-RESEARCH PAPER

International Journal of Mining and Geo-Engineering

IJMGE

IJMGE 59-4 (2025) 377-383

DOI: 10.22059/IJMGE.2025.395689.595261

Efficient Net-B3 for iron ore pellet quality control: real-world image classification with high accuracy

Marziye Salehi [a] and Asma Shams-Kermani [a, *]

[a] *Faculty of Electrical and Computer Engineering, Sirjan University of Technology, Sirjan, Iran.*

Article History:

Received: 20 May 2025.

Revised: 15 August 2025.

Accepted: 01 November 2025.

ABSTRACT

This study proposes a deep learning–based quality control method for iron ore pellet production using real-world image classification. An EfficientNet-B3 architecture classifies pellet images into four size categories: very small, small, medium, and large. Trained on 17492 images captured under realistic conditions, the model achieved a classification accuracy of 99.9%, outperforming alternative architectures, including ResNet50, VGG16, and MobileNet. Additional performance metrics, such as precision, sensitivity, and Matthews correlation coefficient (MCC) which were further confirmed the robustness of the approach. The results demonstrated the potential of deep learning for automating pellet size monitoring and highlight its industrial relevance for improving efficiency and quality in steel production.

Keywords: *Iron ore pelletization, Image classification, Deep learning, efficientnet, Quality control..*

1. Introduction

Quality control of iron ore pellets is a critical aspect of the iron and steel production industries, as pellet characteristics directly affect the quality of the final steel product. Pelletization is one of the most widely used methods for agglomerating of fine iron particles into pellets, which serve as the primary feedstock for steel production [1]. Consequently, pellet quality plays a decisive role in determining the performance and properties of the produced steel.

Whereas demand shows a regular upward trend, in developing countries due to their limited deposit base and over-mining, the natural quality of available iron resources shows a steady decline. Because mineral resources are non-renewable and consumption grows with industrial development, their use must be managed effectively and efficiently [2,3]. In the steel-making industry, iron powder is typically agglomerated into pellets, and pellet quality strongly influences the grading of the resulting steel. Critical quality parameters include pellet shape, size, and purity, and ideally, pellet size should be in range of 9 to 14 millimeters [1].

Controlling physical properties such as size, shape, and structural integrity is essential for optimized steel performance and quality. To meet this demand, advanced computational methods have become vital for real-time quality control, as traditional techniques like sieving or manual sampling are time-consuming and less reliable [4]. Other approaches, such as radiation-based classification, can be effective but involve hazardous radiation and high energy consumption [3].

In image-based methods, pellet size is determined from their images. Techniques used include thresholding, watershed, neural networks, and deep learning. Local thresholding segments images using a threshold at each point, requiring precise settings for acceptable results [5]. The watershed algorithm uses topological analysis to segment images into distinct regions based on image gradients [6]. Challenges in pellet image processing include poor lighting, cluttered backgrounds, overlapping pellets, light reflections, and the need for high-quality images. Real-time measurement speed is also critical. Integrating deep learning-based image analysis offers a breakthrough in efficient pellet quality evaluation [7].

This approach has transformed image processing in agriculture, healthcare, and manufacturing [8]. In mineral processing, deep CNN architectures such as ResNet, VGGNet, and EfficientNet have shown strong results in particle classification and quality assessment [9, 10]. EfficientNet balances computational efficiency and accuracy by uniformly scaling model dimensions [11].

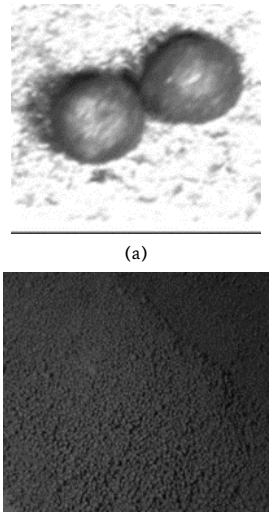
Besides the fact that automatic image-based quality control is currently under development, some drawbacks remain which should be resolved. Most deep learning models for pellet classification are trained on laboratory-generated datasets, which fail to capture real-world variability such as lighting changes, overlapping particles, and background noise. A U-Net-based deep neural network was designed for automatic pellet recognition from images [6]. The EfficientNet-B3 model, with fewer parameters than traditional U-Net, includes batch normalization layers, which significantly reduce computation time and improve network generalization. Reference [7] presents a machine learning algorithm to estimate pellet size classes during production, using images from a rotating disk. Instead of measuring each pellet individually, this method qualitatively estimates the overall size distribution of production pellets. However, most models face computational limitations, restricting their application in industrial real-time monitoring [12].

