Physicochem. Probl. Miner. Process., 61(5), 2025, 211515

http://www.journalssystem.com/ppmp

Physicochemical Problems of Mineral Processing

ISSN 1643-1049

Advancing sustainable lithium mining: semi-industrial validation of the mineral-DeepLabV3+ AI-based XRT segmentation model

Tianyou Yu, Yimin Zhu, Jie Liu, Yuexin Han, Yanjun Li

School of Resources and Civil Engineering, Northeastern University, Shenyang110819, China

National-local Joint Engineering Research Center of High-efficient Exploitation Technology for Refractory Iron Ore Resources, Shenyang 110819, China

Corresponding author: zhuyimin@mail.neu.edu.cn (Yimin Zhu)

Abstract: The global demand rise in lithium, driven by the expansion of the new energy sector and electric vehicle markets, underscores the urgent need for efficient mineral processing of spodumene, the primary lithium ore. Traditional methods encounter challenges such as low-grade ores, fine particle sizes, complex mineral intergrowth, and high energy consumption with notable environmental impacts. This study addresses these issues by developing an intelligent X-ray Transmission (XRT) pre-sorting system enhanced with the novel Mineral-DeepLabV3+ deep learning algorithm. The approach integrates machine vision and artificial intelligence to deliver lightweight, precise ore segmentation under industrial mining conditions, incorporating innovations such as the MobileNetV4 backbone, optimized Atrous Spatial Pyramid Pooling module, and CSWin attention mechanism. Key findings demonstrate that the Mineral-DeepLabV3+ model achieves a mean intersection over union of 95.19 %, boosts segmentation accuracy by 3.28 %, and reduces parameter count by over 44 % compared to baseline models, while maintaining fast inference speed. Semi-industrial trials confirm its superior performance across varying ore sizes, processing rates, and conveyor speeds, achieving Li_2O concentrate grades up to 2.27% and waste rejection rates up to 66%. This technology significantly enhances resource efficiency, reduces environmental footprint, and advances operational sustainability in spodumene beneficiation. The proposed framework offers a scalable solution for driving low-carbon, intelligent practices in mineral processing.

Keywords: spodumene, lithium mining, X-ray transmission (XRT), deep learning; semantic segmentation

1. Introduction

Lithium, as the lightest alkali metal in nature, has unique physicochemical properties, including high specific heat, high electrical conductivity, and strong chemical reactivity (Xiang et al., 2023; Tang et al., 2020; Jung et al., 2020). As the new energy vehicle and energy storage industries rapidly develop, lithium is extensively applied in sectors such as metallurgy, nuclear energy, medicine, and high-energy batteries due to its characteristics (Jamesh, 2019; Gao et al., 2022). Due to the huge demand for lithium and its relatively scarce resources, this supply-demand contradiction has also promoted the advancement and development of lithium ore refining technologies both domestically and internationally. Currently, the main sources of lithium resources include spodumene, lepidolite, phosphor-lithium-alumino-silicate, clay lithium, salt lake brine, and the recycling of used lithium batteries. Among these, spodumene (LiAlSi2O6) ore has a simple chemical composition, an easily controllable extraction process, mature technology, and reliable product quality, making it the largest contributor in the current lithium extraction market (Liu et al., 2022; Song et al., 2019; Zhai et al., 2023).

However, the original ore resources of spodumene generally have low grades and complex compositions, with a grade of only Li_2O 1.02%–2.00%, resulting in high beneficiation and smelting costs, difficulties in efficient and green extraction technologies, and an increase in environmental risks (Greim et al., 2020). Additionally, spodumene typically coexists with other silicate minerals (feldspar, quartz,

DOI: 10.37190/ppmp/211515

beryl, and mica), complicating its properties and making the separation process highly intricate (Tabelin et al., 2021; Qiu et al., 2022). Furthermore, spodumene deposits are mainly found in granite pegmatite deposits, which have experienced extensive weathering and clay alteration. During mining, waste rock and surrounding rocks often mix with the ore and enter the beneficiation process. This not only makes the recovery and utilization of spodumene and its associated minerals more complicated, but also generates a large amount of tailings, consuming considerable energy in the beneficiation process and causing significant environmental impact (Han et al., 2022; Liu et al., 2023; Yelatontsev; Mukhachev, 2021; Sarker et al., 2022).

Common techniques used in the mineral processing of spodumene include manual selection, flotation, magnetic separation, and dense media separation. Currently, flotation has become the mainstream method due to its wide application. The commonly used flotation method currently involves the addition of Na_2CO_3, Na_2S, NaOH, naphthenic acid soap, and oxidized paraffin soap (Xie et al., 2021b; Xie et al., 2020; Shu et al., 2020; Shi et al., 2023). However, the main challenges in spodumene beneficiation include the low grade of spodumene ore, fine ore particle size, fine embedding particle size, and complex intergrowth relationships between minerals. These factors collectively increase the difficulty of spodumene separation, leading to issues with traditional single collectors, such as low recovery efficiency, poor selectivity, and excessive consumption, thereby limiting the development and utilization of spodumene resources. Conversely, intelligent preselection can exploit the differences in properties between spodumene and gangue minerals to discard some low-grade waste rocks before grinding, reducing the load on subsequent operations, enhancing the grade of selected spodumene ore, lowering energy consumption in the beneficiation process, and increasing sorting efficiency. Furthermore, preselection waste disposal also reduces water consumption and the amounts of fine-grained tailings in the beneficiation plant, thereby decreasing environmental pollution from beneficiation wastewater and the accumulation of fine tailings. Therefore, preselection of spodumene ore is beneficial for reducing costs, minimizing pollution, improving quality, and enhancing efficiency, making it an effective means of achieving maximal resource utilization (Xie et al., 2021a).

Presently, the preselection technologies employed in mining operations include color sorting, X-ray fluorescence (XRF), and X-ray transmission (XRT) technologies (Zheng, 1980; Ding et al., 2014; Wu, 2020). Color sorting technology is suitable for cases where there is a significant color difference between the ore and surrounding rock, but it is easily affected by ore particle size and the covering of ore surfaces by mineral sludge, which limits its effectiveness. XRF technology is limited because X-ray fluorescence only excites a depth of 1-2 mm on the surface of the ore, making it unable to reflect the grade information of the ore's interior. Additionally, when multiple elements coexist, their fluorescent signals may interfere with each other, often requiring large ore samples for testing and statistical analysis to determine the feasibility of sorting. XRT technology utilizes the differing X-ray attenuation capabilities of different atoms and sorts based on the atomic number differences in mineral compositions (Robben et al., 2020). It is less affected by factors such as ore surface dust, stains, and particle shape and size. Currently, XRT preselection technology is a hot research topic in the field of ore preselection (Li et al., 2023; Li et al., 2018). In Spain, Rafaella Resources' tungsten-tin intelligent pre-sorting project achieves a waste rock removal rate exceeding 50%. In Canada, Maritime's Hammerdown Gold Mine utilizes this technology to discard 34% of the feed as waste rock, while Osisko's Cariboo Gold Mine employs intelligent optical pre-selecting technology before separation, leading to a 50% reduction in water and energy consumption (2020a; 2021a); results from Vendetta Mining's San Rafael Mine, the world's largest underground tin mine in Peru, show that intelligent optical pre-sorting can increase the Sn grade from 0.6% to 2.8%, with a recovery rate of 90% (2020c; 2020d; 2020b); Additionally, Vast Resources' Baita Plai polymetallic mine in Romania uses intelligent sorting equipment to achieve a waste rock removal rate of approximately 40%, resulting in significant energy savings and efficiency improvements (2021b).

With advancements in artificial intelligence and big data processing technologies, XRT preselection technology has seen rapid progress in waste disposal accuracy and processing capacity. However, in the field of mine preselection, due to the brittleness of pegmatitic spodumene ores, significant particle size variations are caused during the crushing process. The ore often exists in a dispersed state, with some ores having extremely fine particles that are difficult to identify, while others are clumped together, making identification challenging. Moreover, due to the severe intergrowth phenomenon in

spodumene ores, the grayscale images obtained from XRT transmission of the ore show multiple complex intergrowth relationships, which demands very high accuracy for image segmentation (Li et al., 2021). As computer technology has advanced, intelligent ore block segmentation methods based on image processing techniques have been introduced. Many researchers have investigated ore block image segmentation methods. Traditional techniques, such as the watershed method and its optimized versions, thresholding, and theory-based segmentation approaches, are widely used in many contexts, but they often fail to deliver satisfactory results when applied to XRT images of ores. However, with the rapid growth of deep learning technology, this situation has seen considerable improvement. Deep learning, with its ability to automatically learn and extract key features from vast amounts of training data, has provided new opportunities for image segmentation research in the mining field, driving the in-depth application and study of this technology.

