

Article

Not peer-reviewed version

Development of a New Method for Measuring the Diameter of Iron Ore Pellets Using Digital Image Processing and Convolutional Neural Networks in Compliance with ISO 4698

[Rafael M. Campos](#) ^{*,†} and [Gustavo Maia De Almeida](#) [†]

Posted Date: 1 October 2023

doi: [10.20944/preprints202309.2179.v1](https://doi.org/10.20944/preprints202309.2179.v1)

Keywords: Convolutional Neural Networks; Computer Vision; Artificial Intelligence; Iron Ore Pellets; ISO 4698



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Development of a New Method for Measuring the Diameter of Iron Ore Pellets Using Digital Image Processing and Convolutional Neural Networks in Compliance with ISO 4698

Rafael M. Campos ^{†,*} and Gustavo M. de Almeida [†]

Postgraduate Program in Control and Automation Engineering (ProPECAut), Federal Institute of Espírito Santo (IFES), Serra, 29173-087, Brazil; gmaia@ifes.edu.br

* Correspondence: rmofati@gmail.com

† These authors contributed equally to this work.

Abstract: Iron ore processing involves critical steps that affect the quality of the final product. Determining the diameter of the pellets is the necessary initial step for volume measurement and, ultimately, for porosity and bulk density measurement, crucial characteristics for optimizing the burning process in the blast furnace. Traditional measurement methods using mercury present issues related to operator safety and environmental preservation, while the practices described in ISO 4698 standard require lengthy preparation and execution time. In light of environmental needs, operator safety, and the time consumed in tests, implementing a new method based on digital image processing and convolutional neural networks for measuring the diameter of burnt iron ore pellets is proposed. The methodology involves capturing images of the pellets using a high-resolution camera and utilizing digital image processing and neural networks capable of performing pixel-by-pixel object segmentation in the images, providing precise information about the pellets to calculate their volume automatically. The results were compared with those obtained using traditional methods, evaluating their conformity with the ISO 4698 standard. In conclusion, this study offers a new approach to measuring the volume of burnt iron ore pellets, providing accurate and reliable results in compliance with safety and environmental preservation standards.

Keywords: convolutional neural networks; computer vision; Artificial Intelligence; iron ore pellets; ISO 4698

1. Introduction

Technological innovations are making industrial processes increasingly agile, faster, more efficient, and more sustainable. In this context, it is important for companies to maintain the long-term efficiency and sustainability of their processes.

The establishment of ISO 14000, focusing on audit and environmental management, demonstrates the rise of the subject. This significance transcends environmental considerations and extends to economic dimensions as well. As stated by [1], it has evolved into a crucial prerequisite for participation in the global commodities market.

The values of bulk density and porosity of iron ore pellets, as well as the swelling of pellets when subjected to pressureless reduction, are essential characteristics in pelletization processes and are all related to the volume of the pellets. These characteristics directly impact the quality of the pellets and are essential both in the production stage and in the utilization of the pellets as raw material in the blast furnace process, as mentioned by [2].

In their research investigating the influence of porosity on the quality of iron ore pellets, [3] highlighted the significance of controlling this attribute throughout the process. They emphasized that porosity significantly affects both the reducibility index and compression strength, both of which are crucial factors determining pellet quality.

Concerning the significance of porosity and its correlation with strength, [3] asserts that as porosity increases, the strength of iron ore pellets diminishes. Failure to meet the steel mills' porosity criteria may result in pellet breakage during transportation and loading into the blast furnace, thereby reducing permeability and adversely affecting the final product's quality. It is noted that optimal or reference values for porosity can range between 25% and 35%.

The use of mercury in determining the volume of iron ore pellets, as outlined in ISO 4698, ISO 15968, and JIS M 8719 standards, is incongruent with contemporary global environmental concerns. Employing this method complicates the handling, reusability, storage, and disposal of tested pellets, posing hazards and incurring substantial costs for companies [4].

In consideration of environmental preservation imperatives and advancements in technology, alternative methods that eschew mercury in favor of other chemical elements and water have emerged. Nonetheless, challenges persist regarding the reproducibility of results obtained with diverse elements, as they may penetrate pellets to varying depths [5].

Concurrently, as the evaluation of iron ore pellet characteristics has evolved, notable progress has been made in Computer Vision and Artificial Intelligence, particularly with Convolutional Neural Networks (CNNs). These networks have the capacity to extract patterns from a set of images through pre-training and subsequently identify these patterns in subsequent images, as elucidated by [6].

For this study, the Mask R-CNN developed by [7] and the YOLOv8 model introduced by [8] were selected as CNN models. The former builds upon the Faster R-CNN network developed by [9] by simultaneously generating an object mask alongside a bounding box and conducting pixel-wise segmentation of objects in images. This mask proves indispensable as it fills pixels corresponding to each identified object, segregating it from other objects or the background within the same static image or video frame. The latter represents an evolution of the Yolov5 model released in 2020 by Glenn Jocher and lacks an accompanying research paper. The YOLOv8 model offers instance segmentation capabilities, a pivotal feature for this research, as previously mentioned.

2. Related Works

The work by [10] presents an image processing-based approach for estimating the size distribution of raw iron ore pellets. The proposed method is fully automated and utilizes image processing techniques to accurately measure and continuously monitor the pellet diameter distribution at the output of the pelletizing disk. The authors highlight that traditional methods used by operators to measure pellet size have some limitations, such as subjectivity and lack of precision. Therefore, the approach proposed in the work could be a viable alternative for mining companies seeking to improve the efficiency and quality of the pellet production process. Three different image processing techniques were tested for pellet identification: morphological segmentation, active contour, and thresholding method. The results showed that the active contour-based approach was the most accurate, with 99.1% of pixels classified correctly. However, the authors plan to extend this work to use a more intelligent technique to combine the results of the three techniques.