Reference [7] also classifies pellets on a rotating disk into four categories: very small, small, medium, and large. It uses a weighted ensemble of convolutional neural networks, including VGG16, MobileNet, and ResNet50. Experimental results show that image acquisition within the inner area of the pelletizing disk accurately reflects its operational status. Another reference introduces a dataset of pellet images of varying sizes, captured with an industrial camera during

* Corresponding author. *E-mail address:* as_shamsi@yahoo.com (A. S. Kermani).

production. In these images, the background is relatively simple, as shown in Figure 1(a), which depicts laboratory conditions with a small number of pellets. By contrast, Figure 1(b) shows pellets on a rotating disk under realistic conditions.

Most previous deep learning methods for pellet size determination have relied on laboratory images, limiting generalization to industrial environments. In contrast, this study classifies real-world pellet images, such as those in Figure 1(b), which enhances the research's practical utility.



(a)

(b)

Figure 1: Two images of iron ore pellets: (a) laboratory image, and (b) real-world rotating disc image.

This work aims to bridge these gaps by accurately classifying iron ore pellets using the EfficientNet-B3 CNN model. Transfer learning is applied to images captured from industrial pelletizing disks, enabling a more realistic simulation of operating conditions. The proposed approach achieves a classification accuracy of over 99.9%, outperforming classic CNNs such as VGG16, ResNet50, and MobileNet. This contribution enhances automatic quality control in iron ore pellet production, helping ensure both high quantity and quality of steel outputs.

The paper is structured with presenting the EfficientNet-B3 architecture. Then, the criteria and methods for evaluating its performance against comparative models is described which reports experimental results, and compares these results with those in [2] and [12]. Finally, and in last part summarizes the key findings, discusses the implications and limitations of EfficientNet-B3, and compares it with alternative models.

2. 2. Methodology

Convolutional Neural Networks (CNNs) automatically and efficiently capture spatial and temporal dependencies in images by convolving a set of kernels or filters. A typical CNN consists of three main types of layers: convolutional, pooling, and fully connected layers. The convolutional layers hierarchically extract nonlinear features: the initial layers detect general properties such as edges and bright spots, while deeper layers identify more complex and specific patterns. These layers apply filters that transform input images into feature maps, highlighting relevant structures.

Pooling layers then reduce the dimensionality of these feature maps, lowering the number of network parameters and computational cost. The resulting two-dimensional maps from the convolution and pooling stages are flattened into one-dimensional vectors in the fully connected layers. The extracted features can then be passed to a custom classification block, potentially including dropout and batch normalization layers, to improve classification performance [8, 1].

Deep models can achieve superior performance compared to traditional techniques by leveraging pre-trained weights rather than starting from random weights. Several CNN models pre-trained on the ImageNet dataset, containing about one million images across 1000 object categories, have been effectively applied in prior work [13]. Consequently, transfer learning is widely adopted to fine-tune these pre-trained models and address the challenges of limited training data [14].

2.1. EfficientNet-B3 CNN model

Model scaling is a key challenge in designing CNN architectures that balance accuracy gains from increasing model size with the risk of overfitting. Given computational constraints, model design should match available resources, with the option to scale up later as resources increase to achieve higher accuracy. Three main strategies improve model performance:

1. **Increasing Depth:** Adding more layers enables extraction of more complex features and hierarchical representations. However, greater depth increases the risk of overfitting and raises computational costs.
2. **Widening:** Increasing the number of neurons or filters per layer expands the network's capacity to learn detailed features. This can improve representational power but often requires careful tuning to avoid excessive computational demands.
3. **Resolution Enhancement:** Using higher-resolution input images or feature maps can yield more accurate, detailed outputs, particularly in high-resolution tasks, but at the cost of greater memory and computation requirements.

