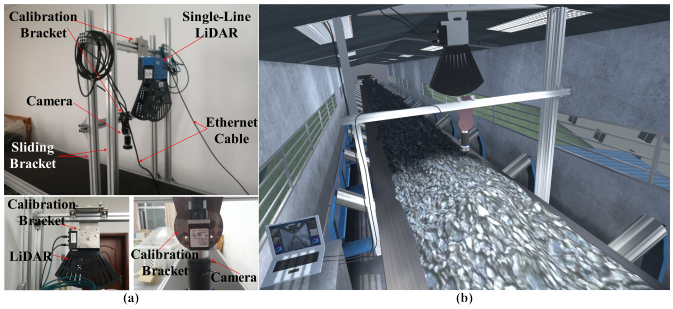


Fig. 2. Structure diagram of the RGB-laser data acquisition system. (a) Partial view of the acquisition system and (b) overall view of the acquisition system.

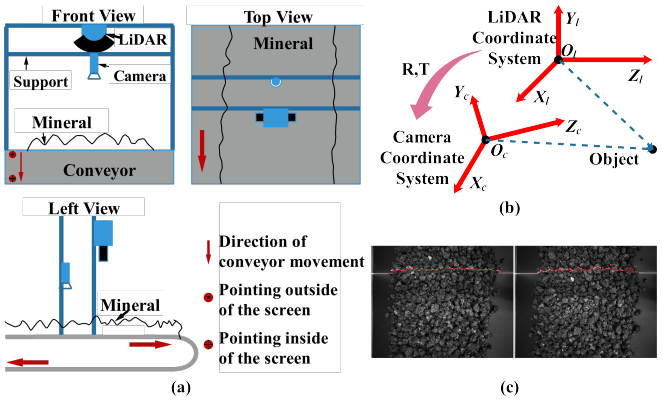


Fig. 3. Joint calibration of the camera and single-line laser scanner. (a) Three views of the acquisition system; (b) the external parameter calibration; and (c) image after calibration.

calibration brackets, Ethernet cables, and a computer, as shown in Fig. 2.

Joint calibration of the camera and LiDAR. Joint calibration of the camera and single-line laser scanner is the basis for obtaining RGB-laser datasets with consistent temporal and spatial distributions. It includes time synchronization, camera internal parameter calibration, and external parameter calibration. Time synchronization and camera internal parameter calibration can be completed offline. External parameter calibration of single-line LiDAR and industrial cameras solves the rotation and translation matrices that can convert the radar coordinate system into the camera coordinate system, as shown in Fig. 3. Specifically, the image's local highest point of the particle surface is selected as the marker point. Then, the corresponding data point pairs are obtained in the image coordinate system and the LiDAR coordinate system. Furthermore, they are formed into a linear system of equations containing at least three datasets. Finally, the PnP algorithm is applied to solve this system of equations to obtain the extrinsic parameter matrix [22].

RGB-laser Dataset of BF materials. More than 10000 2592×2048 RGB images and corresponding laser data were acquired by an industrial camera (ACA2500-14GM Basler) and a single-line laser scanner (LMS4111R-13000) on a conveyor belt moving at a constant speed. The acquisition frequency of the industrial camera with single-line LiDAR is approximately 60ms. To reduce the redundancy and duplication of data, 60 original pictures with almost no duplicated areas, called nonoverlapping pictures, were selected according to a fixed time interval (approximately 10s). Then, the pictures that were unclear, incomplete, and without laser data pictures were eliminated based on manual experience. The label drawing tool Labelme was used to manually label the 56 pictures manually, and the pixels were labeled background, edge, and material internal areas. Forty-two of the 56 original images were used to form the training set, 7 were used to form the validation set, and 7 were used to form the test set. To increase the diversity of the data, we performed five rounds of random cropping, rotation, and flipping for each image, where each image was expanded to 5*8=40 images. Therefore, 42 2592*2048 original images were expanded into 1680 256*256 training images, where 280 256*256 images are manually eliminated due to blurred imaging, too much background, insufficient light, and too many shadows. Finally, the remaining 1400 images formed the training set, where dense depth maps are obtained by complementing the laser maps [23].

III. SEGMENTATION NETWORK ARCHITECTURE

A. Network Backbone

The backbone of RLPNet consists of an RGB encoding branch, a laser encoding branch, a fusion branch, and the decoding process, as shown in Fig. 4. These encoding branches are all feature compression processes with a network-like structure. Each layer of these three branches consists of two groups of 3×3 convolution, batch normalization, and ReLU, i.e., ResNet-18 (R18) [24]. The fusion branch consists of 5 FDR layers, and the FDR input is the features on the RGB and Laser branches. The features in the encoder are concatenated into the features in the decoder through skip connections. Then, the 256*256 segmented images are obtained by four bottom-up sampling processes and 1×1 convolution.

B. Feature Dual-recalibration Module

*1) **Motivation**:* Due to the accuracy limitations of the interpolation and calibration algorithms, the depth information obtained by LiDAR may still contain some noise. Some scholars have proposed solutions to improve this problem in the direction of RGB-D feature fusion [25]. In [26], the relative localization of moving targets is completed through sensor fusion of the RGB-D camera and LiDAR. ACNet exploits an attention complementarity model that uses of depth information effectively [27]. RFNet uses SENet submodules to give independent weights to RGB and depth information in advance [28], [29]. However, these designs simplify the relationship between RGB and depth information. In other words, the interaction and correlation of RGB and depth information are only implicitly modeled. Later, a bilinear feature fusion CNN was proposed to recognize distributed tactile pressure in intelligent industries [30]. Xie et al. proposed an end-to-end CNN architecture to extract features from different images to detect the surface defects of magnetic tiles [31]. These methods have difficulty making good use of the