In recent years, many mining experts have conducted extensive research on deep learning methods for ore image segmentation. Among them, Yang et al. proposed an optimized U-Net network architecture in the field of ore image research, specifically using VGG-16 as the encoder part to improve the network's performance in ore image analysis (Yang et al., 2020). Although significant results were obtained on manually arranged ore datasets, the data source is still limited to artificially set environments. LU et al. analyzed five improved deep learning models for ore image segmentation, achieving ideal segmentation results, but the research focused on mineral segmentation in microscopic images (Liu et al., 2021). In their study, KOH et al. examined the performance of two models, Mask R-CNN and SOLOv2, in identifying grain boundaries and performing mineral classification tasks (Koh et al., 2021). Zhang et al. proposed a threshold segmentation method based on neighborhoods and Otsu, which binarizes each pixel using the minimum threshold to obtain the segmented image (Zhang et al., 2010). Although the above deep learning image segmentation algorithms have demonstrated outstanding learning abilities in image recognition, they are still significantly affected by other factors in practical applications. Especially in the segmentation task of ore XRT images. Firstly, existing algorithms depend on manually curated ore datasets, which lack sufficient diversity and the complexity of real-world environments, leading to poor generalization ability of the models in actual mining contexts. Additionally, ore XRT images often contain significant noise, particularly due to the interference of fine-grained ores and dust, which makes it difficult for deep learning algorithms to extract accurate features. Furthermore, mineral boundaries in ore XRT images are complex, with small differences between minerals, making it difficult for current models to precisely identify these intricate boundaries, especially in cases of multi-mineral coexistence. Moreover, deep learning models typically demand substantial computational resources, and as most mines have limited computational power, this further limits the effectiveness of algorithm training. Therefore, although deep learning technologies offer new opportunities for image segmentation research in the mining sector, algorithms for ore XRT images still require further refinement to improve their adaptability and robustness.

DeepLabV3+ is a deep learning model for semantic segmentation tasks proposed by the Google Brain team. The backbone network MobileNetV2 is a lightweight semantic segmentation model based on CNN. It achieves high accuracy while maintaining a small number of parameters and low computational cost, making it highly suitable for deployment on mining mobile platforms with limited computing resources (Chen et al.; Sandler et al.). While DeepLabV3+ can efficiently recognize mining environments in simple settings, it utilizes multiple down-sampling operations to achieve a larger receptive field and lower-resolution feature maps. These operations can cause the loss of fine details, especially for small and fine-grained target ores. In the segmentation of adhered ores, the recognition of blurred edges and occluded ores in XRT images is less effective, leading to reduced segmentation accuracy.

In conclusion, to tackle the challenges in spodumene XRT image segmentation, this study introduces a spodumene XRT image segmentation approach based on the Mineral-DeepLabV3+ model. The lightweight MobileNetV4 is used as the backbone feature extraction network to reduce the number of parameters and accelerate segmentation speed while maintaining good performance in high-speed preselection. Moreover, the optimized ASPP module introduces controllable functions and multi-scale fusion, facilitating denser pixel sampling and improving detailed feature extraction, leading to better recognition of adhered ores and fine-grained spodumene (Qin et al.; Dong et al.). Furthermore, the

CSWin attention mechanism is employed to enhance the model's focus on ore-related information, reducing background interference from dust and impurities, thereby improving segmentation accuracy. Moreover, comparative experiments, ablation studies, and semi-industrial trials on the XRT intelligent preselection machine were conducted to verify the superiority of this method under different working conditions. The research results indicate that this method significantly enhances spodumene recognition and sorting accuracy, substantially improving preselection grade and waste rejection rates. It provides technical support for optimizing intelligent ore sorting systems and promotes the development of mining production towards greater intelligence and sustainability

2. Related work

2.1. The principle and equipment for identifying spodumene using XRT imaging

X-ray transmission (XRT) preselection technology utilizes a radiation source to emit X-rays, which penetrate ores of suitable particle sizes to generate imaging. Artificial intelligence (AI) is then employed to analyze the images, identifying differences in mineral composition, after which a sorting system is used for separation. The fundamental principle behind this process is that X-rays attenuate as they pass through the ore. Different atoms within the ore absorb X-rays to varying degrees, allowing for the determination of its properties. Due to the strong penetrating power of X-rays, data is generated by detecting the energy signals produced as the X-rays pass through the material. As such, XRT preselection technology is capable of detecting variations in ore density, elemental composition, and other characteristics. Spodumene ores with different grades absorb the excited X-rays to varying degrees, which is then used for sorting. The principle of spodumene preselection using XRT technology is illustrated in Fig. 1. As shown in the Fig., X-rays emitted from the excitation source pass through the ore, attenuating to different extents. The receiver measures the degree of attenuation and forms corresponding images, which are then analyzed to determine the ore composition.

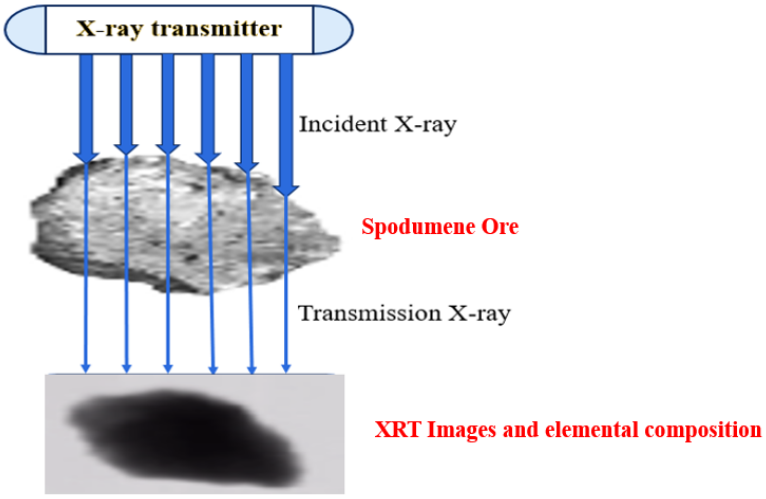

Fig. 1. The principle of using XRT technology for spodumene pre-selection

Based on the principle of spodumene sorting using XRT imaging, we have developed a deep learning-based intelligent X-ray preselection machine for spodumene. The intelligent X-ray preselection machine for spodumene consists mainly of a uniform material distribution and conveying system, an X-ray emission and detection system, an information processing and decision-making system, and a sorting and separation execution system.

The uniform material distribution and conveying system mainly consists of a vibrating feeder and a conveyor belt, which function to ensure that the material is spread in a single, uniform layer and passes smoothly and at a constant speed through the X-ray detection zone, before being thrown into the sorting and separation area at the same initial speed. The X-ray emission and detection system mainly includes the radiation source and the X-ray signal reception and detection unit, which are usually positioned above and below the conveyor belt to allow X-rays to penetrate the material layer, acquiring

fundamental classification and position information of material particles, providing a basis for subsequent scientific separation decisions. The information processing and decision-making system mainly consists of an industrial control computer and an operational display system. It serves as the core of the entire system, processing and analyzing the received X-ray signals, and is responsible for issuing intelligent decisions and separation commands. The sorting and separation execution system mainly includes an air compressor, high-pressure air jet nozzles, and their control valves. The airflow control valve is regulated by the information processing and decision-making system. Upon receiving an activation command, high-speed airflow is ejected from the nozzle to precisely blow the target particles that have left the belt and are undergoing projectile motion, altering their original trajectory. The non-target particles, unaffected and continuing on their original trajectory, land in separate collection bins. After completing one ejection cycle, the nozzle valve is immediately closed, cutting off the airflow, thereby completing a full operation cycle. By iterating this process, the system successfully completes the intelligent XRT preselection and waste removal of spodumene. The main components and working principle of the XRT intelligent ore sorting machine are shown in Fig. 2.