Effective CNN scaling seeks to improve accuracy while conserving resources. Techniques such as progressive resizing optimize both performance and efficiency. Research shows that uniformly scaling depth, width, and resolution by a constant factor is more effective than tuning each dimension separately, yielding consistent accuracy improvements and better efficiency.

In many domains, existing models still suffer from high computational requirements and large parameter counts, limiting their practicality. Thus, improving accuracy while reducing parameters and FLOPS is essential. EfficientNet addresses this challenge, holding the current state-of-the-art in ImageNet classification with 84.4% top-1 accuracy using only 66 million parameters. The EfficientNet family includes eight models (B0–B7), each with progressively more parameters and higher accuracy.

In this work, the EfficientNet-B3 was applied and Figure 2 illustrates its network architecture. A core building block is the MBConv layer, conceptually similar to the inverted residual blocks in MobileNetV2.

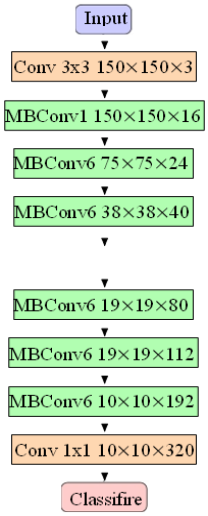


Figure 2: Schematic representation of EfficientNet-B3.

An MBConv convolutional block uses a direct skip connection between its input and output. Its internal convolutions first expand the input activation maps, increasing the depth of the feature maps, and then apply a compression step to reduce the number of output channels. The narrower layers are linked by skip connections, while the wider, high-channel layers are positioned between these connections. This design is both computationally efficient and results in smaller model sizes with reduced overhead.

Additionally, MBConv blocks employ depthwise separable convolutions, which factorize standard convolutions into depthwise and pointwise operations. This greatly reduces computational cost and the total number of parameters, further improving model efficiency. By using MBConv blocks, EfficientNet achieves balanced scaling of model size, depth, and resolution, maximizing performance while minimizing computational demands.

The model used in this work is EfficientNet-B3, recently introduced by Google and publicly available via GitHub. EfficientNet-B3 belongs to the EfficientNet family models pre-trained on the ImageNet dataset and scaled according to model complexity, resource availability, and target objectives. In this study, EfficientNet-B3 was fine-tuned for the classification task and run on an NVIDIA GeForce GTX 1080 Ti GPU. Implementation was carried out in Python, with PyTorch used for tensor operations.

3. Results

The section begins with a description of the dataset, followed by the evaluation metrics. The results of the EfficientNet-B3 model are then presented and compared with alternative approaches using clearly defined performance criteria.

3.1. Database

The dataset used in this work was first introduced in [2]. It contains color images of iron ore pellets captured from a rotating disk. Camera settings and image acquisition details are provided in Figure 3, and a sample image is shown in Figure 1(b). The dataset comprises 17492 images, each with a resolution of 224 × 224 pixels, evenly divided into four size classes: very small, small, medium, and large (4373 images per class). Each class is further split into 2359 training images, 1004 testing images, and 1010 validation images.

Figure 3. Camera installation on the rotary disc used in the data collection process as described in reference [2].

3.2. Evaluation metrics

To assess the performance of the EfficientNet-B3 model against benchmark methods, five metrics were used: precision, accuracy, sensitivity, F1-score, and the Matthews correlation coefficient (MCC), defined in Equations (1) to (5).

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \tag{1}$$

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \tag{2}$$

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP+FN}} \tag{3}$$

$$\text{F1-score} = \frac{2\times\text{Precision}\times\text{Sensitivity}}{\text{Precision+Sensitivity}} \tag{4}$$

$$\text{MCC} = \frac{(\text{TP}\times\text{TN}-\text{FP}\times\text{FN})}{\sqrt{(\text{TP+FP})\times(\text{TP+FN})\times(\text{TN+FP})\times(\text{TN+FN})}} \tag{5}$$

Here, TP represents True Positives, FP represents False Positives, TN represents True Negatives, and FN represents False Negatives.