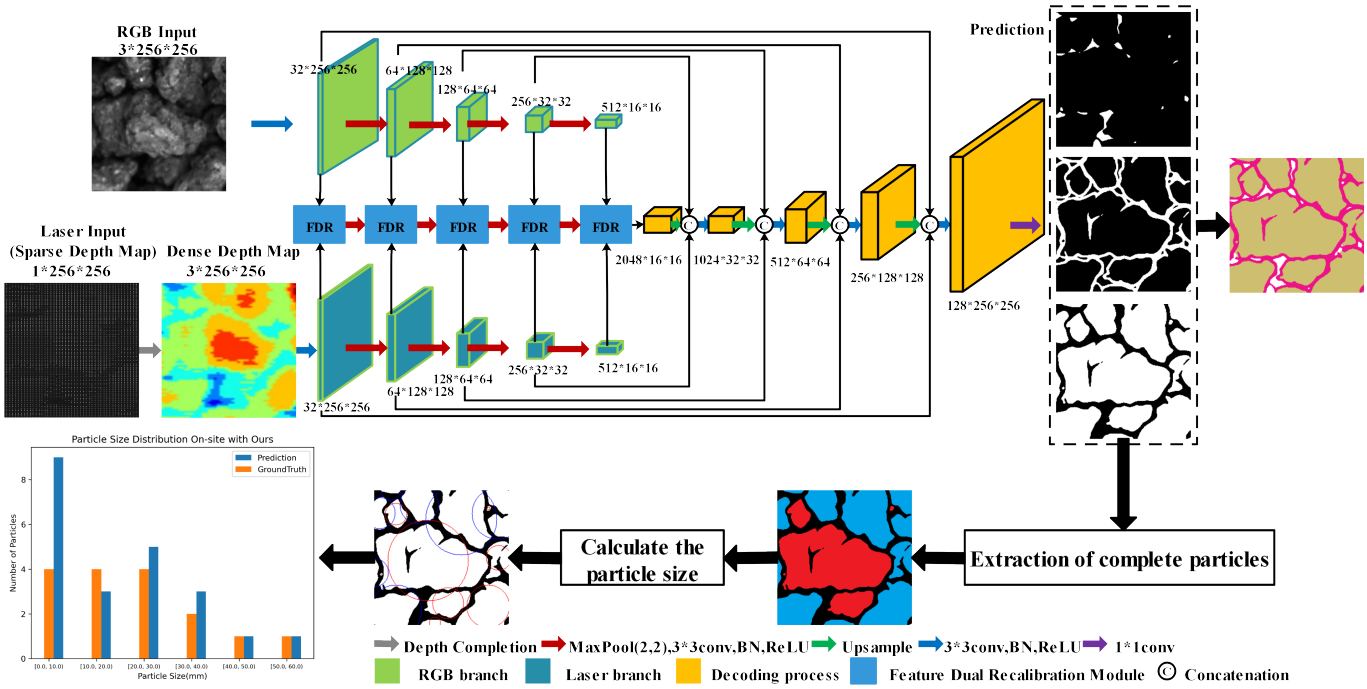


Fig. 4. The structure of the proposed detection system based on RGB-laser particle segmentation. The segmentation of BF materials images is to classify each pixel to determine whether the pixel belongs to the background, edge, or materials' internal area, labeled to 0, 1, and 2, respectively.

complementarity of the two kinds of features. Motivated by this, we design and embed a novel FDR module in RLPNet to obtain sufficient information during fusion, as shown in Fig. 5.

*2) **Independent Recalibration**:* FDR is divided into independent recalibration (IR) and joint recalibration (JR). The purpose of IR is to recalibrate the importance distribution of each feature channel at the self-channel level before feature fusion. Specifically, a C×H×W feature map is first compressed in the spatial dimension, turning each 2D feature channel into a real number. This C×1×1 sequence of real numbers has a global receptive field, representing a global information distribution on the channel. This can be achieved by global average pooling on the two feature maps, denoted $RGB_{in}, LAS_{in} \in \mathbb{R}^{C\times H\times W}$.

$$S_l = F_{gap}\left(LAS_{in}\right) = \frac{1}{H\times W}\sum_{i=1}^{H}\sum_{j=1}^{W} LAS_{in}(i,j) \tag{1}$$

where F_{gap} refers to global average pooling, and S_l represents the global descriptors of the laser feature map on the channel.

A bottleneck structure containing two fully connected (FC) layers is adopted to further learn the nonlinear relationship between channels, which reduces the model complexity and improves the generalization ability. The advantages of choosing a bottleneck structure consisting of two FC layers to model the correlation between channels are twofold. 1. Compared with using only one FC layer, the learning process of this network has more nonlinearity and can better fit the complex correlation between channels. 2. This design also reduces the number of parameters and the computational effort of the network compared to two FC layers without dimensionality reduction. Sufficient descriptions and experimental proofs are provided for this problem in [32]. The first FC layer plays the role of dimensionality reduction, while the second FC layer aims to restore the original dimension. Then, the independent attention vector for the laser input is learned by

$$W_l = F_{2fc}(S_l, w_l) = \sigma\left(w_l^1\delta\left(w_l^0 S_l\right)\right) \tag{2}$$

where F_{2fc} refers to two FC layers, and w_l^0 and w_l^1 represent the parameters in the two FC layers. σ denotes the sigmoid function, and δ refers to ReLU. The network can select more important channels and suppress less important channels by assigning different weights. Then, a less noisy feature representation is obtained by channel-wise multiplication between the input feature map LAS_{in} and the channel weights W_l, taking the laser feature map as an example.

$$LAS_{ir} = LAS_{in} \otimes W_l \tag{3}$$

where $\otimes$ refers to channel-wise multiplication. These two feature maps can be independently recalibrated to form a feature representation with less noise from the perspective of their channel distribution. In addition, it is also necessary to highlight the differences in the contribution of each feature map in the whole integrated level to facilitate fusion. Therefore, a second recalibration is designed to recalibrate the weights of these two features according to the contribution of each feature after IR, denoted joint recalibration.

*3) **Joint Recalibration**:* First, to fully exploit the complementarity of RGB and a laser, the fused features must be aggregated in a complementary manner at a certain location in space according to their representational capabilities. To achieve this, the spatial gates for RGB_{in} and LAS_{in} are generated to control the information flow of each modality