The article by [11] presents another approach for measuring the size distribution of raw iron ore pellets. The proposed method utilizes advanced image processing techniques, including dual morphological reconstruction and circle-scan, to segment overlapping pellets by finding the center of each pellet through the local maximum methodology, resulting in accurate and efficient measurements. Additionally, the method was compared with three other methods known in the literature, demonstrating its superiority by achieving an average accuracy of 94.3%. This approach overcomes the limitations of existing sampling and testing methods, which are known to be time-consuming and inaccurate, becoming a reliable and advanced alternative for analyzing raw iron ore pellets in the industry.

Similarly, [12] proposes an algorithm that uses an active contour-based approach to segment overlapping raw pellets. The results showed that the algorithm is capable of accurately segmenting this type of pellet grouping, improving size measurement accuracy, and reducing human errors in the

steel industry. The authors also suggest that the active contour-based approach can be applied in other areas involving the segmentation of overlapping objects in images.

One of the key quality characteristics of iron ore pellets, which must be measured during their manufacturing process, is the distribution of their diameter. Traditional methods of digital image processing are challenged by the complex constitution, sediments, and residues encountered, and ultimately by the differences in illumination and background during the capture of such images. Thus, [13] developed a method for the autonomous calculation of this distribution using a deep convolutional neural network with a U-NET architecture.

As the quantity of pores and cracks in iron ore pellets essentially has the same X-ray absorption, these features cannot be discriminated by intensity differences during tests of this model. With this information, [14] developed a way to discriminate between these two classes of discontinuity using a Deep Convolutional Neural Network. For this purpose, a network with the U-NET architecture was employed. The neural network used was trained using manually segmented image samples of the two mentioned classes. After optimizing the training parameters, the network was exposed to higher-resolution images of various pellets, successfully discriminating between pores and cracks, though errors in detecting cracks during the validation process were still encountered.

3. Theoretical references

3.1. ISO 3082 standard

The ISO 3082 standard is an important tool to ensure the quality of results obtained from sample analysis. It defines the minimum requirements for sample selection and preparation, as well as for the sampling process itself. The standard also provides recommendations for the selection of the sampling device and the quantity of material to be taken.

The ISO 3082 standard is based on statistical principles and includes a series of procedures to ensure the representativeness of the sample. The sampling process is divided into three main stages: sample selection, sample preparation, and actual sampling. In each of these stages, it is important to strictly follow the guidelines of the standard to ensure that the sample is representative of the bulk material.

Sample selection is the first stage of the sampling process. In this stage, it is necessary to determine the sample size and select the sampling location. Additionally, factors such as material homogeneity, size variation, and moisture distribution need to be considered.

Sample preparation is the second stage of the sampling process. In this stage, the material needs to be prepared for sampling. This includes fragmentation of the material, if necessary, and separation of foreign particles.

Actual sampling is the third and final stage of the sampling process. In this stage, a representative quantity of the material needs to be taken for analysis. The ISO 3082 standard provides recommendations on the quantity of material to be taken and the type of sampling device to be used.

For use in tests according to the ABNT ISO 4698 standard, as in the case of this study, Table 13 of the ISO 3082 standard defines a quantity of 4 test portions containing 18 pellets each, randomly chosen from a sample of at least 1Kg.

3.2. ISO 4698 standard

The ABNT ISO 4698 standard is an important standard for determining the volume of iron ore pellet samples, although its main objective is to achieve the swelling index, which is calculated as the difference between the volume of pellets before and after reduction in equipment that simulates the reduction zone of a blast furnace.

As mentioned in the previous section, batch sampling and preparation should be carried out according to the methods described in the ISO 3082 standard to maintain sample representativeness.

For the volume measurement of a sample portion, the most important topic in the ISO 4698 standard is section 6.6, which specifies that the volume measurement equipment must have an accuracy of 0.2ml or 0.2cm³ for each test portion. In addition to this topic, Annex B of the standard provides four methods for determining the volume of the test sample, two of which involve the use of mercury and are not preferable due to environmental concerns.

3.3. Convolutional Neural Networks and Their Further Developments

The Convolutional Networks, originally introduced by [15] and commonly referred to as Convolutional Neural Networks (CNN), represent a specialized category of Deep Neural Networks tailored for processing data structured as one-dimensional or multidimensional arrays, such as images. Images can be conceptualized as grids of pixels in two dimensions. In [15], was developed a multilayer Artificial Neural Network known as LeNet-5, capable of classifying handwritten digits based on small pattern representations extracted from the input images.

Subsequent advancements following [15] are elucidated by [16]. The author highlights that since 2006, various techniques have been devised to address the challenges associated with training deep CNNs. However, it was the introduction of the [17] architecture that marked the initial breakthrough. This network shared similarities with LeNet-5 by [15] but incorporated a greater number of layers. Furthermore, as noted by [16], the ResNet architecture [18], which secured victory in the ILSVRC 2015 competition, boasted a depth twenty times greater than AlexNet and eight times more extensive than VGGNet developed by [19]. Augmenting the number of layers in CNNs enhances the network's ability to approximate the activation function closer to the target with heightened non-linearity, resulting in improved representation of objects within the input images.

A new CNN configuration was proposed by [20] to extract distinct regions from a single image. This setup (Figure 1) was coined as the Region-based Convolutional Neural Network (R-CNN). Rather than attempting to classify a vast number of regions, this approach identifies around 2000 regions of interest within the examined images. Following this, a CNN analyzes these regions, and supervised learning models (SVM) are employed for the classification of each region. The authors' novel configuration yielded notable results, achieving a mean Average Precision (mAP) of 53.3%. Impressively, this outcome surpassed the best result achieved by other VOC 2012 competitors by more than 30%.