3.3. Results and discussion

This section presents a detailed evaluation of the EfficientNet-B3 model, addressing both its effectiveness and comparisons with benchmark methods. The experiments were conducted on the dataset described in subsection 3.1 and benchmarked against state-of-the-art models, including VGG16, ResNet50, and MobileNet. Additional insights from References [2] and [12] are incorporated, where these models, along with YOLOv8, were applied to iron ore pellet image classification tasks.

Figure 4 shows the training and validation accuracy curves over multiple epochs. Accuracy increased rapidly during the initial epochs and stabilized after the 8th epoch. Validation accuracy closely matched training accuracy, indicating strong generalization. The model achieved peak accuracies of 0.999 (validation) and 0.998 (training).

Figure 5 depicts the cross-entropy loss curves for training and validation. Loss values were already relatively low in the first epoch, demonstrating effective early learning. Gradients converged after the 8th epoch, as indicated by stabilized loss curves. The close alignment of validation and training loss further confirms the model's robust generalization to unseen data.

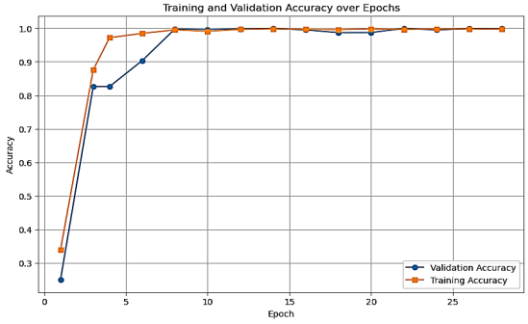


Figure 4: Training and validation accuracy over epochs.

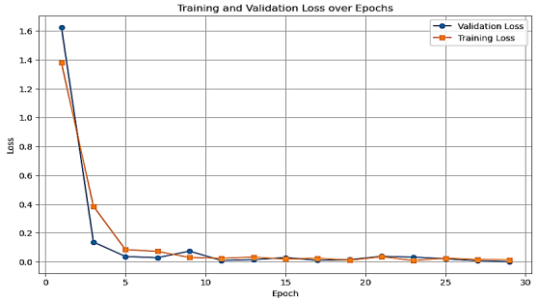


Figure 5: Training and validation loss over epochs.

Further details are presented in the following subsections. The confusion matrices, offering a fine-grained view of model performance across classes are presented and accordingly discussed. Then, the evaluation metrics used to assess the EfficientNet-B3 model's efficiency is explained. Together, these analyses highlight the model's strengths and limitations when benchmarked against existing approaches.

3.3.1. Confusion matrix

Confusion matrices provide a clear view of classification model performance by comparing true labels with predicted labels. Each cell indicates the number of samples for a specific true, predicted class pair, enabling evaluation of accuracy, precision, recall, and class discrimination ability. Figures 6 to 11 present the confusion matrices for five benchmark models and the EfficientNet-B3 model. Figure 11 shows the matrix for EfficientNet-B3, while the preceding figures depict ResNet50, MobileNet, VGG16, and the models reported in References [2] and [12]. The dataset contains 4016 images evenly distributed across four classes: very small, small, medium, and large, each with 1004 images. Model comparisons are based on the performance patterns evident in these confusion matrices.

ResNet50: Figure 6 shows the confusion matrix for ResNet50. The model performs strongly for the 'large' class, achieving the highest correct predictions with minimal misclassification. However, notable errors occur in the 'medium' and 'small' classes, with 'small' often misclassified as 'medium'. Overall, while ResNet50 performs well, it struggles to distinguish between smaller and medium-sized classes.

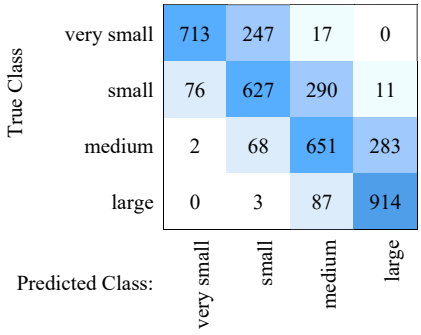


Figure 6. Confusion matrix of the ResNet50.