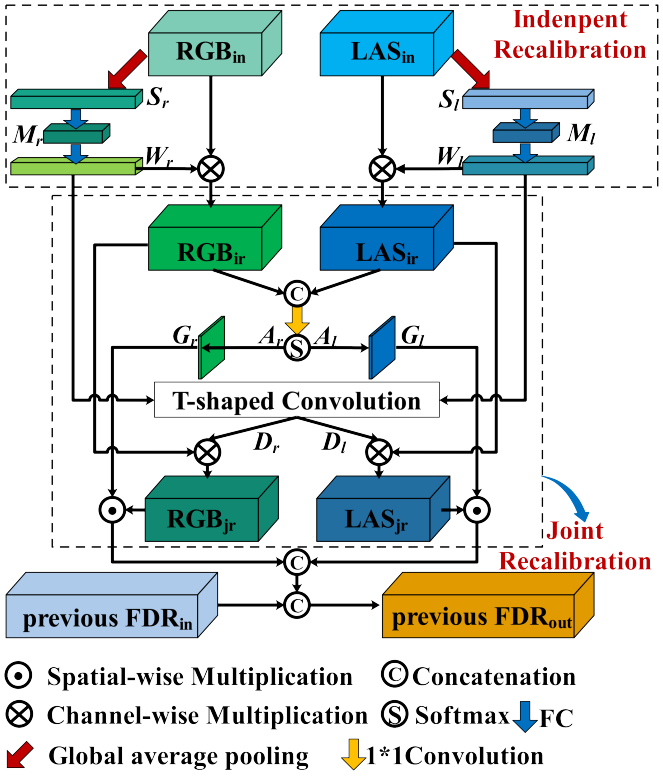


Fig. 5. Structure of the feature dual-recalibration module. FDR can recalibrate the feature weights at the self-channel level and whole integrated level for the input RGB and laser features. Its output is the fusion feature map for this weight distribution, which provides more effective input for subsequent decoders. The dimension of both M_r and M_l are $C/r * 1 * 1$. From the experimental results, this bottleneck design is able to improve the segmentation accuracy by sacrificing some efficiency. A balance between accuracy and efficiency can be found with a suitable setting of r. r indicates the scaling factor, we set $r = 16$.

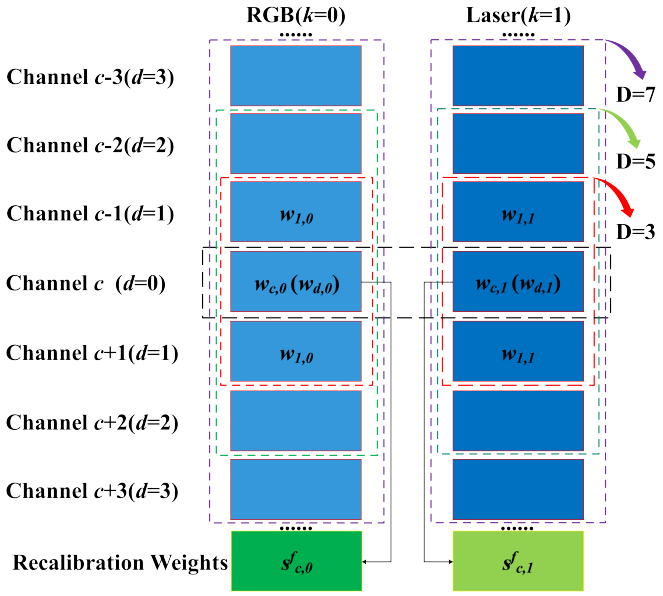


Fig. 6. TC is composed of a D×1 convolution and a 1×W convolution, where D depends on the neighborhood size d of the current channel (D=2d+1) and W indicates the number of feature maps participating in feature fusion (W=2). The value of D determines the perceptual field of the T-shaped convolution. The problem of choosing the size of the perceptual field of T has some similarities to the problem of choosing the size of the standard convolution kernel. The minimum parameters and computational effort are required for D=3. The value of D determines which nearby channels are computed with the current channel. D=3 means selecting the channel closest to the current channel for operation, that is, the previous channel and the next channel most related to the current channel. Channels that are farther from the current channel are less correlated with the current channel, so ignoring the farther channels in our calculations will not affect the results much.

feature map using a soft attention mechanism [33]. These two feature maps are first concatenated and then mapped to two different spatial gates through a 1×1 convolution:

$$A_l = F_{1\times1con}(RGB_{ir} \oplus LAS_{ir}) \tag{4}$$

$$G_l(i,j) = \frac{e^{A_l(i,j)}}{e^{A_r(i,j)} + e^{A_l(i,j)}} \tag{5}$$

Similarly, it requires adjusting the fusion feature weight from the channel level according to the different contributions in the whole integrated level. [34] concludes that the location information is encoded according to the order of channels in feature extraction. Thus, if the positional distributions of the RGB and laser input in the spatial dimension correspond, their feature maps also have a certain similarity in the corresponding channels. Moreover, the sequential encoding of channels also indicates that there must be some connection between each channel and nearby channels. Therefore, a T-shaped convolution (TC) is designed to adjust the fused feature weights of RGB and laser again, as shown in Fig. 6.

$$D_l^t(c) = \lambda_{l,c}W_l(c) + \lambda_{r,c}W_r(c) + \sum_{i=-d}^{(D-1)/2} \lambda_{l,c+i}W_l(c+i) \tag{6}$$

$$D_l(c) = \frac{D_l^t(c)}{D_l^t(c) + D_r^t(c)} \tag{7}$$

where $D_r^t(c)$ and $D_l^t(c)$ represent the intermediate outcomes of the c-th channel after the weighted operation of $W_r(c)$ and $W_l(c)$. $D_r(c)$ and $D_l(c)$ are the outcomes of linear changes for $W_r(c)$ and $W_l(c)$, which represent the outcomes of the recalibration of $W_r(c)$ and $W_l(c)$ from the perspective of the whole integration. λ refers to the intermediate parameters that need to be trained. Then, the weighted feature map LAS_{jr} can be obtained by a channel-wise multiplication between the feature maps LAS_{ir} and the fusion weight D_l.

$$LAS_{jr} = LAS_{ir} \otimes D_l \tag{8}$$

Finally, the fusion feature map M_{out} can be obtained by weighting RGB_{jr} and LAS_{jr} at the spatial level.

$$M_{out} = RGB_{jr} \odot G_r + LAS_{jr} \odot G_l \tag{9}$$

where $\odot$ refers to spatial-wise multiplication. To fully use the feature information of different scales, the fusion feature maps of each layer obtained by FDR are integrated as the first layer input of the decoding process. Additionally, the RLPNet is trained with a cross-entropy function.

Algorithm 1 Particle size distribution calculation

Input: A segmented image I_s
Output: Particle size distribution $R = (r_1, r_2, \ldots, r_n)$

$P, N \leftarrow f_{cd}(I_s)$
//Extracts and counts all connected domains in the image
for each $i \in [1, N]$ **do**
  if RECOGNITION$(p_i, size(I_s))$ **then**
    $d_i \leftarrow f_{MCD}(p_i)$
  else
    $d_i \leftarrow f_{heywood}(p_i)$
  end if
end for
$R \leftarrow f_{pss}(D)$ //f_{pss} represents the particle size statistic
function RECOGNITION$(Array, size)$
  //Input a particle contour array and image size
  $length = size[0]$
  $width = size[1]$
  for each $j \in [0, len(Array))$ **do**
    $x = Array[j][0]$
    $y = Array[j][1]$
    if $x * y * (x - length) * (y - width) = 0$ **then**
      return $True$ //Edge particle
    else
      return $False$ //Complete particle
    end if
  end for
end function

IV. PARTICLE SIZE DISTRIBUTION CALCULATION

After obtaining the segmented image, an efficient particle size calculation (PSC) method based on segmented images is needed to quantify the final result. A single particle's size representation method is related to the detection method. Currently, the main representation methods are triaxial diameter, projection diameter, spherical equivalent diameter, etc., as shown in Fig. 7. The triaxial diameter calculation method can be used to describe the particle size for detection methods that can obtain the particles' three-dimensional shape. However, the PSD method based on image segmentation essentially computes the size in the projection of the particle in a specific direction. Therefore, the projected diameter is used in this paper to calculate the single particle size. The projected diameter mainly includes the minimum circumcircle diameter (MCD), Feret, Martin, and Heywood [35]. Heywood, $d_H = \sqrt{\frac{4A}{\pi}}$, calculates the particle size by converting the projected area of an object into a circle of equal area, where A refers to the projected area.