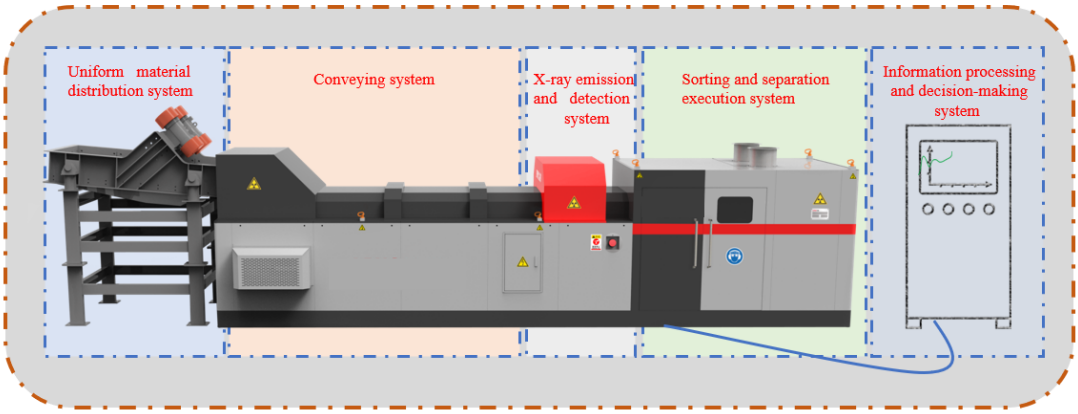

Fig. 2. The main components and working principle of the XRT intelligent ore sorting machine

It can be seen that the information processing and decision-making system is the core of the entire system. Its function is to process and analyze the received X-ray signals to generate the corresponding XRT image and perform sorting based on this image. This system mainly consists of three modules: image segmentation, image recognition, and image classification. An accurate image segmentation algorithm is the foundation for achieving precise sorting. However, in the XRT intelligent sorting process, the physical properties of the ore directly affect the sorting performance. Specifically, adhered ores and fine-grained ores may present challenges for X-ray imaging, image segmentation, and subsequent intelligent decision-making and sorting execution processes. Adhered ore refers to multiple ore particles that attach to each other or remain in close contact during transportation, making it difficult to accurately distinguish individual particle shapes and boundaries in the X-ray imaging process. When two or more particles are in close contact, XRT imaging may fail to clearly distinguish their boundaries, making it difficult for image segmentation algorithms to correctly delineate the ore contours, thereby affecting the accuracy of mineral classification. Additionally, due to the smaller size of fine-grained ore particles, their X-ray transmission signals experience weaker attenuation, resulting in increased noise and decreased sorting accuracy. Therefore, we need to use deep learning to build a segmentation model with strong generalization ability, high segmentation accuracy, and fast processing speed to eliminate the impact of adhered ore and fine-grained ore on recognition results. Based on this, feature extraction can be performed on each identified ore region to accurately recognize ore images in industrial settings.

2.2. Contributions of this paper

Different from previous literature, the main contributions of this paper are as follows. (1) Propose an improved Mineral-DeepLabV3+ model, using the lightweight MobileNetV4 as the backbone network to improve computational efficiency. By optimizing the ASPP module and introducing the CSWin attention mechanism, the model enhances detail feature extraction and improves the segmentation

accuracy of adhered ores and fine-grained spodumene. (2) Conduct comparative experiments and ablation studies to evaluate the model's performance against other algorithms and analyze the contribution of each improved module through ablation experiments, providing theoretical support for the optimization of intelligent ore sorting systems. (3) Deploy different algorithms on the XRT intelligent preselection machine for semi-industrial trials, evaluating sorting metrics under different ore particle sizes, feeding rates, and conveying speeds to verify the model's superiority in practical applications. This study offers novel perspectives for advancing intelligent ore sorting technology and establishes a robust technical foundation for the intelligent and sustainable development of mining operations.

3. Methods

3.1. Properties of spodumene raw ore

The samples in this study were collected from the Dahongliutan spodumene mine in Xinjiang, one of the major spodumene-producing areas in China. The ore is extracted from an open-pit mining operation in a uniform and random manner, making it highly representative. Representative samples were randomly selected and analyzed for chemical composition using inductively coupled plasma emission spectrometry (ICP), with the results of the raw ore's chemical composition shown in Table 1. According to Table 1, the primary recoverable component in the raw ore is Li_2O, with a content of 1.01%, while SiO_2, Al_2O_3, Na_2O, and K_2O exhibit relatively high concentrations of 7 .51%, 16.39%, 3.72%, and 2.58%, respectively. This suggests that the raw ore contains a substantial amount of quartz and feldspar minerals. This indicates that the content of valuable minerals is relatively low, and conventional flotation processes struggle to achieve effective recovery, requiring large amounts of flotation reagents and high temperatures, which result in environmental pollution. Therefore, intelligent optical preselection is essential to enhance the ore grade before beneficiation.

Table 1. Chemical composition of the raw ore

Chemical	Li_2O	SiO_2	CaO	MgO	Fe	Al_2O_3	K_2O	Na_2O
Quantity / %	1.01	73.51	0.44	0.12	0.55	16.39	2.58	3.92
Chemical	Rb_2O	TiO_2	MnO	BeO	Sn	Nb_2O_5	P_2O_5	
Quantity / %	0.12	0.02	0.08	0.05	0.07	0.013	0.38	

3.2. Dataset building and data enhancement

To enhance the spodumene recognition algorithm, we independently constructed a spodumene image dataset using X-rays. The XRT images of spodumene ore were collected using an XRT intelligent preselection machine, totaling 2,350 images, which include spodumene samples of different grades, particle sizes, and quantities. The ore regions in the images were manually labeled using Labeme software, forming the dataset required for model training. To adapt to the characteristics of ore image segmentation and align with real-world mining ore sorting scenarios, we proposed the RailAugment data augmentation method. We initially applied standard data augmentation operations and then selected appropriate ones based on the characteristics of the ore, removing operations such as inversion, equalization, and reversal, which could potentially degrade XRT image details and features. Subsequently, we divided data augmentation techniques into pixel-level and spatial-level operations, forming two augmentation spaces: the Pixel Augmentation Space (SP) and the Geometric Augmentation Space (SB). The SP space comprises {Gaussian blur, Gaussian noise, sharpness, color transformations, brightness, contrast, posterize}, whereas the SB space includes {translate x, translate y, shear x, shear y, scale, rotate, horizontal flip, vertical flip}.

A random sampling of M augmentation operations from each space was conducted, with M set to {2,3} in this study. We generated four combinations sampled from SP and SB: 1+2, 0+3, 1+1, and 0+2. The M-Augment data augmentation method simulates real-world scenarios such as large-grained ores, fine-grained ores, high-quantity ores, and low-quantity ores, as shown in Fig. 3. After manual selection, a dataset of 6,940 images was obtained. The dataset was randomly split in a 7:2:1 ratio, resulting in 4,858

training images, 1,388 validation images, and 694 test images, forming the spodumene XRT segmentation dataset (XSD) for experimentation.

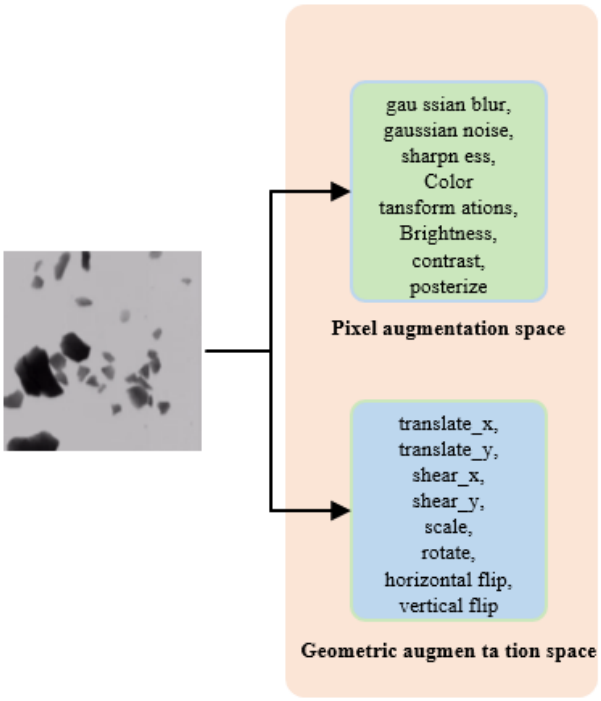

Fig.3. Effect diagram of spodumene XRT segmentation dataset augmentation method

3.3. Mineral-DeepLabV3+ algorithm construction and improvement

DeepLabV3+ primarily follows an Encoder-Decoder structure, consisting of an encoder module that progressively reduces feature maps to capture higher-level semantic information and a decoder module that gradually restores spatial details. The encoder comprises two key components: the backbone (DCNN) and the ASPP module. The first module in the encoder is the DCNN, which serves as the backbone network for extracting image features. Following the DCNN is the ASPP network, which processes the backbone output using a 1×1 convolution, three 3×3 dilated convolutions, and global pooling. The results are then concatenated and processed with a 1×1 convolution to reduce the number of channels.