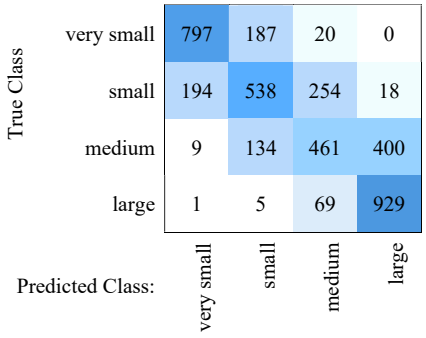


Figure 7. Confusion matrix of the Mobilenet.

VGG16: Figure 8 presents the confusion matrix for VGG16. The model classifies the 'large' class with high accuracy and few misclassifications. Performance is also good for the 'very small' class, though some confusion with other categories remains. Significant misclassifications occur between the 'small' and 'medium' classes, indicating a need for better separation of these categories.

Reference [2] model: For the Reference [2] model as well, the confusion matrix in Figure 9 demonstrates strong performance overall, with high scores across all classes, particularly for the class 'large'. This suggests that an ensemble-based approach can provide robust performance by aggregating multiple models. As a result, it produces fewer misclassifications than various individual models. However, the model still struggles to clearly differentiate between the 'very small' and 'small' classes.

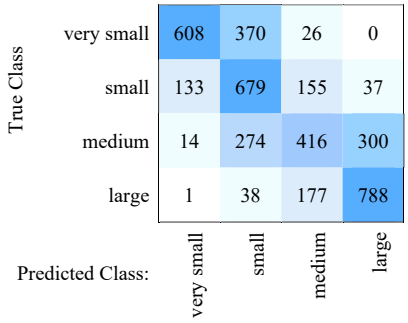


Figure 8. Confusion matrix of the VGG16.

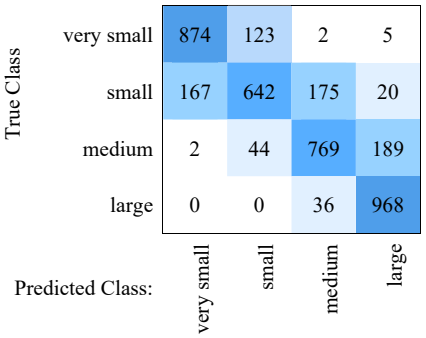


Figure 9. Confusion matrix of Reference [2] model.

Reference [12] Model: According to Figure 10, the confusion matrix of the Reference [12] approach shows near-perfect classification performance. This model achieves high accuracy across all classes with minimal errors. These results demonstrate that the method proposed in Reference [12] is highly effective in classifying all categories with high accuracy.

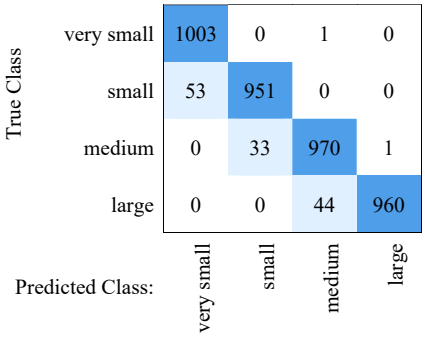


Figure 10. Confusion matrix of Reference [12] model.

EfficientNet-B3 model: Figure 11 presents the EfficientNet-B3 model confusion matrix, which demonstrates outstanding performance and near-perfect accuracy across all categories. The model shows minimal misclassification and is highly effective in distinguishing between all classes. Given this high accuracy, it can be regarded as the most effective among the evaluated methods.

In summary, although all models performed excellently in classifying the data, the EfficientNet-B3 model achieved the highest accuracy and the fewest misclassifications. The confusion matrices confirm that EfficientNet-B3 delivers the greatest precision and reliability in data classification, making it the most effective model among those compared. Its superior performance highlights its potential for practical applications requiring precise and dependable classification.