Due to the limited field of view of the camera, the particles at the image's border are missing some pixels and are recorded as edge particles. The pixels missing from edge particles have significant errors in their particle size calculations, so the same calculation method cannot be used for all particles. An ERP is proposed to solve this problem, as shown in Algorithm. 1. We judge whether a particle is an edge particle by determining whether at least two of the contour points of each particle appear at the image edges.

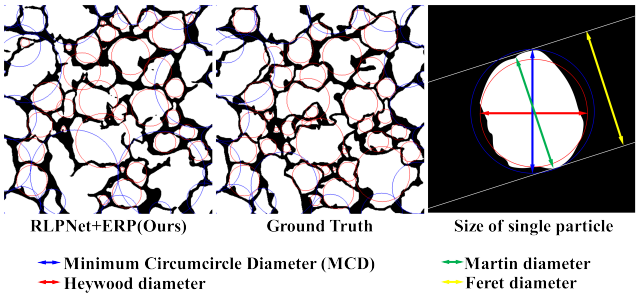


Fig. 7. Calculation of particle size. Martin diameter refers to the length of the line segment that bisects the projected area in a specific direction. Feret diameter refers to the spacing of 2 parallel lines in the same direction that holds the particle.

We assume that the i-th particle size interval is $[t * s_0, (t+1) * s_0)$ $(t \in \mathbb{N})$, and the number of particles in the i-th interval is N_i, where the largest particle size in the measured particles is s_{max}. There, the particle size distribution of the measured particle pile is as follows:

$$r_i = \frac{N_i}{\sum_{j=1}^{s_{\max}//s_t+1} N_j} \tag{10}$$

where $//$ refers to a division operation with rounding down and r_i represents the ratio of the amount of i-th interval particles to the total particles as a percentage. s_0 represents the size of the artificially determined particle level interval.

The mass distribution of the particles, denoted $M = (m_1, m_2, \ldots, m_n)$, is obtained from field acquisition, while the quantity distribution of the particles is obtained based on the segmented image, denoted $R = (r_1, r_2, \ldots, r_n)$. Therefore, we also need to establish a conversion relationship between these two.

$$m_i = \frac{\sum_{j-1}^{N_i} w_j}{M} = \frac{\sum_{j-1}^{N_i} f_m(d_j)}{N\bar{w}} = \frac{\sum_{j-1}^{N_i} f_m(d_j)}{\frac{N_i}{r_i} \cdot \frac{1}{N_i} \sum_{j-1}^{N_i} f_m(d_j)} = r_i \tag{11}$$

where $\bar{w}$ refers to the mean mass of the particles and M indicates the total mass of the pile of particles. $f_m(d)$ represents the relationship between particle size and mass; if the particles are considered spheres, then $f_m(d) = \frac{1}{6}\pi\rho d^3$. N denotes the total quantity of the pile of particles, and N_i refers to the number of particles in the ith particle level interval.

V. EXPERIMENTAL VERIFICATION

A. Verification Analysis of the Segmentation Algorithm

We prioritize cutting-edge algorithms that have structures similar to ours, and they are widely compared, including RGB-based segmentation algorithms (UNet, PSPNet, and DeepLabv3) and RGB-D-based segmentation algorithms (RedNet, RDFNet, RFNet, and ACNet), as shown in Fig. 8 and Tab. I. Since our individual encoder branches have some structural similarities to networks such as UNet, we first chose them as the baseline for our RGB-based segmentation algorithm.

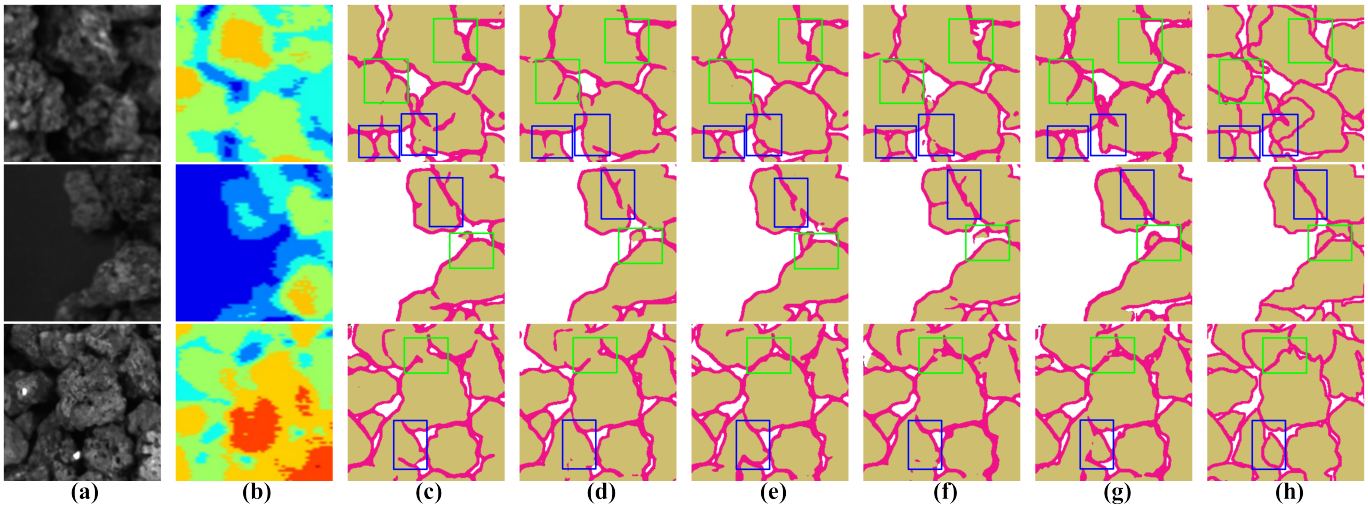


Fig. 8. Segmentation outcomes of part of the test dataset. The locations marked by blue boxes indicate locations where the RGB-D-based algorithms are significantly better than the RGB-based algorithms, and the locations marked by green boxes indicate where RLPNet is significantly better than the mainstream RGB-D algorithms. (a)RGB; (b)laser; (c)DeepLabv3; (d)UNet; (e)RFNet; (f)ACNet; (g)RLPNet (Ours); (h)Ground truth.