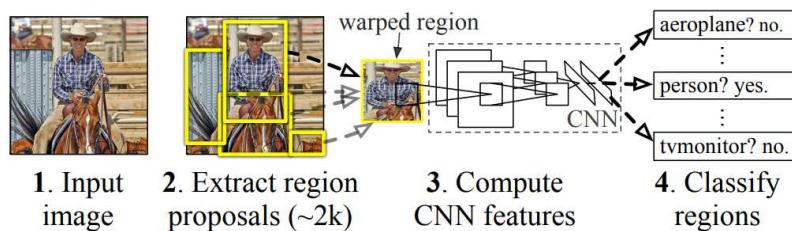


Figure 1. R-CNN model steps. Source: [20].

The subsequent research by [21], building upon his prior contributions and those of others, aimed to enhance the efficiency of object detection using Deep Convolutional Neural Networks (DCNNs). The preceding model suffered from several drawbacks, including multi-stage training, substantial computational costs, high memory utilization, and sluggish object detection. These associated limitations were mitigated by substituting the ConvNet, which was reinitialized for each object analysis, with a SPPnet (Spatial Pyramid Pooling network). This SPPnet harnessed a shared feature model among the objects under analysis. In comparison to Girshick's earlier work [21], the new model (Fig. 2) incorporated several innovations to accelerate training and testing, while also improving the accuracy of object detection. Notably, this novel model accomplished training the VGG16 network nine times faster and achieved higher mean Average Precision (mAP) values in the VOC 2012 challenge when compared to its predecessor.

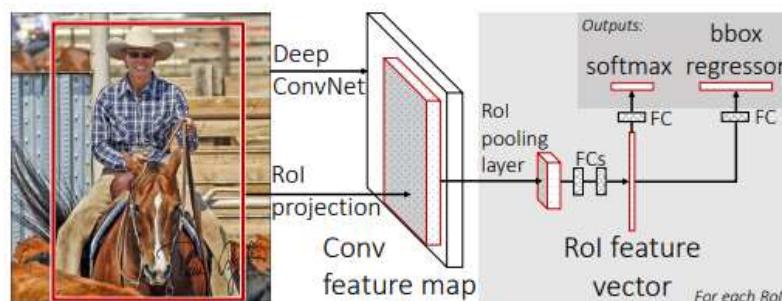


Figure 2. Architecture of the Fast R-CNN Model. Source: [21].

Enhancements to the Fast R-CNN model introduced by [21] were undertaken by [9] to facilitate real-time object detection, giving rise to the Faster R-CNN model. A significant reduction in processing overhead was achieved by sharing features among the networks, an approach inherited from the Fast R-CNN architecture. This adaptation of the model brought about almost real-time detection capabilities, with the notable exception of the time invested in region generation, which emerged as the primary bottleneck in detection systems. Addressing this challenge, the authors devised a Region Proposal Network (RPN) aimed at leveraging shared convolutional layers with the networks responsible for object detection. As the research progressed, it became apparent that the feature maps utilized by region detectors like Fast R-CNN could double as region proposers. By incorporating convolutional layers, the boundaries of each region were refined, culminating in improved object detection outcomes.

Utilizing the framework of the Faster-RCNN model, [7] introduced a novel model designed to execute object segmentation within images fed to the network by employing pixel-to-pixel masks. This innovative model, named Mask R-CNN, served as the foundation for the research presented in this work.

3.4. Mask R-CNN

The technique introduced by [7], named Mask R-CNN, extends the Faster R-CNN architecture initially proposed by [9] by concurrently developing an object mask alongside the bounding box. This generated mask precisely covers the pixels corresponding to the identified object, effectively segregating it from other objects within the same static image or analyzed video frame. While the original Faster R-CNN by [9] was not inherently designed for pixel-level alignment between network inputs and outputs, a straightforward and quantization-free modification known as RoIAlign was introduced. This RoIAlign layer faithfully preserves exact spatial locations, thereby directly enhancing mask accuracy, yielding improvements ranging from 10% to 50%. With these enhancements in place, the revised method outperformed all competitors in the 2016 COCO challenge, achieving a frame rate of 5fps on a GPU (equivalent to 200ms per video frame).

As described by [7], the object detection system in images comprises four distinct modules. The first module is responsible for extracting features from each object within the images. The second module employs a CNN to generate Regions of Interest (RoIs) based on the outputs of the preceding feature extraction module. Essentially, this is accomplished through the features extracted earlier. The third module employs another CNN to classify the objects present in each RoI generated by the second module. The final module is tasked with segmenting the objects within the image, achieving this through the application of a binary mask to each object.

3.5. YOLO - You Only Look Once

The YOLO (You Only Look Once) neural network is a computer vision model capable of real-time object detection. Its first version, YOLOv1, was proposed by [22], in a paper titled "You Only Look Once: Unified, Real-Time Object Detection" and published at the CVPR computer vision conference.

Before YOLO, object detection models would divide the image into regions and apply the classifier to each of them, which was slow and imprecise. YOLO, on the other hand, takes a different approach: it divides the image into a grid and, for each cell, predicts a set of bounding boxes and the probabilities of each box containing an object of a specific class.

YOLOv1 was trained on an object detection dataset known as VOC (Visual Object Classes) and achieved an average precision of 63.4% in real-time object detection. However, the model had some limitations, such as difficulty in detecting small objects and bounding boxes overlapping with nearby objects.