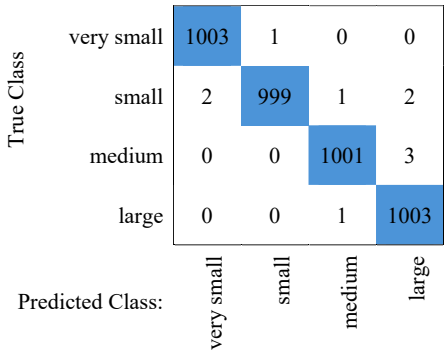


Figure 11. Confusion matrix of the EfficientNet-B3 model.

3.3.2. Evaluation metrics Results

The metrics from previous part (subsection 3.2) were computed for all models, and the results are summarized in Table 1. The EfficientNet-B3 model consistently outperforms the alternatives, achieving the highest values for sensitivity, precision, MCC, F1-score, and accuracy across all classes. For the 'very small' class, EfficientNet-B3 achieved 99.93% accuracy and 99.90% sensitivity. Accuracy in the 'small', 'medium', and 'large' classes was also high, 99.85%, 99.88%, and 99.85%, respectively. These results indicate that EfficientNet-B3 not only surpasses ResNet50, MobileNet, and VGG16, but also improves upon earlier reference results reported in the literature. Overall, this in-depth analysis underscores the significant progress represented by the EfficientNet-B3 model and its strong potential for real-world applications.

4. Discussion

The experimental results demonstrate that EfficientNet-B3 consistently outperforms ResNet50, MobileNet, VGG16, and previously published approaches, achieving near-perfect classification performance across all four size categories. Accuracy, sensitivity, MCC, and F1-scores exceed 99.8% for each class, highlighting the model's robustness and precision.

Although performance is exceptional, the confusion matrix (Figure 11) shows a few residual misclassifications, all occurring between adjacent size categories. This pattern can be explained by the dataset preparation method adopted from [2], where only 'very small' pellet images were captured in situ and the 'small', 'medium', and 'large' classes were generated by cropping the originals by fixed percentages (5%, 10%, 15%) and resizing. While this approach ensured balanced data, it also introduced subtle visual overlaps between classes, especially near category boundaries.

Cropping increments produce gradual rather than sharp differences in pellet scale, and real-world factors, such as dust, lighting variation, and pellet overlap patterns are inherited across all derived classes. As a result, borderline samples may share nearly identical spatial and textural characteristics, making them inherently ambiguous. The rare errors observed, such as 'small' classified as 'very small' or 'medium' as 'large', likely stem from these intrinsic similarities rather than model deficiencies.

Table 1: Performance metrics for different models across all classes.

Model	Class	Sensitivity	Precision	MCC	F1-Score	Accuracy
ResNet50	Very Small Small Medium Large All Classes	0.7102 0.6245 0.6484 0.9104 0.7234	0.9014 0.6451 0.623 0.7566 0.7234	0.745 0.5156 0.5108 0.7546 0.6272	0.7944 0.6346 0.6354 0.8264 0.7234	0.9081 0.8202 0.814 0.8909 0.8573
Mobilenet	Very Small Small Medium Large All Classes	0.7938 0.5359 0.4592 0.9253 0.6785	0.7962 0.6227 0.5734 0.6897 0.6785	0.7268 0.4506 0.3736 0.7017 0.5659	0.795 0.576 0.51 0.790 0.6785	0.8977 0.8028 0.7794 0.8564 0.8332
VGG16	Very Small Small Medium Large All Classes	0.6056 0.6763 0.4143 0.7849 0.6203	0.8042 0.4989 0.5375 0.7004 0.6203	0.6164 0.4115 0.3244 0.6337 0.4891	0.6909 0.5742 0.4679 0.7403 0.6203	0.8645 0.7493 0.7644 0.8463 0.8051
Reference [2]	Very Small Small Medium Large All Classes	0.8705 0.6394 0.7659 0.96410.81	0.838 0.7936 0.7831 0.81890.81	0.8043 0.6305 0.7 0.8486 0.742	0.8539 0.7082 0.7744 0.88560.81	0.9255 0.8683 0.8884 0.9216 0.8998
Reference [12]	Very Small Small Medium Large All Classes	**0.999** 0.9472 0.9661 0.9562 0.967	0.9498 0.9665 0.9557 0.999 0.967	0.9653 0.9426 0.9478 0.9663 0.955	0.9738 0.9567 0.9609 0.9771 0.967	0.9866 0.9786 0.98 0.9848 0.9824
EfficientNet-B3 model	Very Small Small Medium Large All Classes	0.999 0.995 0.997 0.999 0.9975	0.998 0.999 0.998 0.995 0.9975	0.998 0.996 0.9967 0.996 0.9967	0.9985 0.997 0.9975 0.997 0.9975	0.9992 0.9985 0.9988 0.9985 0.9988