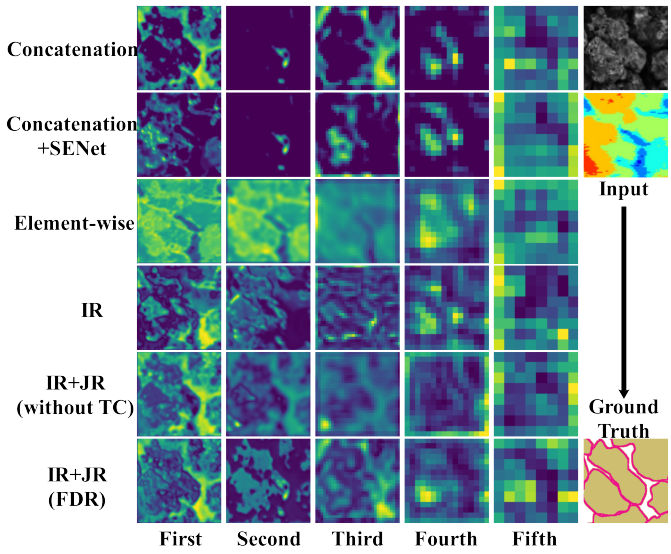


Fig. 9. Comparison with the visualized feature maps of the mainstream feature fusion methods. The five layers in this Figure represent the output features of the corresponding five-layer FDR in RLPNet.

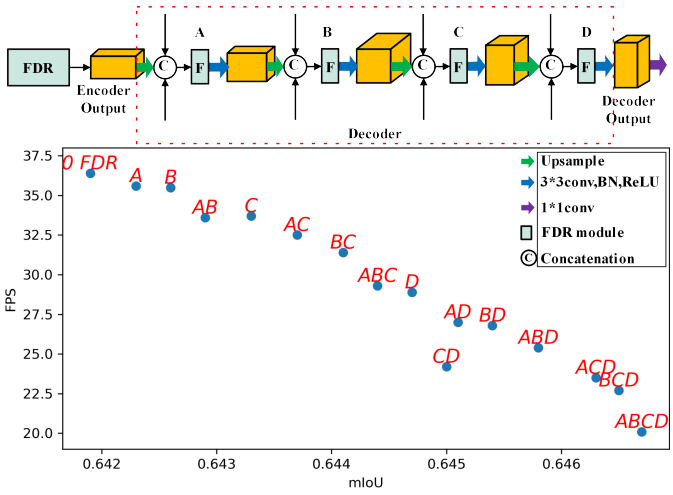


Fig. 10. Comparison of accuracy and efficiency by adding different FDR combinations. As more FDRs are added, mIoU gradually increases, but by a smaller amount, while FPS gradually decreases and by a larger amount.

RFNet, a real-time segmentation network, was chosen as the baseline for evaluating the efficiency of our algorithm. ACNet is a type of encoder-decoder framework and uses ResNet in its backbone, so it is used to evaluate the accuracy of RGB-D-based segmentation.

In the blue boxes, RLPNet generates more accurate connected domains and edges than RGB-based segmentation methods such as UNet, mainly thanks to reliable laser depth information, as shown in Fig. 8. The green boxes show that depth information does not always complement RGB well and may even reduce the overall accuracy. However, RLPNet-R18 has no noticeable wrong edges in the green box and can achieve better segmentation than other algorithms. The segmentation result of RLPNet-R18 reaches 64.2% on mIoU, which is 1.75% and 0.93% better than that of DeepLabv3 and that of ACNet, respectively. Compared with RLPNet-R18, RLPNet-R34 and RLPNet-R50 can only improve by 0.99% and 1.15%, respectively. Regarding algorithmic efficiency, the FPS of RLPNet-R18 is very close to that of the real-time segmentation network RFNet and slightly higher than that of ACNet. However, the efficiencies of RLPNet-R34 and RLPNet-R50 are considerably lower.

B. Verification Analysis of the Fusion Strategy

To further verify the efficient fusion capability of FDR, we replace the position of FDR in RLPNet with several mainstream fusion modules, including concatenation, element-wise add, concatenation+SENet, IR, and JR without TC. Fig. 9 and Tab. I show that the feature maps obtained by FDR are visually closer to the labels, which is also proven by comparing

TABLE I
ACCURACY AND EFFICIENCY COMPARISON OF PARTICLE SIZE DISTRIBUTION UNDER DIFFERENT ALGORITHMS

	Method	Backbone	D[1]	PSC	Segmentation Accuracy				Final Distribution Accuracy				FPS
					Background	Edge	Ore	mIoU	RS	eKL	eJS	mean	
RGB	UNet [14]	-	-	ERP	0.6096	0.4243	0.8309	0.6216	0.6225	0.7013	0.8337	0.7192	10.9
	PSPNet [36]	-	-	ERP	0.6140	0.4252	0.8301	0.6231	0.6411	0.7117	0.8400	0.7310	9.2
	DeepLabv3 [37]	-	-	ERP	0.6159	0.4262	0.8311	0.6244	0.6528	0.7296	0.8515	0.7447	7.6
RGB-D	RedNet [25]	R18[2]	-	ERP	0.6096	0.4319	0.8374	0.6263	0.6727	0.7180	0.8436	0.7448	9.6
	RDFNet [38]	R18	-	ERP	0.6119	0.4321	0.8382	0.6274	0.6885	0.7277	0.8495	0.7553	1.2
	RFNet(Concatenation+SENet) [28]	R18	-	ERP	0.6087	0.4294	0.8375	0.6252	0.6541	0.7073	0.8371	0.7329	9.3
	ACNet [27]	R18	-	ERP	0.6241	0.4356	0.8381	0.6326	0.7201	0.7350	0.8536	0.7696	7.2
	RLPNet(Ours)	R18	3	MCD					0.5198	0.6429	0.7861	0.6497	7.5
	RLPNet(Ours)	R18	3	HEY	**0.6394**	**0.4394**	**0.8469**	**0.6419**	0.3739	0.6134	0.7663	0.5846	7.5
	RLPNet(Ours)	**R18**	3	**ERP**					**0.7675**	**0.7477**	**0.8606**	**0.7919**	**7.4**
	RLPNet-JR(Without JR)	R18	-	ERP	0.6298	0.4302	0.8258	0.6286	0.7043	0.7249	0.8474	0.7589	9.1
	RLPNet-TC(Without TC)	R18	-	ERP	0.6359	0.4321	0.8397	0.6359	0.7402	0.7409	0.8568	0.7793	7.4
	RLPNet-con(with concatenation)	R18	-	ERP	0.6252	0.4230	0.8136	0.6206	0.5880	0.6949	0.8300	0.7044	10.3
	RLPNet-ele(with Element-wise)	R18	-	ERP	0.6277	0.4197	0.8096	0.6190	0.5407	0.6741	0.8170	0.6773	9.4
	RLPNet-FDR(Add 4 FDRs to decoder)	R18	3	ERP	0.6401	0.4525	0.8475	0.6467	0.7806	0.7503	0.8619	0.7976	4.8
	RLPNet-D7(Ours)	R34	7	ERP	0.6600	0.4525	0.8450	0.6525	0.7895	0.7570	0.8659	0.8042	3.0
	RLPNet-D5(Ours)	R34	5	ERP	0.6593	0.4524	0.8449	0.6522	0.7790	0.7564	0.8658	0.8004	3.2
	RLPNet(Ours)	R34	3	ERP	0.6584	0.4523	0.8447	0.6518	0.7854	0.7539	0.8641	0.8011	3.3
	RLPNet(Ours)	R50	3	ERP	0.6590	0.4627	0.8489	0.6534	0.7921	0.7591	0.8672	0.8062	1.1