The decoder receives low-level feature maps from the backbone's intermediate layers and the output from the ASPP module as input. The ASPP module is used to fuse deep and shallow features. Dilated convolutions with different dilation rates capture more contextual information, allowing each convolutional output to encompass a broader range of features, thereby improving feature extraction capability. However, dilated convolutions may also lead to the loss of fine local details and insufficient feature representation, particularly when dealing with blurred ore boundaries in XRT images. In ore preselection applications, ore boundary details are often very fine or share similar colors and textures with adjacent ores, making dilated convolutions prone to ignoring critical local details and failing to preserve boundary information effectively. Additionally, DeepLabV3+ incorporates a complex multi-scale information processing approach that theoretically improves performance, but in real-world applications, the increased number of parameters and computations may result in longer training and inference times. In mining environments with limited computational resources, real-time detection is essential, and prolonged computations may hinder system efficiency and response time

3.3.1. Mineral-DeepLabV3+ algorithm

Due to the complexity of real-world conditions in ore XRT preselection, where fine-grained ores and dust are prevalent and prone to noise interference, the standard DeepLabV3+ model suffers from local detail loss, insufficient fine-grained feature representation, and long computation times. To better

address the requirements of practical scenarios, this study proposes the Mineral-DeepLabV3+ model, specifically designed for ore segmentation, with its network architecture shown in Fig. 4. The Mineral-DeepLabV3+ model builds on DeepLabV3+ by integrating the lightweight MobileNetV4 as its backbone network for feature extraction, reducing parameter count and enhancing segmentation speed. Furthermore, the ASPP module is optimized by introducing a controllable function during pyramid pooling, followed by multi-scale feature fusion. This modification allows the network to perform denser pixel sampling, strengthening its ability to extract fine details and improving feature reuse and representation. Moreover, by incorporating the CSWin attention mechanism after each convolutional layer and shallow feature in the improved ASPP module, the model can compute horizontal and vertical attention in parallel, allowing it to better focus on ore-related information and improve segmentation accuracy. In convolutional structures, the 1×1 convolution is not merely a simple scaling operation of the input; instead, it plays an important role in semantic segmentation tasks. First, the 1×1 convolution enables information fusion across channels without altering the spatial resolution of the feature map, thereby recombining semantic features from different channels. Second, it is often used for dimensionality reduction, effectively decreasing the number of feature channels while preserving the most relevant information, which reduces both parameters and computational cost. Finally, by combining with non-linear activation functions, the 1×1 convolution enhances the non-linear representation capability of the network, allowing it to better capture complex ore boundary information. Therefore, in the proposed Mineral-DeepLabV3+ model, the 1×1 convolution plays a crucial role in feature compression and cross-channel feature recombination, contributing to improved segmentation accuracy while maintaining computational efficiency.

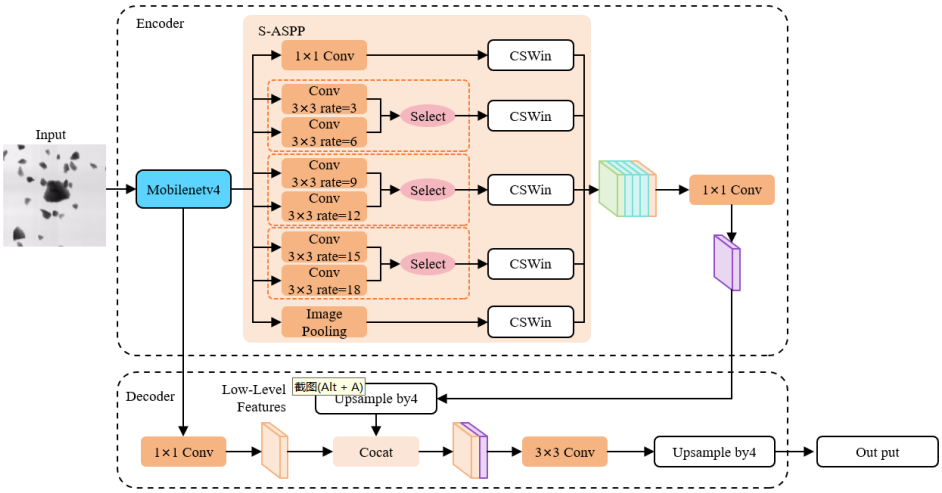

Fig. 4. Mineral-DeepLabV3+ model structure

3.3.2. Lightweight backbone feature network

The images in the ore preselection process are complex, and the DCNN network in the original model is overly cumbersome for ore segmentation tasks, leading to local information loss, long computation times, and insufficient fine-grained feature representation. This study adopts the lightweight MobileNetV4 architecture as the backbone for feature extraction. The core components of MobileNetV4 (MNV4) include the Universal Inverted Bottleneck (UIB) and the Mobile MOA module, combined with an optimized Neural Architecture Search (NAS) strategy. The inverted bottleneck block includes two optional depthwise convolutions (DW), one preceding the expansion layer and the other situated between the expansion and projection layers. These adjustments enable the model to achieve stronger feature representation while maintaining low computational cost.

3.3.3. CSWin attention module

Spodumene XRT images are influenced by various factors such as variations in X-ray excitation, different ore orientations, particle sizes, and dust coverage, which pose significant challenges to

segmentation tasks. Cross-shaped Window Self-Attention (CSWin) is a cross-shaped attention mechanism module that computes horizontal and vertical attention in parallel, as shown in Fig. 5. By integrating the CSWin module after each convolution operation, the model can effectively capture both local and global semantic information across different scales, thereby improving segmentation performance in complex scenarios.

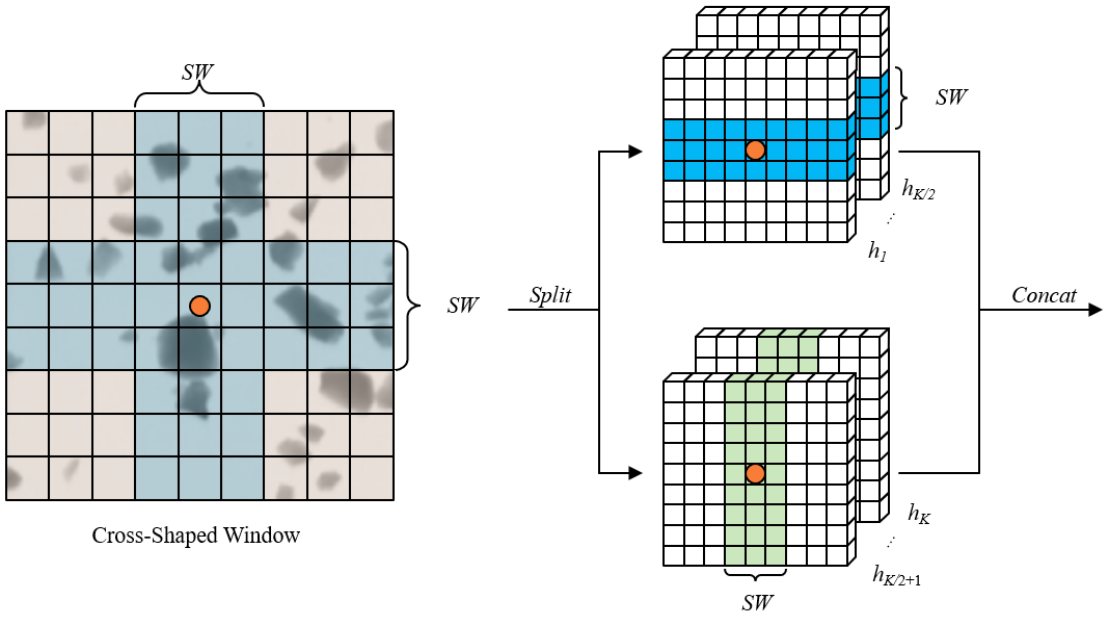

Fig.5. CSwin self-attention module structure

3.4. Training platform and embedded platform

The training model platform hardware configuration includes an I9 9900K CPU, Nvidia RTX4090 GPU with 8GB memory, and a software environment of 64-bit Windows 10, Python 3.7, Pytorch 2.0.0 framework, and Cuda 11.0. The training parameters are shown in Table 2.

Table 2. Training parameters

Training paramers	Values
Initial learning rate	0.01
Optimizer	SGD
Optimizer momentum	0.937
Optimizer weight decay rate	0.0005
Number of images per batch	16
Number of epochs	100

In industrial experiments, the model needs to be deployed on the XRT intelligent pre-selection equipment for spodumene, using Nvidia Jetson Nano with an A57 quad-core ARM CPU, 128-core Maxwell GPU, 4GB of RAM, and 64-bit LPDDR4. The software environment is Ubuntu 18.04, with runtime configurations including Jetpack 4.5, Python 3.6, Pytorch 1.8, and TensorRT 8.0.1.6

3.5. Practical experiments

To assess the algorithm's performance, practical experiments were carried out on spodumene ore sourced from the Dahongliutan mine in Xinjiang. The Mineral-DeepLabV3+ algorithm, alongside other benchmark algorithms, was applied to intelligent preselection equipment under varying operational conditions. The preselection results were analyzed across different ore particle sizes, processing volumes, and conveyor speeds. These outcomes were then compared against those obtained using traditional X-ray preselection methods. The industrial experiment conditions are summarized in Table 3.