To improve performance, [23] introduced YOLOv2, which included improvements such as the use of Batch Normalization to stabilize training, the utilization of a residual network to learn deeper features, and the incorporation of a technique called Multi-Scale Training to enhance object detection at different scales.

In 2018, [24] released YOLOv3, which introduced the idea of multi-scale detection, allowing the model to detect objects of different sizes and aspect ratios more accurately. Additionally, YOLOv3 also utilized the concept of a Feature Pyramid Network to capture features at different scale levels and improve object detection accuracy.

Since then, various variations of YOLO have been proposed by researchers. [25], for instance, introduced a series of advanced techniques to improve object detection accuracy and speed, including the use of a deeper network architecture, training on high-resolution images, and employing a more advanced data augmentation mechanism.

Other variations include YOLO Nano, a smaller and faster model aimed at devices with limited processing capacity, and YOLOv5, released in 2020, which employed a completely different approach based on single-stage object detection and advanced data augmentation techniques. This version, along with the most recent YOLOv8, was developed by the Ultralytics company, with no published papers available at the time of writing.

Since its first version in 2016, several advancements have been made to improve the accuracy and speed of object detection, making YOLO a suitable tool for a variety of applications such as video surveillance, robotics, autonomous driving, and object recognition in medical images.

4. Motivations

The development of this work was driven by several motivations, all aimed at improving the efficiency, safety, and sustainability of industrial processes related to iron ore pelletizing.

One of the main motivations was the growing recognition of the importance of environmental management in the global commodities market. The ISO 14000 standard, which deals with environmental management and auditing, highlighted the need for companies to prioritize sustainability as a fundamental requirement for market participation. With this in mind, the use of mercury and petroleum derivatives in determining the volume of iron ore pellets, as prescribed by ISO 4698 and other standards, has become a concern due to its environmental impact and the challenges associated with handling and disposal. Finding alternative methods that are both eco-friendly and cost-effective has become imperative.

Furthermore, the quality of iron ore pellets is directly influenced by their porosity and apparent density, which are characteristics that rely on accurate volume measurements. Achieving optimal levels of porosity is critical for pellet strength, reducibility, and overall quality. Deviations from the desired porosity range can result in breakage during transport and loading into blast furnaces, impacting product quality and the efficiency of the production process.

Advances in Computer Vision and Artificial Intelligence, particularly Convolutional Neural Networks (CNNs), have provided an opportunity to improve the analysis of iron ore pellets. CNN models such as Mask R-CNN and YOLOv8 offer instance segmentation capabilities, allowing the identification and differentiation of pellets, other objects, and the background. By leveraging these

advanced models, it becomes possible to automate and improve the accuracy of pellet analysis, reducing manual intervention and increasing overall process efficiency.

The combination of environmental concerns, the need for precise porosity control, and advances in computer vision technologies served as the driving forces behind this research. By replacing mercury-based methods and automating the analysis process, the aim was to achieve more sustainable and efficient pellet production while meeting the industry's stringent quality requirements. The results of this work have the potential to contribute to the optimization of iron ore pelletizing processes, reducing the environmental impact, and improving overall operational performance.

5. Materials and Methods

5.1. Samples

The samples analyzed in this research were provided by the metallurgical laboratory of a major iron ore mining company in Brazil. A total of four sample sets were supplied, with each set consisting of 18 pellets randomly selected from a larger batch of over 1 kilogram. The random selection and adequate sample size ensured representativeness and followed the ISO 3082 standard recommendations for iron ore pellets. The pellets came from different stages of the iron ore pelletizing process, specifically production, loading, shipping, and client site. Obtaining samples from various process stages provided data encompassing the variability present across production.

For all sample sets, the mass of each pellet was measured using a properly calibrated precision balance capable of accuracy up to four decimal places. The metallurgical laboratory conducted reference volume measurements on the pellets following method B.4 of ISO 4698 standard, which involves water immersion. Based on these reference volume results, the laboratory calculated the diameter of the pellets using the standard sphere volume equation. These reference measurements represented the baseline data for comparison with the image-based measurement approaches investigated in this study.

5.2. Prototype

A semi-automated prototype was developed to enable consistent and efficient capturing of pellet sample images. The prototype was fabricated using 3D printing technology and its design incorporated several components optimized for multi-position imaging. An illuminated rotating stage formed the base of the prototype. This white translucent acrylic stage could perform a complete 360-degree rotation along its central axis in precise increments of 24 degrees. The rotational motion was driven by a stepper motor that allowed controlled incremental imaging of the pellet samples.

A Raspberry Pi HQ camera with a 12-megapixel resolution was mounted at a 45-degree angle and 10 cm distance from the center of the rotating stage. This positioning ensured optimal framing and clarity in the images captured. The prototype was outfitted with both front and back lighting elements. The backlighting created shadow images that precisely isolated the pellets from the background, while the front lighting provided standard RGB images.

The Jetson Nano developer kit integrated into the prototype enabled automated control of the imaging process and rotational motion. In summary, the thoughtful design and integration of components in this semi-automated prototype enabled the consistent and efficient generation of multiple images per pellet sample in a variety of positions and orientations. This capability facilitated the automated generation of a robust dataset of diverse pellet images to support the research objectives.

5.3. Dataset

Using the prototype, a novel dataset was produced containing 6480 images of segmented pellets. For each of the 18 pellets in a sample set, 90 distinct images were captured from different positions and orientations. At each stage position, paired RGB and backlit shadow images were acquired before incrementing the rotation.

The backlit shadow images provided precise shading to delineate the contour and shape of the pellets. Otsu thresholding appropriately isolated the pellet from the white background. A custom Python algorithm was developed to order the pellet boundary pixels in a clockwise sequence. This order was an important requirement for the convolutional neural network models utilized later in the research.