1. D is the kernel size of TC proposed in this paper. 2. R18 indicates ResNet-18, R34 indicates ResNet-34, and R50 indicates ResNet-50.

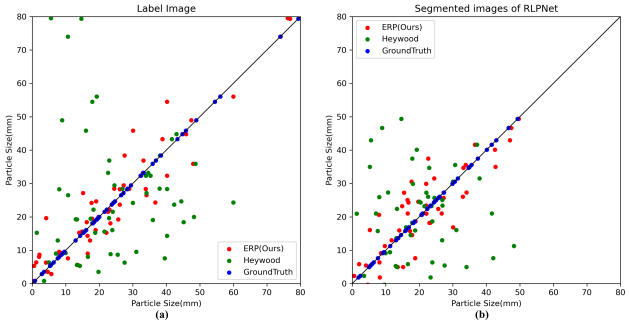


Fig. 11. Comparative graph of particle size calculation results. (a) Comparison of particle size calculation results based on labeled images and (b) comparison of particle size calculation results based on segmented images of RLPNet. The blue dots indicate the true results, and the red dots indicate the results calculated by ERP, and the green dots indicate the results calculated by Heywood. The distribution of the red dots can be observed through the 45-degree line to be closer to the distribution of the blue dots than the distribution of the green dots, i.e., the results of ERP are closer to the labels.

segmentation accuracy. In addition, we did not add FDR to the decoder for the following reasons. 1. The output features of the first layer of the decoder have already been denoised by FDR, and adding FDR to the decoder again is methodologically repetitive and redundant. 2. The accuracy improvement is not apparent, and the efficiency decreases more. We tested the effect of adding 4-layer FDR in the decoder, which only improves the mIoU by 0.14% compared with RLPNet-R18 but decreases the efficiency by 2.6FPS, as shown in Fig. 10.

We tested the effect when D=3, 5, and 7 in TC. As stated in Tab. I, the accuracies of RLPNet-R34 at D=5 and 7 are only 0.04% and 0.03% higher, respectively, than the accuracy at D=3, and its efficiency is still decreasing.

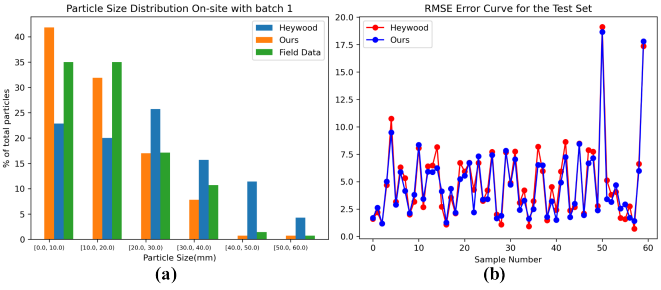


Fig. 12. Particle size distribution calculation results and error analysis of different batches using ERP and Heywood. (a) Particle size distribution; and (b) RMSE.

C. On-site Verification of Particle Size Distribution

Not only will the segmentation accuracy affect the final particle size distribution accuracy, but the particle size calculation method will also have a considerable impact on the final particle size distribution.

To quantitatively assess the accuracy of the final particle size distribution, we used the coefficient of determination ($RS = \frac{\sum_{i=1}^{n} w_i(\hat{y}_i - \bar{y}_i)^2}{\sum_{i=1}^{n} w_i(y - \bar{y}_i)^2}$), KL-divergence ($KL = \sum_{x \in X} P(x) \log \frac{P(x)}{Q(x)}$), and JS-divergence ($JS = \frac{1}{2}(KL(P||\frac{P+Q}{2}) + KL(Q||\frac{P+Q}{2}))$) to calculate the distance between the predicted particle size distribution and the actual particle size distribution [39]. $\hat{y}_i$ represents the predicted value, and $\bar{y}_i$ denotes the mean of the true value. P and Q indicate the two different distributions. $eKL = e^{-\sqrt{KL}}$ and $eJS = e^{-\sqrt{JS}}$ are exponential transformations of KL-divergence and JS-divergence, respectively.

To verify the accuracy of ERP, we obtained segmentation results with RLPNet-R18 and then calculated the particle size distribution with ERP, Heywood, and MCD. Fig. 11 shows the size of each particle for a given sample calculated by the combination of RLPNet-R18 and ERP with Heywood. The

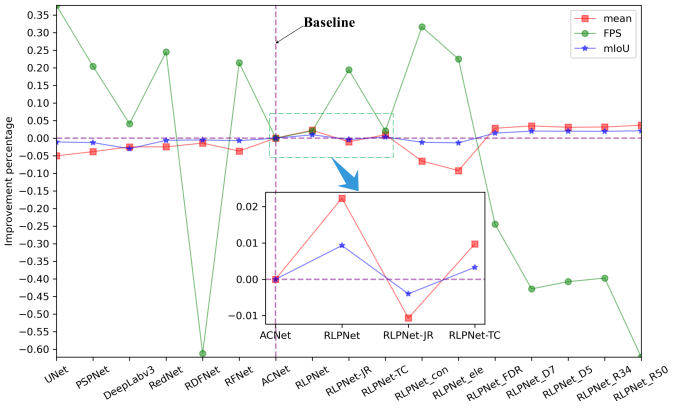


Fig. 13. Improvement percentage in the final particle size distribution obtained by different segmentation algorithms compared to ACNet. *mean* indicates the average accuracy of the final particle size distribution. The purple dashed line indicates our baseline, i.e., the performance of ACNet. For easier observation, we have zoomed in on the green boxed area. The green boxed area shows the comparison results of the four methods ACNet, RLPNet, RLPNet-JR and RLPNet-TC. The small picture in the middle is an enlarged version of the comparison between RLPNet and ACNet.