Table 3. industrial experiment conditions

Experiments	particle size, (mm)	processing volumes, (kg/h)	conveyor speeds, (m/s)
1	0-5; 5-10; 10-15; 15-20; 20-25	50	4
2	10-15	10; 20; 30; 40; 50; 60; 70	4
3	10-15	50	1; 2; 3; 4; 5; 6; 7

3.6. Evaluation metrics

The evaluation metrics are mainly composed of two parts: model evaluation metrics and industrial experiment metrics. The main model evaluation metrics include Accuracy, Mean Pixel Accuracy (mPA), and Mean Intersection over Union (mIoU).

Accuracy refers to the proportion of correctly predicted samples to the total number of samples, making it one of the key factors for evaluating a model. The calculation formula is shown in Formula (1). Where TP represents the number of pixels correctly identified as fasteners, TN represents the number of pixels correctly identified as background, FP represents the number of background pixels incorrectly identified as fasteners, and FN represents the number of fastener pixels incorrectly identified as background.

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+FP+TN+FN}} \tag{1}$$

Meanwhile, mPA is also an important metric. It is calculated by averaging the pixel accuracy of each category, reflecting the classification ability of different categories at the pixel level. The calculation formula is given in Formula (2). Where k+1 represents the number of categories, piidenotes the number of correctly classified pixels, pij represents the number of pixels of class iii that are predicted as class j, and pji represents the number of pixels of class j that are predicted as class iPA (Pixel Accuracy) represents the ratio of correctly classified pixels to the total number of pixels in each class, while mPA (Mean Pixel Accuracy) is obtained by computing PA for each class and then averaging the results.

$$\text{mPA} = \frac{1}{\text{k+1}} \sum_{\text{i=0}}^{\text{k}} \frac{\text{p}_{\text{ii}}}{\sum_{\text{j=0}}^{\text{k}} \text{p}_{\text{ij}}} \tag{2}$$

MIoU is another crucial metric for model evaluation. It is typically used to measure the overlap between predicted results and ground truth annotations. IoU is calculated as the ratio of the intersection to the union of the predicted and actual regions, while mIoU is the average IoU across all categories. The calculation formula is shown in Formula (3).

$$\text{mIoU} = \frac{1}{\text{k+1}} \sum_{\text{i=0}}^{\text{k}} \frac{\text{p}_{\text{ii}}}{\sum_{\text{j=0}}^{\text{k}} \text{p}_{\text{ij}} + \sum_{\text{j=0}}^{\text{k}} \text{p}_{\text{ji}} - \text{p}_{\text{ii}}} \tag{3}$$

The assessment criteria for industrial trials include the grade of the spodumene pre-concentrate and the waste rejection rate of the raw ore. The waste rejection rate quantifies the proportion of waste material removed during the sorting process, providing a reliable indicator of sorting efficiency. It is defined as the %age of waste rock or tailings discarded during beneficiation in relation to the total raw ore processed. A higher waste rejection rate signifies a more efficient elimination of undesired material, thereby improving the grade and overall quality of the ore. This enhancement further contributes to reducing carbon emissions and minimizing environmental pollution associated with beneficiation. The waste rejection rate is generally calculated using the formula presented in Formula 4, where mc represents the mass of waste rock removed, and mw denotes the total mass of raw ore. These metrics play a crucial role in evaluating both the effectiveness and economic viability of the beneficiation process.

$$\gamma = \frac{\text{m}_\text{c}}{\text{m}_\text{w}} \tag{4}$$

4. Results and discussion

4.1. Ablation studies and comparative experiments

4.1.1. Ablation study on module optimization

To further validate the segmentation performance of the proposed algorithm in the pre-selection of spodumene ores and investigate the effectiveness of each improvement, an ablation study was

conducted. Using the XSD dataset for training and MobileNetV4 as the backbone network, a comparative analysis was performed on the S-ASPP module and the introduction of the CSWin attention mechanism, as shown in Table 4. According to the results in Table 4, the introduction of the S-ASPP module enhanced the capability to handle small objects and fine details in segmentation tasks, particularly under complex backgrounds, and its multi-scale feature extraction improved segmentation accuracy. The CSWin attention mechanism strengthened the model's ability to focus on target objects and suppressed background noise; although its individual effect was not as strong as S-ASPP, it positively contributed to the model's overall background suppression and target discrimination capabilities. Combining S-ASPP with CSWin is crucial to obtaining the best segmentation performance. In the segmentation task of spodumene ores under fine-grained and high-speed recognition conditions, the complementary nature of these two modules played an important role in enhancing overall model performance, achieving an mPA of 97.24%, Accuracy of 97.62%, and mIoU of 94.77%, thus demonstrating optimal results.

Table 4. Performance comparison results of improved modules

S-ASPP	CSWin	mPA/%	Accuracy/%	mIoU/%
×	×	93.12	92.11	90.67
√	×	96.11	96.62	92.22
×	√	96.22	96.88	93.21
√	√	97.24	97.62	94.77

4.1.2. Comparative experiments on backbone network replacement and controllable atrous spatial pyramid pooling

To verify the effectiveness of MobileNetV4, comparative experiments were conducted on the XSD dataset using various networks as the backbone feature extraction networks for the DeepLabV3+ model. As shown in Table 5, MobileNetV4 achieved an mPA of 96.35%, which was only 1.32% lower than Xception, yet exceeded ResNet101 and VGG16 by 10.24% and 14.34%, respectively. In terms of Accuracy and mIoU, MobileNetV4 reached 95.88% and 91.64%, slightly lower than Xception's 95.97% and 91.95%, but still substantially higher than ResNet101 and VGG16. With only 13.92M parameters, MobileNetV4 was significantly smaller compared to other networks, having 19.52M fewer parameters than VGG16, and markedly fewer parameters than ResNet101 and Xception. Regarding inference time, MobileNetV4 was faster than ResNet101, Xception, and MobileNetV2, particularly being 24.1ms faster than Xception. MobileNetV4 achieved an excellent balance between accuracy, parameter count, and inference time. Its accuracy was nearly comparable to Xception, yet with significantly lower parameters and inference time, making it ideal for resource-constrained environments, particularly in mining scenarios where real-time accurate judgment of a large amount of ore is required with stringent demands on response speed and computational resources.

Table 5. Comparison results of different backbone networks

Backbone network	mPA /%	Accuracy /%	mIoU /%	Parameters /M	Inference time /ms
VGG16	83.15	86.26	83.21	31.96	18.92
ResNet101	87.25	89.63	87.59	28.71	22.31
Xception	98.81	95.97	91.95	26.44	40.42
MobileNetv4	97.49	95.88	91.64	12.84	16.32

Controllable Atrous Spatial Pyramid Pooling (S-ASPP) applies atrous convolutions with different dilation rates to extracted high-level features, then utilizes a control function to select the optimal result. For ore XRT image segmentation tasks, the appropriate dilation rate needs to be selected according to specific circumstances to achieve the best segmentation outcome. Taking the original ASPP dilation rates (6, 12, 18) as a reference set, the following dilation rate combinations were set up for comparative experiments to determine alternative dilation rates for atrous convolutions. The experiments utilized

the XSD dataset as the training dataset and MobileNetV4 as the backbone network. As indicated in Table 6, the dilation rate combination of (3, 9, 15) exhibited the best performance, achieving an mPA of 96.71%, an Accuracy of 97.07%, and an mIoU of 93.74%. Thus, it was selected as the candidate dilation rate for S-ASPP.

Table 6. Performance comparison experiment of S-ASPP module

Serial Number	Dilation rate combination	mPA/%	Accuracy/%	mIoU/%
1	(2,5,7)	95.53	96.49	92.62
2	(3,9,15)	96.71	97.07	93.74
3	(4,16,24)	95.93	96.71	92.43
4	(5,15,25)	95.64	96.1	92.13

4.1.3. Comparative experiment of semantic segmentation models

Under the same environment, five classical semantic segmentation models (U-Net, SegNet, DenseNet, PSPNet, DeepLabV3+) as well as the recently proposed Swin-UNet and Segmenter models were trained on the RFS dataset alongside the proposed Sc-DeepLabV3+ model, with all experiments conducted under identical experimental settings and dataset. Table 5 Comparison results of different backbone networks. The segmentation accuracy metrics and parameter counts of the different models were obtained, as shown in Table 7.

As shown in Table 7, the mIoU of Mineral-DeepLabV3+ reached 95.19%, and the mPA was 97.16%, representing an improvement of approximately 3.28% and 0.41% over the baseline DeepLabV3+. This indicates that the innovative introduction of S-ASPP and CSWin based on DeepLabV3+ has had a significant effect, enhancing the precise segmentation of fastener details and edges. Moreover, compared to other models, Minerals-DeepLabV3+ also shows outstanding accuracy, with its mPA, Accuracy, and mIoU metrics being close to those of Swin-UNet and Segmenter, indicating that this model is highly competitive at this stage. Sc-DeepLabV3+ has a parameter count of 30.5M, which is significantly lower than that of the baseline DeepLabV3+, Swin-UNet, and Segmenter. This reduction is due to Sc-DeepLabV3+ employing the lightweight MobileNetV4 backbone, which allows the model to execute tasks efficiently under limited computational resources, thereby lowering labor and compute costs. Additionally, it further boosts computational speed and reduces memory usage, making deployment on mobile devices in mining and other resource-constrained settings possible.