Beyond classification performance, industrial implementation is a critical consideration. The dataset used in this study originates from an active iron ore pelletization plant, where images were captured from the inside area of a disc pelletizer under real production conditions. This ensures that the trained model has already been exposed to many of the environmental challenges typical of industrial operations, including dust, lighting variability, and background color changes. The image acquisition setup described in [2], a GigE camera with controlled illumination mounted 650 mm from the disc, demonstrates a practical, non-contact configuration suitable for production environments.

For deployment, several operational aspects must be addressed:

- **Real-time throughput:** Pelletization discs operate continuously, and classification must occur at high frame rates to provide timely process feedback. EfficientNet-B3's computational efficiency, combined with hardware acceleration (e.g., GPU-based processing or optimized inference on edge devices), can support this requirement.
- **Integration with process control systems:** Classification outputs (very small, small, medium, and large) can directly inform disc speed, feed rate, and moisture adjustments in near real time, enabling closed-loop control of pellet growth.
- **Robustness to environmental variation:** While the source dataset demonstrates tolerance to operational challenges, periodic recalibration and retraining with updated plant data will be essential for maintaining accuracy over time.
- **Hardware durability and maintenance:** Industrial-grade cameras and lighting must be used to withstand vibration, temperature changes, and dust ingress, with regular cleaning and alignment checks to preserve image quality.

By leveraging an architecture already tested in a real-world pelletization plant, the transition from laboratory validation to full industrial deployment is substantially de-risked. However, operational success will still depend on robust integration into plant automation frameworks and long-term model adaptability.

Consequently, and looking forward, several future development directions emerge. Integrating multi-modal data—such as hyperspectral imagery, 3D laser scanning, or weight measurements, could improve discrimination between visually similar size classes. Sensor fusion approaches might combine visual and non-visual measurements for greater robustness. Evaluating model generalization on datasets collected from different mines or under varied conditions would further test adaptability. In parallel, exploring model compression through pruning and quantization could enable deployment on resource-constrained hardware without compromising accuracy.

It is also important to acknowledge study limitations. The dataset originates from a single operational environment, which may not fully capture variability in ore type, texture, or imaging conditions across different industrial sites. All experiments were conducted using high-performance computing resources, which may not reflect the computational constraints in production settings. Additionally, the scope of this work was limited to image-based classification, without incorporating other sensing modalities that may provide complementary information.

Finally, the model explainability remains as an essential consideration for industrial adoption. In mineral processing environments, the transparency of decision-making is critical for operator trust, process validation, and regulatory compliance. While EfficientNet-B3 achieves near-perfect accuracy, its internal reasoning is inherently opaque due to the complexity of deep convolutional networks. Explainability techniques, such as Gradient-weighted Class Activation Mapping (Grad-CAM), can generate heatmaps to highlight the image regions most influential in the model's decision, enabling verification that the model is focusing on relevant pellet features rather than irrelevant background noise. More advanced methods, including Layer-wise Relevance Propagation (LRP) and SHapley Additive exPlanations (SHAP), can quantify the contribution of each pixel or feature to the output. Integrating such tools into the operational pipeline would allow real-time auditing of decisions, strengthen stakeholder confidence, and provide actionable feedback for refining both model performance and dataset quality.