The dataset was automatically annotated to provide the x-y coordinates tracing the pellet boundaries in each image. For the Mask R-CNN model, COCO JSON formatted files containing this annotation data were created. For YOLOv8, the existing YOLOv5 PyTorch TXT annotation format was used, necessitating conversion of the Mask R-CNN data via the Roboflow online service.

The images were divided into training (80%, 5184 images) and validation (20%, 1296 images) sets. The considerable variety of pellet poses and orientations represented in the dataset improved the generalization of the models. Moreover, the automated image capture and annotation process boosted efficiency compared to manual approaches.

A critical prototype calibration step established the pixel-to-millimeter ratio for converting diameter measurements from pixels to real-world units. This involved imaging a 12 mm diameter calibration sphere and calculating the pixel distance between edges. The calibration provided the ability to obtain accurate physical measurements from the image data.

In summary, the heterogeneous sample sets provided representative test data encompassing process variability in pellet production. The novel semi-automated prototype enabled consistent and efficient image acquisition to build the dataset. Precise calibration ensured reliable real-world measurements from pixel data. The dataset itself offered standardized training and evaluation data with pellet diversity and accurate annotations. Together, these elements provided a robust foundation for developing and evaluating the pellet measurement approaches using digital image processing and convolutional neural networks.

5.4. Model training

To construct the neural network algorithm, the online platform Google Colaboratory (Colab), built upon the Jupyter Notebook framework, was leveraged. This selection stemmed from its user-friendly interface for incorporating third-party packages and effortlessly accessing neural network files, which could be directly cloned from GitHub servers and stored on Google Drive. Furthermore, this platform offers restricted access to GPUs hosted on dedicated Google servers, adding to its utility and convenience.

At the time of writing this work, the servers predominantly used Nvidia Tesla T4 model GPUs, which significantly reduced the training time and allowed for the necessary inferences during this development.

The Mask R-CNN and YOLO models were trained for 3500 and 9000 epochs, respectively. Both models used the same dataset, which comprises 6480 instances of pellets, as mentioned previously.

6. Results

6.1. Results Using Digital Image Processing

Upon subjecting the images of pellet samples provided by the mining company to digital image processing analysis, the summarized outcomes illustrated in Table 1 were attained.

Table 1. Volume Results Achieved Using PDI.

Sample	Reference (ISO 4698) (cm ³)	Lower Acceptance Limit (cm ³)	Upper Acceptance Limit (cm ³)	PDI (cm ³)
Sample 1	15.99	15.79	16.19	16.08
Sample 2	16.21	16.01	16.41	16.02
Sample 3	14.31	14.11	14.51	14.13
Sample 4	13.49	13.29	13.69	13.36

The volume value was calculated using the sphere volume Formula (1), based on the diameter provided by the equipment developed during the tests conducted with each set of samples.

$$V = \frac{4}{3}\pi \left(\frac{d}{2}\right)^3 \quad (1)$$

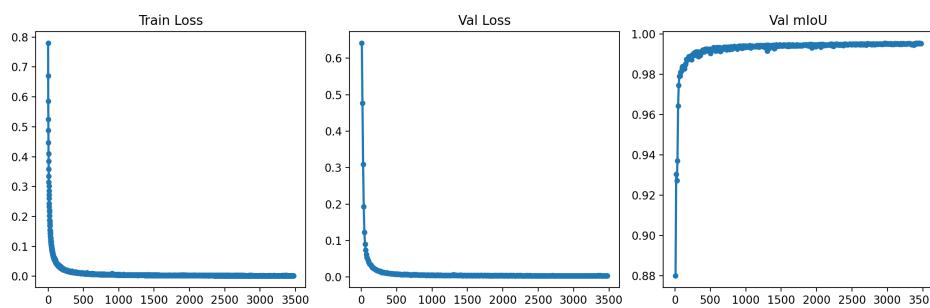
V: Pellet Volume

d: Pellet Diameter

6.2. Results Using Neural Network Models

6.2.1. Mask R-CNN Model

The developed configuration yielded promising outcomes after 3500 epochs of training using the mentioned dataset, with a duration of 14 hours on Google Colab GPUs. Throughout this period, low loss was observed during both the training and validation phases. Moreover, the model exhibited an average Intersection over Union (IoU) of 99.4%, indicating excellent delimitation and segmentation capability of regions of interest. These results imply the effectiveness of the Mask R-CNN model in the undertaken task, providing a detailed analysis of visual information and potentially reliable outcomes. The training progress is graphically depicted in Figure 3.

**Figure 3.** Loss and average IoU graphs during training and validation of the Mask R-CNN network.

After detecting and segmenting the pellets in each image using the Mask R-CNN model, as depicted in Figure 4, diameter measurements in pixels for each pellet were conducted.

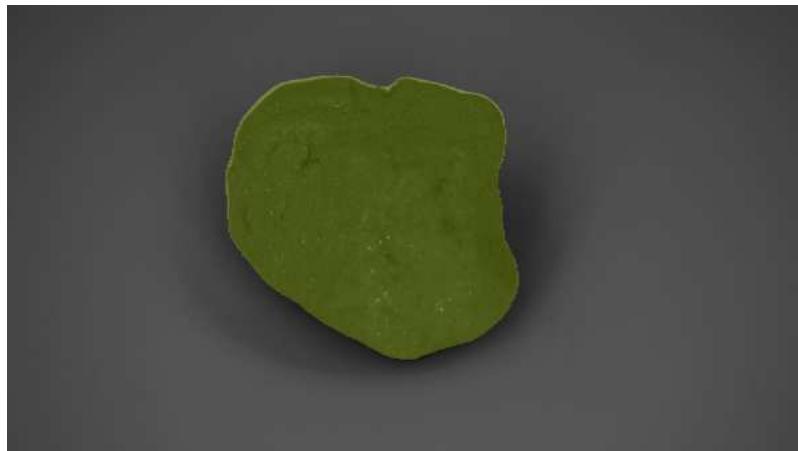


Figure 4. Sample pellet segmented by the Mask R-CNN model.