The complete visualization results of the segmentation process, including detected edges and object contours for different methods, have been provided in a publicly accessible GitHub repository: https://github.com/98123111/DeepLabV3-.

Table 7. Comparison of the performance of different models on the RFS dataset

Algorithm	mPA/%	Accuracy/%	mIoU/%	Number of parameters/M
U-Net	85.29	94.89	81.71	7.52
SegNet	86.87	96.09	83.64	27.61
DenseNet	91.88	96.49	89.04	28.7
PSPNet	96.18	94.1	91.11	46.85
DeepLabV3+	96.75	95.59	91.91	54.26
Swin-UNet	97.01	97.26	94.07	59.64
Segmenter	97.12	97.51	93.81	53.51
Mineral-DeepLabV3+	97.16	97.88	95.19	30.5

4.2. Semi-industrial experiments under algorithm deployment

To rigorously validate the practical effectiveness, robustness, and industrial applicability of the proposed Mineral-DeepLabV3+ algorithm, a series of carefully designed semi-industrial experiments were carried out. In this study, the Mineral-DeepLabV3+ model was deployed onto a semi-industrial

intelligent optical-electronic preselection device specifically configured for spodumene sorting tasks. The experimental design comprehensively evaluated the sorting performance under a range of operational conditions, including variations in ore particle size, processing volumes, and conveyor speed. These factors were systematically varied to assess not only the algorithm's adaptability to different material characteristics but also its resilience under the dynamic throughput conditions typically encountered in real-world mineral processing operations. In addition, a conventional X-ray sorting system, widely regarded as a standard in current industrial practice, was employed as a control benchmark. By comparing the performance of the deep learning-based system against this traditional technology, the study aimed to quantify the relative improvements in sorting precision, efficiency, and overall processing capability. The evaluation focused particularly on the ability of the proposed algorithm to enhance the detection and separation of fine-grained spodumene particles, improve the rejection of gangue material, and maintain stable performance even at high processing speeds and large throughput volumes. The comprehensive experimental results thus provide strong empirical evidence supporting the superior competitiveness of the Mineral-DeepLabV3+ framework, highlighting its potential for integration into future intelligent mineral processing systems operating under resource-constrained or mobile deployment conditions.

4.2.1. Preselection experiments of spodumene under different particle sizes

Initially, the preselection experiments were carried out under varying ore particle size conditions. Particle size is one of the key factors affecting preselection outcomes, mainly because spodumene minerals are finely distributed in the ore and closely associated with gangue minerals, often forming complex intergrowths with quartz, feldspar, and others. Larger particle sizes result in incomplete spodumene liberation, leading to gangue contamination in the concentrate and lowering sorting performance; conversely, overly fine particles pose challenges for sensor detection. Under a processing capacity of 30 kg/h and conveyor speed of 4 m/s, Fig. 6 presents the X-ray preselection results of spodumene ores with different particle sizes using the Mineral-DeepLabV3+ algorithm.

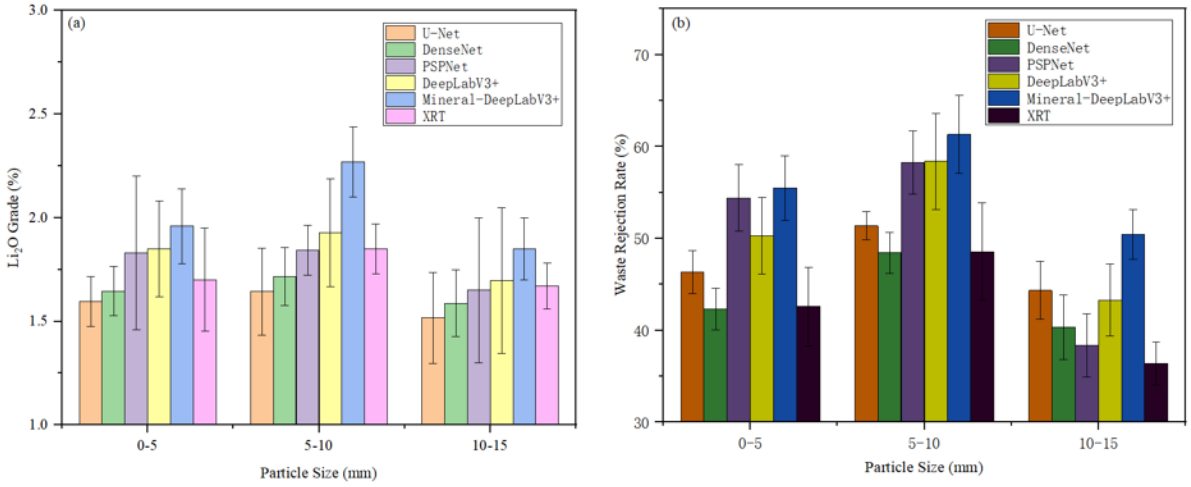

Fig. 6. Pre-selection results under different particle sizes

Fig. 6(a) shows that among various algorithms, as the particle size increases, the concentrate grade first rises and then falls, indicating that excessively large particles perform poorly, and excessively small particles also make it difficult to achieve ideal grades. However, the Mineral-DeepLabV3+ algorithm proposed in this study still shows significant advantages, as it employs the cross-shaped CSWin self-attention mechanism to fully preserve fine-grained feature map information, enhancing model detection performance and accurately identifying fine-grained ores. Under smaller particle sizes, the concentrate Li_2O grade can reach 1.95%, with the best preselection result reaching 2.27% in the 5–10 mm size range. Overall, the algorithm can effectively detect and separate spodumene across all particle size ranges. As seen in the Fig. 7 (a), compared to traditional XRT preselection methods, the deep learning-

based X-ray preselection significantly improves sorting performance, directly demonstrating its effectiveness.

Fig. 6(b) shows that with increasing particle size, the waste rejection rate first rises and then falls. This phenomenon mainly arises because although fine particles are well liberated, they are difficult for sensors to fully recognize, and the air-sorting system may mistakenly blow some light mineral particles into the concentrate, reducing the rejection rate. When the particle size is large, although recognition improves, the poor degree of ore liberation leads to some valuable components being discarded with the tailings, causing resource waste. Under the Mineral-DeepLabV3+ algorithm in this study, the rejection rate in the 5–10 mm size range reached as high as 61%, further validating the effectiveness of the deep learning algorithm.

4.2.2. Preselection experiments of spodumene under different feeding speeds

Additionally, the study carried out preselection experiments under varying processing volumes. Variations in processing volumes have a direct impact on the operational capacity of the sorting equipment. An increase in processing volume leads to a higher number of ore blocks on the conveyor per unit time, but the number of blocks the sensor can process per unit time is limited, thereby affecting preselection efficiency. Therefore, under conditions of 10–15 mm particle size and a conveyor speed of 4 m/s, this study investigated the effects of processing volume variations on spodumene preselection results, as shown in Fig. 7.

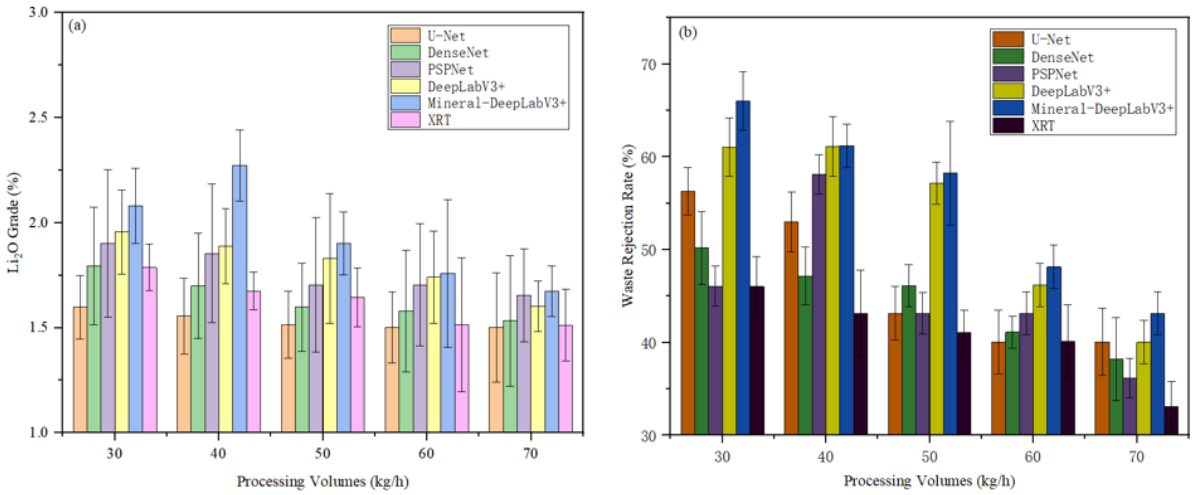

Fig. 7. Pre-selection results under different processing volumes

Fig. 7 indicates that under low processing volumes, performance metrics are relatively stable, but when the volume increases, sorting performance drops sharply, demonstrating that the equipment's processing capacity is inherently limited by sorting effectiveness. Fig. 7(a) shows that increasing processing volume leads to a continuous decline in concentrate grade, reflecting the device's limited processing capacity per unit time. However, the Mineral-DeepLabV3+ algorithm proposed in this study still performs excellently, being able to detect a large number of spodumene blocks per unit time better than other algorithms, significantly improving production efficiency, mainly due to enhanced feature fusion capabilities after algorithm improvements. Under large sample conditions, its detection efficiency surpasses that of traditional X-ray preselection methods, maintaining a high concentrate grade even under high processing volumes.