Overall, the results position EfficientNet-B3 as a high-performance and industrially viable solution for particle size classification. By addressing residual error cases, optimizing deployment pipelines, incorporating explainability, and exploring multi-modal extensions, future research can further enhance the model's robustness, interpretability, and readiness for large-scale real-world applications.

5. Conclusions

Real-time monitoring of pelletizing disks in iron ore production is a complex task, crucial for ensuring both operational efficiency and the quality of the final pellet product. The proposed algorithms classify pellets into four size categories: very small, small, medium, and large, providing actionable insights into whether the disk operates under optimal conditions. For example, a persistent classification trend toward 'very small' or 'small' pellets may indicate excessive production of fine particles, which can negatively affect overall performance. In such cases, disk settings should be adjusted to allow particles to grow before discharge. Conversely, a steady increase in larger particles suggests that growth must be controlled to maintain the target size range. In this work, the EfficientNet deep learning model achieved over **99% accuracy**, outperforming previously proposed techniques. These results highlight the potential of advanced deep learning methods to significantly improve operational control and efficiency in industrial environments.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this article.

Acknowledgment

We acknowledge the Sharif High-Performance Computing (HPC) Center at Sharif University of Technology, Iran, for providing the computational resources and facilities that were crucial for the completion of this research project.

Ethics approval and data use consent

This research did not involve human or animal subjects, so ethics committee approval was not required.

Data availability

The data supporting the findings of this study are provided in [2] and are publicly available at https://bit.ly/3b4eTQ4.

References

[1] Duan, J., Liu, X., Wu, X., & Mao, C., (2020) Detection and segmentation of iron ore green pellets in images using lightweight U-Net deep learning network. *Neural Comput Appl,* 32, 5775-5790.

[2] Deo, AJ., Sahoo, A., Behera, SK., & Das, DP., (2022) Vision-based size classification of iron ore pellets using ensembled convolutional neural network. *Neural Comput Appl* , 34, 18629-18641.

[3] Liu, Y., Wang, X., Zhang, Z., & Deng, F., (2023) A review of deep learning in image classification for mineral exploration. *Miner Eng*, 204, 108433.

[4] Liao, C., & Tarng, Y., (2009) On-line automatic optical inspection

system for coarse particle size distribution. *Powder Technol* , 189, 508-513.

[5] Budzan, S., & Pawelczyk, M., (2018) Grain size determination and classification using adaptive image segmentation with shape-context information for indirect mill faults detection. In: Advances in Technical Diagnostics. *Proc 6th Int Congr Tech Diagn*, 6, 215-224.

[6] Ma, W.,Wang, L., Jiang, T., Yang, A., & Zhang, Y., (2023) Overlappingpelletsizedetectionmethod based on marker watershed and GMM image segmentation. *Metals*, 13, 327.

[7] Subramanyam, V., Patra, P., & Singh, MK., (2017) Automatic image processing-based size char- acterization of green pellets. *Int J Autom Smart Technol*, 7, 85-91.

[8] Kamilaris, A., & Prenafeta-Boldu, FX., (2018) Deep learning in agriculture: A survey. *Comput Electron Agric*, 147, 70-90.

[9] Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M. & Berg, A.C., (2015) ImageNet large scale visual recognition challenge. *Int J Comput Vis*, 115, 211-252.

[10] He, K., Zhang, X., Ren, S., & Sun, J., (2016) Deep residual learning for image recognition. *In: Proc IEEE Conf Comput Vis Pattern Recognit*, 770-778.

[11] Tan, M., (2019) EfficientNet: Rethinking model scaling for convolutional neural networks. arXiv.

[12] Salehi, M., & Shamsi-Koshki, A., (2024) Classification of iron ore pellets based on size using deep learning. In: *Proc 13th Iranian and 3rd Int Conf Comput Vis Image Process*. [In Persian]

[13] Weiss, K., Khoshgoftaar, TM., & Wang, D., (2016) A survey of transfer learning. *J Big Data*, 3, 1-40.

[14] Ashtiani, S.H.M., Javanmardi, S., Jahanbanifard, M., Martynenko, A., & Ver- beek, F.J., (2021). Detection of mulberry ripeness stages using deep learning models. *IEEE Access*, 9, 100380-100394.