Upon detecting the pixel-based diameter results for each iron ore pellet sample subjected to the Mask R-CNN analysis, it was feasible to compute the actual diameter in millimeters for each pellet and subsequently determine their corresponding volume.

Table 2. Results Achieved Using the Mask R-CNN Model.

Sample	Reference (ISO 4698) (cm ³)	Lower Acceptance Limit (cm ³)	Upper Acceptance Limit (cm ³)	Mask R-CNN (cm ³)
Sample 1	15.99	15.79	16.19	16.77
Sample 2	16.21	16.01	16.41	17.04
Sample 3	14.31	14.11	14.51	15.16
Sample 4	13.49	13.29	13.69	14.36

6.2.2. YOLO Model

Using the YOLO model, promising results were achieved after 9000 epochs of training with the mentioned dataset. The training process took approximately 26 hours on Google Colab GPUs. Following training, low loss was observed in both the training and validation phases. Furthermore, the model exhibited an average Intersection over Union (IoU) of 99.5%, indicating exceptional delimitation and segmentation capability of regions of interest. These results underscore the efficacy of the YOLO model in the specific task, providing comprehensive visual analyses and presumably reliable outcomes. The obtained results are graphically shown in Figure 5.

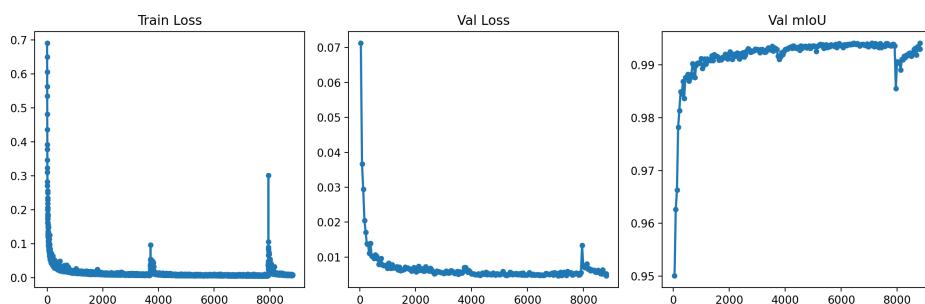


Figure 5. Loss and average IoU graphs during training and validation of the YOLO network.

Similar to the previous model, after detecting and segmenting the pellets in each image exposed to the YOLO model, diameter measurements in pixels for each pellet were conducted.

By identifying the pixel-based diameter values obtained for all iron ore pellet samples subjected to the neural network analysis, it became possible to compute the true diameter in millimeters for each pellet and subsequently determine the corresponding volume.

Table 3. Results Achieved Using the YOLO Model.

Sample	Reference (ISO 4698) (cm ³)	Lower Acceptance Limit (cm ³)	Upper Acceptance Limit (cm ³)	YOLO (cm ³)
Sample 1	15.99	15.79	16.19	17.05
Sample 2	16.21	16.01	16.41	17.34
Sample 3	14.31	14.11	14.51	15.41
Sample 4	13.49	13.29	13.69	14.52

6.3. Results Under Low Illumination

Analyses were performed under low illumination conditions to verify the outcomes using the three methods previously highlighted. In Figure 6, image (a) displays the original image captured under low illumination. Image (b) presents the result of the analysis using the PDI method. Images (c) and (d) showcase the outcomes obtained using the Mask R-CNN and YOLO models, respectively.

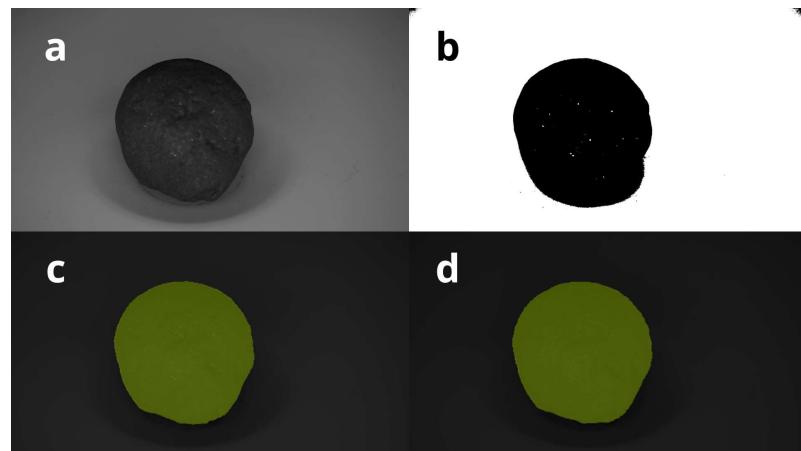


Figure 6. Analyses conducted under low illumination. Original image with low illumination (a). Segmentation performed by PDI (b). Results achieved using the Mask R-CNN and YOLO models, respectively (c) and (d).

These analyses were conducted to assess the methods' capability to process and detect objects accurately in scenarios with insufficient lighting. The results obtained provided insights into the effectiveness and applicability of the employed methods, enabling a comprehensive understanding of their performance in challenging lighting conditions.