Fig. 7(b) shows that the rejection rate decreases as processing volume increases, indicating that the device's sorting capability declines under high processing volumes. This is mainly because high processing volumes cause ore blocks to obscure each other, affecting X-ray imaging quality. However, the Mineral-DeepLabV3+ algorithm enhances performance by introducing an ASPP module, adding controllable functions during pyramid pooling, and applying multi-scale feature fusion, which provides denser pixel sampling, better detail extraction, and improved feature reuse and representation—thus

improving detection accuracy, enhancing concentrate-waste separation efficiency, and raising overall equipment throughput.

4.2.3. Preselection experiments of spodumene under different conveyor speeds

Lastly, preselection experiments were performed at various conveyor speeds. As the conveyor speed increases, the movement speed of ore blocks on the belt rises, increasing the number of images captured by the camera per unit time, thereby placing greater demands on the algorithm's processing capacity. Therefore, changes in conveyor speed also influence the efficiency of the preselection process. Under conditions of 10–15 mm particle size and 30 kg/h processing volume, this study investigated the effects of conveyor speed changes on the UV fluorescence preselection results of spodumene, as shown in Fig. 8.

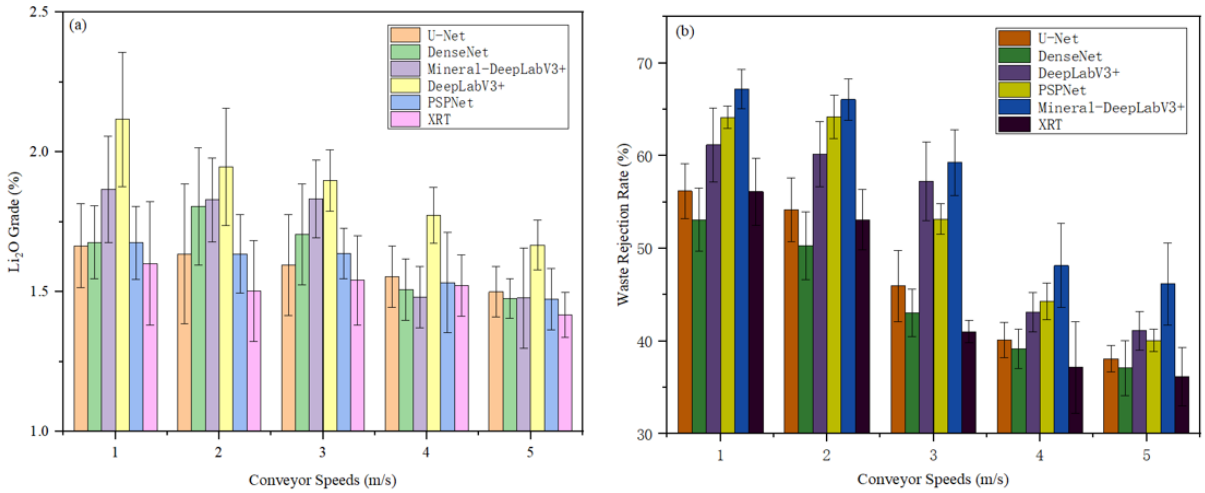

Fig. 8. Pre-selection results under different conveyor speeds

Fig. 8 indicates that with increasing conveyor speed, both the concentrate grade and rejection rate decrease, demonstrating that despite enhanced equipment capacity, sorting performance diminishes accordingly. However, the preselection device using the Mineral-DeepLabV3+ algorithm can still maintain relatively stable performance metrics under high-speed conveyor conditions. Fig. 8(a) shows that the concentrate grade decreases as conveyor speed increases, indicating that per-unit-time processing capacity has a significant impact on sorting performance. Nevertheless, the Mineral-DeepLabV3+ algorithm proposed in this study still shows excellent performance. Compared to other algorithms, this method offers a higher FPS, enabling the analysis of more images per unit time and enhancing production efficiency. Compared to traditional X-ray sorting technologies, the deep learning approach can still achieve higher concentrate grades at high conveyor speeds, mainly because conventional detection equipment has limited processing capacity and is better suited for identifying individual ore blocks.

Fig. 8(b) indicates that as conveyor speed increases, the rejection rate decreases, showing that preselection performance declines at high speeds due to limited model recognition, which prevents timely commands to the mechanical sorting device, resulting in waste rock mixing into the concentrate. Furthermore, the increased throughput per unit time further constrains the sorting system's operational efficiency. However, the Mineral-DeepLabV3+ algorithm in this study integrates a lightweight MobileNetV4 backbone into the DeepLabV3+ framework, significantly reducing parameter count and increasing segmentation speed, thus improving computational efficiency, shortening runtime, and enhancing sorting performance while maintaining high precision.

5. Conclusion

This paper proposes a novel ore pre-selection method that combines deep learning and X-ray transmission technologies, aiming to enhance the accuracy of spodumene ore image segmentation,

improve ore grades entering the beneficiation process, reduce energy consumption, and achieve low-carbon, sustainable mining operations.

(1) This study developed an intelligent segmentation and sorting system for spodumene X-ray images based on the Mineral-DeepLabV3+ deep learning model. The system integrates a conveyor and feeder unit, X-ray excitation and detection components, and an air-jet separation device. A proprietary dataset of 6,900 X-ray transmission images covering spodumene ores of varying particle sizes and Li_2O grades was constructed. To address limited image data, pixel-level and spatial-level augmentation techniques were applied, laying a foundation for model optimization and future segmentation algorithm development.

(2) The paper presents the Mineral-DeepLabV3+ model, specifically designed for ore image segmentation tasks. It integrates the lightweight MobileNetV4 backbone, an optimized Atrous Spatial Pyramid Pooling module, and the CSWin attention mechanism, significantly improving segmentation precision while reducing computational load. Compared to baseline models, Mineral-DeepLabV3+ achieves a 3.28% increase in segmentation accuracy, reduces parameter count by over 44%, and maintains fast inference speed, enabling efficient, real-time segmentation and detection of fine-grained and adhered spodumene ores under industrial conditions.

(3) Comparative experiments against leading semantic segmentation models, including U-Net, SegNet, PSPNet, DeepLabV3+, Swin-UNet, and Segmenter, demonstrated that Mineral-DeepLabV3+ consistently outperforms alternatives, achieving a mean intersection over union of 95.19% and excelling in ore boundary identification, particularly under high-speed, large-throughput conditions.

(4) Semi-industrial trials further validated the real-world effectiveness of Mineral-DeepLabV3+ in spodumene ore pre-selection based on image segmentation. Results showed that the system achieves concentrate grades up to 2.27% Li_2O and waste rejection rates up to 66% across varying particle sizes, feeding rates, and conveyor speeds, demonstrating superior industrial applicability and real-time segmentation capability.

In summary, this paper presents an innovative ore image segmentation and pre-selection framework that combines deep learning and X-ray transmission, significantly enhancing spodumene sorting performance. It provides critical technical support for intelligent mineral processing systems and low-carbon, green mining, while paving the way for future advancements in sustainable mining technologies.

Acknowledgments

This work was supported by the Study on multi-field pre-enrichment of spodumene and synchronous enrichment of low-grade beryllium tantalum niobium [grant numbers 2021YFC2903201]. The funding agencies had no role in study design, data collection, and analysis, decision to publish, or preparation of the manuscript.

References

CHEN, L.-C., ZHU, Y., PAPANDREOU, G., SCHROFF, F., ADAM, H. *Encoder-decoder with atrous separable convolution for semantic image segmentation*. 2018. 801-818.