statistics of the calculated results for the subsample sets were obtained from Fig. 12(a) and Tab. I. Fig. 12(b) shows the error results obtained for the calculations. The accuracy of the overall particle size calculation is improved because ERP considers more incomplete edge particles. Therefore, the results in Fig. 12 and Tab. I show that the particle size distribution obtained by ERP is closer to the actual distribution.

Fig. 13 quantitatively visualizes the effect of enhancing the segmentation algorithm on the final particle size distribution. The final particle size distribution obtained by RLPNet-R18 achieves 76.75%, 74.77% and 86.06% accuracies on RS, eKL, and eJS, respectively, which are 4.74%, 1.27% and 0.70% improvements compared to those of ACNet, respectively. The overall efficiency of RLPNet-R18 is also very close to that of the real-time RGB-D segmentation network RFNet, with a decrease of only 1.9 FPS.

VI. CONCLUSION

Unlike traditional particle size detection methods, we propose a new particle size detection system based on RGB-laser particle segmentation, which can calculate the particle size distribution of blast furnace materials with high efficiency and accuracy. First, to extract particle targets in BF material images with high accuracy, an RGB-laser-based particle segmentation network is proposed to reduce texture interference in edge recognition. Moreover, to further enhance the expression of edge features, a feature dual-recalibration module is designed and embedded in the particle segmentation network to improve the fusion efficiency of RGB and laser features. Through independent and joint recalibration, the two cross-modal features are efficiently fused from the self-channel and whole integration levels. In addition, to reduce the error caused by missing edge particle pixels, an edge-recognition-based particle size calculation strategy is proposed. In the experimental verification part, the effectiveness and high precision of the proposed algorithm are first verified comparison with the mainstream segmentation algorithms, which achieve an accuracy of 64.19% in mIoU. Then, by comparison with other mainstream feature fusion strategies through ablation experiments, the strong fusion ability of feature dual-recalibration for cross-modal features is verified. Finally, the mean similarity between the particle size distribution predicted by the proposed method and the actual distribution reaches 79.19%.

REFERENCES

[1] M. Geerdes, R. Chaigneau, and O. Lingiardi, *Modern Blast Furnace Ironmaking: An Introduction (2020)*. Ios Press, 2020.

[2] P. Zhou, R. Zhang, J. Xie, J. Liu, H. Wang, and T. Chai, "Data-driven monitoring and diagnosing of abnormal furnace conditions in blast furnace ironmaking: An integrated pca-ica method," *IEEE Transactions on Industrial Electronics*, vol. 68, no. 1, pp. 622–631, 2020.

[3] P. Zhou, Y. Lv, H. Wang, and T. Chai, "Data-driven robust rvflns modeling of a blast furnace iron-making process using cauchy distribution weighted m-estimation," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 9, pp. 7141–7151, 2017.

[4] H. Wei, C. Shusen, X. Wenxuan, and B. Jiawen, "Effect of burden-size on burden-distribution and air permeability in high furnace surface," *Journal of Iron and Steel Research*, vol. 30, no. 10, pp. 769–775, 2018.

[5] J.-I. Park, U.-H. Baek, K.-S. Jang, H.-S. Oh, and J.-W. Han, "Development of the burden distribution and gas flow model in the blast furnace shaft," *ISIJ international*, vol. 51, no. 10, pp. 1617–1623, 2011.

[6] K. Yang, S. Choi, J. Chung, and J.-i. Yagi, "Numerical modeling of reaction and flow characteristics in a blast furnace with consideration of layered burden," *ISIJ international*, vol. 50, no. 7, pp. 972–980, 2010.

[7] W. Hartig, K. Langner, H. B. Lungen, and K. P. Stricker, "Measures for increasing the productivity of blast furnaces," *Metallurgical Plant and Technology International(Germany)*, vol. 4, pp. 42–44, 1996.

[8] J. Zheng, P. F. Johnson, and J. S. Reed, "Improved equation of the continuous particle size distribution for dense packing," *Journal of the American Ceramic Society*, vol. 73, no. 5, pp. 1392–1398, 1990.

[9] Z. Pengchao, "Experimental analysis of air permeability of blast furnace block zone," *China metallurgical*, vol. 30, no. 8, pp. 10–14, 2020.

[10] J. P. Schwarz, R. Gao, D. Fahey, D. Thomson, L. Watts, J. Wilson, J. Reeves, M. Darbeheshti, D. Baumgardner, G. Kok *et al.*, "Single-particle measurements of midlatitude black carbon and light-scattering aerosols from the boundary layer to the lower stratosphere," *Journal of Geophysical Research: Atmospheres*, vol. 111, no. D16, 2006.

[11] L. Zhang and J. Kulon, "Comparative study of square and sine-wave excitation methods for the measurement of aerosol particles charge and size distribution using phase doppler anemometry," *IEEE Transactions on Instrumentation and Measurement*, vol. 60, DOI 10.1109/TIM.2010.2082498, no. 4, pp. 1397–1407, 2011.

[12] R. J. Bomphrey, N. J. Lawson, N. J. Harding, G. K. Taylor, and A. L. Thomas, "The aerodynamics of manduca sexta: digital particle image velocimetry analysis of the leading-edge vortex," *Journal of Experimental Biology*, vol. 208, no. 6, pp. 1079–1094, 2005.

[13] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 3431–3440, 2015.

[14] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, pp. 234–241. Springer, 2015.

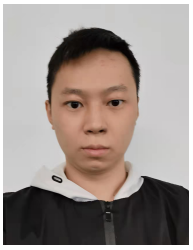
[15] V. Badrinarayanan, A. Kendall, and R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 12, pp. 2481–2495, 2017.

[16] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," *IEEE transactions on pattern analysis and machine intelligence*, vol. 40, no. 4, pp. 834–848, 2017.

[17] C. Peng, X. Zhang, G. Yu, G. Luo, and J. Sun, "Large kernel matters–improve semantic segmentation by global convolutional network," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4353–4361, 2017.