7. Discussion

7.1. Diameter Measurement and Consequently, Volume

Upon obtaining the results through Digital Image Processing methods and the Mask R-CNN and YOLO neural network models, detailed analyses were performed on the outcomes achieved by each method, comparing them to the mining company's sample data.

Analyzing the results obtained from the measurements of samples using different methods yielded significant information about the accuracy and reliability of these techniques, as shown in Table 4.

Table 4. Absolute Errors Encountered in the Developed Methods.

Sample	Reference (ISO 4698) (cm ³)	Difference PDI (cm ³)	Difference Mask R-CNN (cm ³)	Difference YOLO (cm ³)
Sample 1	15.99	0.09	0.78	1.06
Sample 2	16.21	-0.19	0.83	1.13
Sample 3	14.31	-0.18	0.85	1.10
Sample 4	13.49	-0.13	0.87	1.03

It is observed that all measurements obtained through Digital Image Processing are within the acceptance limits. This indicates that the measurement method based on Digital Image Processing is accurate within the specified tolerances. This finding suggests that the Digital Image Processing method is a viable and reliable option for measuring the volume of the studied samples.

On the other hand, it is important to note that all measurements obtained through the Mask R-CNN and YOLO models are above the upper acceptance limit, as shown in Table 4. This observation reveals that these two methods consistently overestimate the measurements compared to ISO 4698 reference measurements. This inconsistency might indicate limitations or challenges faced by the Mask R-CNN and YOLO models in the task of measuring the studied samples. This limitation can be verified in Figure 4, where the neural network models often segment background pixels as if they were pixels of the pellet sample.

Analyzing the specific results of each method, it can be seen that the YOLO model provided the highest results for all samples, followed by the Mask R-CNN model, while Digital Image Processing provided results closest to the ISO 4698 reference. This difference in results suggests that each method has its own characteristics and possibly different levels of accuracy in its segmentation, which might be improved with an increased sample size in the dataset.

When comparing the differences between the results obtained by each method and the ISO 4698 reference, it is important to observe that the difference between the Digital Image Processing results and the reference is always smaller than the acceptable error established. This indicates that the Digital Image Processing method is reliable and produces measurements consistent with the ISO 4698 immersion method reference. However, the differences between the Mask R-CNN and YOLO model results about the reference are consistently larger than the acceptable error, suggesting that these methods might exhibit greater variability and uncertainty in measurements, mainly due to the segmentation errors mentioned earlier.

Another relevant observation is the tendency for the difference between the results of each method and the reference to increase as the ISO 4698 reference decreases. This trend might indicate a possible inverse relationship between these variables, i.e., the lower the reference value, the higher the discrepancy between the results obtained by the methods. This relationship could be investigated in future analyses to better understand the limitations and behavior of the measurement methods across different value ranges. It is thus suggested that in future work, regression techniques could be employed to correct the observed difference between the results of each method and the reference achieved by the ISO 4698 method. This approach can be explored in future work to adjust the results obtained by the Mask R-CNN, YOLO models, and even the Digital Image Processing method, to reduce the discrepancy with the reference.

In summary, based on the presented results and conducted analyses, it can be concluded that the Digital Image Processing method demonstrated precision and reliability in measuring the studied samples, providing results within acceptable limits. On the other hand, the Mask R-CNN and YOLO models exhibited consistent deviations from the reference, indicating the need to consider their limitations and uncertainties for this specific application. However, it is important to emphasize the possibility of addressing these deviations and seeking correction through the use of linear or quadratic regression techniques.

7.2. Measurement under Extreme Illumination Conditions

The results obtained through neural network-based methods, such as Mask R-CNN and YOLO, demonstrated significant advantages over the Digital Image Processing (DIP) method. Notably, these methods excelled in accurately segmenting objects under low illumination conditions.

When comparing the performance of the Mask R-CNN and YOLO models, it was found that Mask R-CNN outperformed in distinguishing deformities in the pellets with greater accuracy. Mask R-CNN was able to identify and segment the irregular areas of the iron ore pellets with greater accuracy. This demonstrates the model's ability to handle subtle variations in object shapes and textures, resulting in a more detailed and accurate analysis of deformities present in the pellets. This enhanced differentiation provided by Mask R-CNN highlights its superiority over the YOLO model in the specific context of detecting the shape of iron ore pellets.

By employing deep learning techniques, Mask R-CNN, and YOLO were able to comprehend and identify objects even in scenarios with low lighting, yielding superior results. These convolutional neural networks demonstrated their capability to handle extreme conditions and perform accurate segmentations, which is crucial in practical applications like image processing in low-light environments. Thus, methods based on neural networks, such as Mask R-CNN and YOLO, offer clear advantages over Digital Image Processing, providing superior and reliable results even in challenging lighting conditions.

8. Conclusions

This dissertation addressed the development of an innovative method for measuring the diameter and consequently the volume of burnt iron ore pellets, according to the specifications of ISO 4698, through digital image processing and the application of a well-known and widely disseminated deep convolutional neural network models.

The results obtained using the digital image processing testing method demonstrated high precision and reliability, meeting normative requirements and respecting safety and environmental preservation standards, as mercury is not used in its implementation.

Despite the safety and preservation standards also being found in the use of the Mask R-CNN and YOLO models, it is important to highlight that the same precision found in the Digital Image Processing method could not be reproduced. Although Mask R-CNN and YOLO were efficient in identifying the objects of interest, in this case, the burnt iron ore pellets, they exhibited discrepancies when it came to measuring the diameter and, consequently, the volume, in relation to the ISO 4698 reference.