DING, T., XU, H., WANG, J., LIU, D., WANG, Z. 2014. R*esearch of Online Ore Sorter based on Vision Recognition Technology.* Equipment Manufacturing Technology, 7, 106-108.

DONG, X., BAO, J., CHEN, D., ZHANG, W., YU, N., YUAN, L., CHEN, D., GUO, B. *Cswin transformer: A general vision transformer backbone with cross-shaped windows.* 2022. 12124-12134.

GAO, H., DENG, N., WANG, G., WANG, X., LIU, Y., ZHANG, L., LIANG, Y., YAN, J., CHENG, B., KANG, W. 2022. *Overview of the Latest Developments and Perspectives about Noncarbon Sulfur Host Materials for High Performance Lithium–Sulfur Batteries.* Energy & Fuels, 36, 7284-7320.

GREIM, P., SOLOMON, A. A., BREYER, C. 2020. *Assessment of lithium criticality in the global energy transition and addressing policy gaps in transportation.* Nat Commun, 11, 4570.

HAN, S., SAGZHANOV, D., PAN, J., VAZIRI HASSAS, B., REZAEE, M., AKBARI, H., MENSAH-BINEY, R. 2022. *Direct extraction of lithium from a-spodumene by salt roasting–leaching process.* ACS Sustainable Chemistry & Engineering, 10, 13495-13504.

JAMESH, M.-I. 2019. *Recent advances on flexible electrodes for Na-ion batteries and Li–S batteries.* Journal of Energy Chemistry, 32, 15-44.

JUNG, W., JEONG, J., KIM, J., CHANG, D. 2020. *Optimization of hybrid off-grid system consisting of renewables and Li-ion batteries*. Journal of Power Sources, 451, 227754.

KOH, E. J. Y., AMINI, E., MCLACHLAN, G. J., BEATON, N. 2021. *Utilising convolutional neural networks to perform fast automated modal mineralogy analysis for thin-section optical microscopy*. Minerals Engineering, 173, 107230.

LI, G., KLEIN, B., SUN, C., KOU, J. 2021. *Lab-scale error analysis on X-ray fluorecence sensing for bulk ore sorting*. Minerals Engineering, 164, 106812.

LI, N., ZHANG, S., PENG, H., FU, L., LU, Y. 2018. A*pplication Practice and Evaluation of Photoelectric Separation in a Phosphate Mining Industry.* Non-Metallic Mines, 41, 73-75.

LI, Y., YU, Y., TONG, X., LIU, H., WEI, X. 2023. *Study on the Application of X-ray Separation Technology in Phosphate Mine Dressing*. Multipurpose Utilization of Mineral Resources, 100-105.

LIU, Y., MA, B., Lü, Y., WANG, C., CHEN, Y. 2023. *A review of lithium extraction from natural resources.* International Journal of Minerals, Metallurgy and Materials, 30, 209-224.

LIU, Y., MA, B., LV, Y., WANG, C., CHEN, Y. 2022. *Selective recovery and efficient separation of lithium, rubidium, and cesium from lepidolite ores.* Separation and Purification Technology, 288, 120667.

LIU, Y., ZHANG, Z., LIU, X., WANG, L., XIA, X. 2021. *Efficient image segmentation based on deep learning for mineral image classification.* Advanced Powder Technology, 32, 3885-3903.

QIN, D., LEICHNER, C., DELAKIS, M., FORNONI, M., LUO, S., YANG, F., WANG, W., BANBURY, C., YE, C. , AKIN, B. *MobileNetV4: universal models for the mobile ecosystem.* 2024. Springer, 78-96.

QIU, S., ZHU, Y., JIANG, Y., LIU, C., YU, J. 2022. K*inetics and mechanism of lithium extraction from a-spodumene in potassium hydroxide solution.* Industrial & Engineering Chemistry Research, 61, 15103-15113.

ROBBEN, C., CONDORI, P., PINTO, A., MACHACA, R., TAKALA, A. 2020. *X-ray-transmission based ore sorting at the San Rafael tin mine.* Minerals Engineering, 145, 105870.

SANDLER, M., HOWARD, A., ZHU, M., ZHMOGINOV, A., CHEN, L.-C. *Mobilenetv2: Inverted residuals and linear bottlenecks.* 2018. 4510-4520.

SARKER, S. K., HAQUE, N., BHUIYAN, M., BRUCKARD, W., PRAMANIK, B. K. 2022. *Recovery of strategically important critical minerals from mine tailings.* Journal of Environmental Chemical Engineering, 10, 107622.

SHI, X., XU, L., WANG, D., MA, Z., XUE, K., MENG, J. 2023. *The selective adsorption and flotation separation of spodumene from feldspar using a novel iron-group multiple ligand collector.* Colloids and Surfaces A: Physicochemical and Engineering Aspects, 677, 132363.

SHU, K., XU, L., WU, H., PENG, L., XU, Y., LUO, L., YANG, J., TANG, Z. 2020. *In situ adsorption of mixed collectors BHA/DDA in spodumene-feldspar flotation system.* Separation and Purification Technology, 251, 117325.

SONG, Y., ZHAO, T., HE, L., ZHAO, Z., LIU, X. 2019. *A promising approach for directly extracting lithium from a-spodumene by alkaline digestion and precipitation as phosphate*. Hydrometallurgy, 189, 105141.

TABELIN, C. B., DALLAS, J., CASANOVA, S., PELECH, T., BOURNIVAL, G., SAYDAM, S., CANBULAT, I. 2021. *Towards a low-carbon society: A review of lithium resource availability, challenges and innovations in mining, extraction and recycling, and future perspectives.* Minerals Engineering, 163, 106743.

TANG, X., JIA, Q., YANG, L., BAI, M., WU, W., WANG, Z., GONG, M., SA, S., TAO, S., SUN, M. 2020. *Towards the high-energy-density battery with broader temperature adaptability: Self-discharge mitigation of quaternary nickel-rich cathode.* Energy Storage Materials, 33, 239-249.

WU, Z. 2020. *Ore pre concentration and discarding waste technology and selection of intelligent photoelectric beneficiation equipment.* World Nonferrous Metals, 202-205.

XIANG, H., WANG, Y., ZHANG, X., CHEN, Z. 2023. *Two-Level Battery Health Diagnosis Using Encoder–Decoder Framework and Gaussian Mixture Ensemble Learning Based on Relaxation Voltage.* IEEE Transactions on Transportation Electrification, 10, 3966-3975.

XIE, R., ZHU, Y., LIU, J., LI, Y. 2021a. *Effects of metal ions on the flotation separation of spodumene from feldspar and quartz.* Minerals Engineering, 168, 106931.

XIE, R., ZHU, Y., LIU, J., LI, Y. 2021b. *A self-assembly mixed collector system and the mechanism for the flotation separation of spodumene from feldspar and quartz*. Minerals Engineering, 171, 107082.

XIE, R., ZHU, Y., LIU, J., WANG, X., LI, Y. 2020. *Differential collecting performance of a new complex of decyloxy-propyl-amine and a-bromododecanoic acid on flotation of spodumene and feldspar*. Minerals Engineering, 153, 106377.

YANG, H., HUANG, C., WANG, L., LUO, X. 2020. *An improved encoder–decoder network for ore image segmentation*. IEEE Sensors Journal, 21, 11469-11475.

YELATONTSEV, D., MUKHACHEV, A. 2021. Processing of lithium ores: Industrial technologies and case studies–A review. Hydrometallurgy, 201, 105578.

ZHAI, Q., SONG, Y., LIU, R., WANG, Y., ZHU, X., XIE, Z., MAO, Z., SUN, W. 2023. *An innovative scheme for the separation of potassium from lithium-bearing solution: a jarosite precipitation method*. Separation and Purification Technology, 322, 124283.

ZHANG, G. Y., LIU, G. Z., ZHU, H., QIU, B. 2010. *Ore image thresholding using bi-neighbourhood Otsu's approach*. Electronics letters, 46, 1666-1668.

ZHENG, F. 1980. *Development of photoelectric sorter abroad*. Nonfer rous Met. Mineral Process, 40-45.

2020a. *Rafaella considers processing options after positive TOMRA XRT ore sorting tests*. International Mining.

2020b. *TOMRA's XRT sensor-based ore sorting technology significantly improves productivity and extends life of San Rafael tin mine*. International Mining, 184-185.

2020c. *Vimy senses Angulari gold-uranium project boost following TOMRA XRT trial* International Mining.

2020d. *XRT ore sorting shows promise at Vendetta Mining's Pegmont project*. International Mining.

2021a. *Osisko Development achieving hard rock cutting results with Sandvik MT720 roadheader plus more positive ore sorting testwork*. International Mining.

2021b. *Vast Resources to leverage new equipment and XRT ore sorting at Baita Plai*. International Mining.