[18] Z. Wang, D. Li, X. Zheng, and D. Xie, "A novel coal dust characteristic extraction to enable particle size analysis," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–12, 2021.

[19] A. Singh, H. Kalke, M. Loewen, and N. Ray, "River ice segmentation with deep learning," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 11, pp. 7570–7579, 2020.
[20] W. Liu, X. Yang, X. Yang, and H. Gao, "A novel industrial chip parameters identification method based on cascaded region segmentation for surface-mount equipment," *IEEE Transactions on Industrial Electronics*, vol. 69, DOI 10.1109/TIE.2021.3082072, no. 5, pp. 5247–5256, 2022.
[21] K. M. Rao and A. Dempster, "Modification on distance transform to avoid over-segmentation and under-segmentation," in *International Symposium on VIPromCom Video/Image Processing and Multimedia Communications*, pp. 295–301. IEEE, 2002.
[22] H. Zhang, Z. Liang, C. Li, H. Zhong, L. Liu, C. Zhao, Y. Wang, and Q. M. J. Wu, "A practical robotic grasping method by using 6-d pose estimation with protective correction," *IEEE Transactions on Industrial Electronics*, vol. 69, DOI 10.1109/TIE.2021.3075836, no. 4, pp. 3876–3886, 2022.
[23] Y. Zhang, Q. Fan, F. Bao, Y. Liu, and C. Zhang, "Single-image super-resolution based on rational fractal interpolation," *IEEE Transactions on Image Processing*, vol. 27, DOI 10.1109/TIP.2018.2826139, no. 8, pp. 3782–3797, 2018.
[24] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.
[25] J. Jiang, L. Zheng, F. Luo, and Z. Zhang, "Rednet: Residual encoder-decoder network for indoor rgb-d semantic segmentation," *arXiv preprint arXiv:1806.01054*, 2018.
[26] H. Song, W. Choi, and H. Kim, "Robust vision-based relative-localization approach using an rgb-depth camera and lidar sensor fusion," *IEEE Transactions on Industrial Electronics*, vol. 63, DOI 10.1109/TIE.2016.2521346, no. 6, pp. 3725–3736, 2016.
[27] X. Hu, K. Yang, L. Fei, and K. Wang, "Acnet: Attention based network to exploit complementary features for rgbd semantic segmentation," in *2019 IEEE International Conference on Image Processing (ICIP)*, pp. 1440–1444. IEEE, 2019.
[28] L. Sun, K. Yang, X. Hu, W. Hu, and K. Wang, "Real-time fusion network for rgb-d semantic segmentation incorporating unexpected obstacle detection for road-driving images," *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 5558–5565, 2020.
[29] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 7132–7141, 2018.
[30] J. Chu, J. Cai, L. Li, Y. Fan, and B. Su, "Bilinear feature fusion convolutional neural network for distributed tactile pressure recognition and understanding via visualization," *IEEE Transactions on Industrial Electronics*, vol. 69, DOI 10.1109/TIE.2021.3086714, no. 6, pp. 6391–6400, 2022.
[31] L. Xie, X. Xiang, H. Xu, L. Wang, L. Lin, and G. Yin, "Ffcnn: A deep neural network for surface defect detection of magnetic tile," *IEEE Transactions on Industrial Electronics*, vol. 68, DOI 10.1109/TIE.2020.2982115, no. 4, pp. 3506–3516, 2021.
[32] A. Howard, A. Zhmoginov, L.-C. Chen, M. Sandler, and M. Zhu, "Inverted residuals and linear bottlenecks: Mobile networks for classification, detection and segmentation," 2018.
[33] X. Chen, K.-Y. Lin, J. Wang, W. Wu, C. Qian, H. Li, and G. Zeng, "Bi-directional cross-modality feature propagation with separation-and-aggregation gate for rgb-d semantic segmentation," in *European Conference on Computer Vision*, pp. 561–577. Springer, 2020.
[34] M. A. Islam, M. Kowal, S. Jia, K. G. Derpanis, and N. D. Bruce, "Global pooling, more than meets the eye: Position information is encoded channel-wise in cnns," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 793–801, 2021.
[35] Y. Wu, G. Castle, and I. Inculet, "Particle size analysis in the study of induction charging of granular materials," *Journal of Electrostatics*, vol. 63, DOI https://doi.org/10.1016/j.elstat.2004.09.008, no. 3, pp. 189–202, 2005, selected Papers from the ESA 2004 Annual Conference. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0304388604001949
[36] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, "Pyramid scene parsing network," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2881–2890, 2017.
[37] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," in *Proceedings of the European conference on computer vision (ECCV)*, pp. 801–818, 2018.
[38] S. Lee, S.-J. Park, and K.-S. Hong, "Rdfnet: Rgb-d multi-level residual feature fusion for indoor semantic segmentation," in *2017 IEEE International Conference on Computer Vision (ICCV)*, pp. 4990–4999. IEEE, 2017.
[39] J. Goldberger, S. Gordon, H. Greenspan *et al.*, "An efficient image similarity measure based on approximations of kl-divergence between two gaussian mixtures." in *ICCV*, vol. 3, pp. 487–493, 2003.



Jinshi Liu received a bachelor's degree and a master's degree in automation from Central South University, China in 2017 and 2020, respectively. He is currently pursuing a Ph.D. in Control Engineering at Central South University. His research interests include deep learning, detection technology, image segmentation, industrial VR and digital twins.

Zhaohui Jiang received a M. Eng. degree in Automatic Control Engineering and a Ph.D. degree in Control Science and Engineering from Central South University, China in 2006 and 2011, respectively. He is currently a professor at Central South University. His research interests include Detection Technology and Automatic Equipment, image processing, industrial VR, modeling and optimal control of complex industrial processes.

Weihua Gui received the degree of the B.Eng. in Electrical Engineering and the M.Eng. in Automatic Control Engineering from Central South University, China in 1976 and 1981, respectively. From 1986 to 1988 he was a visiting scholar at Universität-GH-Duisburg, Germany. Since 1991, he has been a full professor in Central South University. Since 2013, he has been an Academician with the Chinese Academy of Engineering, China. His main research interests are in modeling and optimal control of complex industrial process, distributed robust control, and fault diagnoses.

Zhiwen Chen (M'17) received his B.S. degree in electronic information science and technology and M.S. degree in electronic information and technology from Central South University, China in 2008 and in 2012, respectively, Ph.D. degree in electrical engineering and information technology from University of Duisburg-Essen, Germany in 2016. He is currently an associated professor at Central South University. His research interests are model-based and data-driven fault diagnosis and health monitoring, data analytics.