Despite the fact that the Mask R-CNN and YOLO methods did not present results with precision similar to that of the Digital Image Processing method, they demonstrated the potential of deep convolutional neural networks in solving complex problems in the field of image processing, especially in situations that present greater complexity, such as the recognition and segmentation of burnt iron ore pellets in scenarios of low illumination.

The use of deep learning methods in the field of image processing represents a significant advancement in the area, with applications in a wide range of industries. The results obtained indicate that these methods are highly promising and can be optimized with improvements in the neural network architectures and by increasing the dataset size.

Finally, it is essential to mention that the validation of these methodologies in industrial environments should be carried out with great care, as this dissertation was developed under laboratory conditions. Nevertheless, the encouraging results obtained pave the way for future research and application in the mining sector, contributing to increased efficiency and safety in the production process, and to the preservation of the environment.

Funding: To the CAPES/FAPES Cooperation - Postgraduate Development Program - PDPG, through the project "TIC+TAC: Information and Communication Technology + Automation and Control Technology, Priority Intelligent Technologies," for the financial support of the research, by through FAPES/CNPq Notice No. 23/2018 - PRONEM (Term of Grant 133/2021 and Process No. 2021-CFT5C).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. D. S. de Oliveira, C. V. B. Nunes, L. de Jesus, and S. A. S. Loiola, "Impactos do mercúrio no meio ambiente e na saúde," 2011, pp. 5.
2. M. Cota Fonseca, "Influência da distribuição granulométrica do pellet feed no processo de aglomeração e na qualidade da pelota de minério de ferro para redução direta," Rede Temática em Engenharia de Materiais, 2004.
3. D. V. Bayão, R. A. L. Solé, J. J. Mendes, F. L. v. Krüger, and P. S. Assis, "Influência da porosidade na qualidade de pelotas de Minério de ferro," *ABM Proceedings*, pp. 152–164, Dec. 2018. [Online]. Available: <http://abmproceedings.com.br/ptbr/article/influncia-da-p\orosidade-na-qualidade-de-pelotas-de-minrio-de-ferro>.
4. Y. Omori and E. Kasai, "A new method for measuring apparent volume of iron ore pellets in conjunction with determination of the swelling index," *Transactions of the Iron and Steel Institute of Japan*, vol. 24, no. 9, pp. 751–753, 1984.
5. G. M. Silva, "Evolução da automação em ensaios de porosidade em pelotas de minério de ferro," *ABM Proceedings*, pp. 273–281, Oct. 2019. [Online]. Available: <http://abmproceedings.com.br/ptbr/article/evoluo-da-automao-em-ensaios-de-porosidade-em-pelotas-de\minrio-de-ferro>.
6. L. M. Gonzaga and G. M. de Almeida, "Artificial Intelligence usage for identifying automotive products," 2020, pp. 7.
7. K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," *arXiv:1703.06870 [cs]*, Jan. 2018. [Online]. Available: <http://arxiv.org/abs/1703.06870>.
8. Ultralytics, "YOLOv8," [Online]. Available: <https://ultralytics.com/yolov8>. [Accessed: 2nd February 2023].
9. S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *arXiv:1506.01497 [cs]*, Jan. 2016. [Online]. Available: <http://arxiv.org/abs/1506.01497>.
10. Mahdi Heydari, Rassoul Amirfattahi, Behzad Nazari, and Pooyan Rahimi, "An industrial image processing-based approach for estimation of iron ore green pellet size distribution," *Powder Technology*, vol. 303, pp. 260–268, 2016.
11. Xin Wu, Xiao-Yan Liu, Wei Sun, Chuan-Gang Mao, and Cao Yu, "An image-based method for online measurement of the size distribution of iron green pellets using dual morphological reconstruction and circle-scan," *Powder Technology*, vol. 347, pp. 186–198, 2019.
12. Mohammad Moghimi, Rassoul Amirfattahi, and Mohammad Khodabandeh Samani, "A new algorithm for segmenting overlapped green pellets to improve automatic pelletizing process," in *2017 10th Iranian Conference on Machine Vision and Image Processing (MVIP)*, pp. 82–86, 2017.
13. J. Duan, X. Liu, X. Wu, and C. Mao, "Detection and segmentation of iron ore green pellets in images using lightweight U-net deep learning network," *Neural Computing and Applications*, vol. 32, no. 10, pp. 5775–5790, May 2020.
14. E. T. V. Bezerra, K. S. Augusto, and S. Paciornik, "Discrimination of pores and cracks in iron ore pellets using deep learning neural networks," *REM - International Engineering Journal*, vol. 73, no. 2, pp. 197–203, Jun. 2020.
15. Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation Applied to Handwritten Zip Code Recognition," *Neural Computation*, vol. 1, no. 4, pp. 541–551, Dec. 1989.
16. J. Gu et al., "Recent advances in convolutional neural networks," *Pattern Recognition*, vol. 77, pp. 354–377, May 2018.
17. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, May 2017.

18. K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 770–778.
19. K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *arXiv:1409.1556 [cs]*, Apr. 2015. [Online]. Available: <http://arxiv.org/abs/1409.1556>.
20. R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," *arXiv:1311.2524 [cs]*, Oct. 2014. [Online]. Available: <http://arxiv.org/abs/1311.2524>.
21. R. Girshick, "Fast R-CNN," *arXiv:1504.08083 [cs]*, Sep. 2015. [Online]. Available: <http://arxiv.org/abs/1504.08083>.
22. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 779–788.
23. J. Redmon and A. Farhadi, "YOLO9000: Better, Faster, Stronger," *arXiv preprint arXiv:1612.08242*, 2016.
24. J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," *arXiv preprint arXiv:1804.02767*, 2018.
25. A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," *arXiv preprint arXiv:2004.10934*, 2020.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.