Deep learning for segmentation and classification of rock grains in the aggregates industry

Ricardo Ramos Nunes[1] , Renato Moraes Silva[2]

[1]Department of Computer Engineering
Facens University, Sorocaba, São Paulo, Brazil

[2]Institute of Mathematics and Computer Science
University of São Paulo (USP), São Carlos, Brazil

`ramos.rnunes@gmail.com, renatoms@icmc.usp.br`

Abstract. *This paper describes the application of a convolutional neural network (CNN) to detect rock grains in aggregates that exceed a specified size. In civil construction, the quality of aggregates is crucial and generally assessed by granulometry, with traditional sieving methods being time-consuming and susceptible to human error. This study proposes using machine learning to measure grain size continuously during the production process. Using CNN with the U-net architecture, the models were trained to evaluate grain size in a simulated condition that reproduces the environment of a conveyor belt. The results indicate that the developed models have strong generalization capabilities and can effectively identify contamination by rock grains that exceed the permitted size.*

1. Introduction

Aggregates are crucial in civil construction, providing the foundation for durable and stable structures. These materials, consisting of rock fragments such as sand, gravel, and crushed stone, are used in a variety of construction applications such as concrete production and pavement construction [Přikryl 2021, Associação Brasileira de Normas Técnicas 2005]. According to [Zhong et al. 2022], the global demand for sand and aggregates is projected to increase by 45% by 2060, highlighting the economic and strategic importance of the sector. In Brazil, the aggregates industry comprises approximately 3,100 companies, generating around BRL 34 billion in revenue in 2022 alone, as reported by Fernando Valverde, the Chief Executive Officer of the National Association of Entities of Aggregate Producers for Construction (Anepac) [MT Expo 2022]. The importance of construction aggregates is highlighted by [Přikryl 2021] who affirm that aggregates is the most voluminous mineral raw material exploited by humans nowadays. These figures highlight the economic impact and scale of the aggregates industry both in Brazil and globally.

The Brazilian standard ABNT NBR 7211 categorize aggregates into two types: fine aggregate, which is retained on a 0,15 mm sieve and passes through a 4.75 mm sieve, and coarse aggregate, which is retained on a 4,75 mm sieve and passes through a 75 mm sieve. In this context, the effectiveness of classifying rock grains by size is fundamental to ensure compliance with the technical standard ABNT NBR 7211 [Associação Brasileira de Normas Técnicas 2005] and to prevent structural failures [Přikryl 2021]. Moreover, there is a increasing demand for more efficient and precise methods in the granulometric analysis of aggregates, driven by the need for process optimization and cost reduction.

Traditionally, granulometric analysis of aggregates has been done manually, requiring human intervention and consuming both time and resources. However, the adoption of advanced computer vision and deep learning techniques, such as convolutional neural networks (CNNs), offers a unique opportunity to automate and opti-

mize this process. Deep learning is particularly effective in analyzing complex images. Specifically, CNNs can extract discriminating features from images, enabling accurate identification and classification of rock particles [Leiva et al. 2021]. For example, [Bamford et al. 2021] applied a CNN, more specifically ResNet50, to analyze blast-induced rock fragments from 2D images. According to the authors, in the experiments with piles containing very fine and very coarse fragments, the model's accuracy was lower compared to manual labeling. Nevertheless, for the majority of the piles used to evaluate the model, its accuracy was comparable to manual labeling performed by humans. These findings underscore the potential of CNNs in automating tasks related to granulometric analysis of rocks or similar correlated tasks.

Among the various CNN architectures available, U-Net has recently stood out [Ronneberger et al. 2015]. Originally proposed for different biomedical segmentation applications, this architecture offers the advantage of requiring few annotated images for training, as highlighted by [Ronneberger et al. 2015]. This characteristic makes it promising for aggregate analysis, given the scarcity of labeled datasets for this task and the time-consuming nature of the labeling process.

Given the potential of CNNs, particularly the U-Net architecture, in addressing the challenges of aggregate analysis, this study aims to apply the U-Net architecture to identify and classify rock particles from aggregate samples in images generated under simulated aggregate production conditions. Our hypothesis is that applying CNNs can result in consistent, high-quality aggregate products that meet market expectations. This work aims to: (1) develop and evaluate a U-net-based model for the semantic segmentation of gravel images, categorizing them into grain surface, grain edge or background; and (2) perform size measurements of the identified grains, classifying them as either Gravel 1 (fine gravel) or Gravel 2 (coarse gravel). It is important to clarify that the U-Net model will not directly classify the grains by size. Instead, size classification into Gravel 1 or Gravel 2 will be performed after the segmentation process, based on straightforward measurements of the segmented objects.

The remainder of this paper is structured as follows. Section 2 presents the main related work in the field. Section 3 details the materials and methods employed in the research. Section 4 covers the experimental setup and results. Finally, Section 5 concludes the paper and offers guidelines for future research.

2. Related work

Advancements in image processing and computer vision research, specifically in the context of mining and related disciplines, are continually evolving. Novel techniques and applications are regularly emerging in this field. Previous studies have highlighted the potential of image processing and deep learning for analyzing granular materials. A recent study by [Leiva et al. 2021] demonstrates ongoing progress in this research area, developing and validating an online analyzer for particle size distribution on conveyor belts in a copper ore plant, using image processing techniques such as noise reduction algorithms, morphological operations, and segmentation algorithms.

Furthermore, other studies applied advanced techniques to related problems. For example, [Wang et al. 2021b] applied hierarchical clustering to extract descriptors to use as input to traditional regression models (linear regression, ridge regression, random forest, extreme gradient boosting, and light gradient-boosting machine) to analyze grain crushing strength. In addition, [Maitre et al. 2019] employed traditional machine learning methods, such as classification and regression trees (CART), k-nearest neighbors (k-NN) and random forest, for mineral grains recognition.

Recent studies have applied CNNs [LeCun et al. 1995], a specialized type of neural network designed for processing data with a predefined grid-like topology, such as

images [Goodfellow et al. 2016]. In studies related to rock grains, CNNs have been employed, including ResNet50 for analyzing blast-induced rock fragments from images [Bamford et al. 2021], and Mask R-CNN for instance segmentation and shape evaluation of densely-packed particles like ballast and cobble [Yang et al. 2021].

In this study, we investigate the hypothesis that CNNs can also be employed to measure rock grains in an aggregate production line and identify those that exceed the allowed size limit. A fundamental step towards achieving this objective is the segmentation of grains captured in each image, allowing for the individualized measurement of each grain.

To achieve this goal, we chose the U-net architecture [Ronneberger et al. 2015], a widely recognized and utilized architecture for semantic segmentation tasks within CNNs. Since its introduction by [Ronneberger et al. 2015], it has become a benchmark in the field due to its adaptability to various segmentation tasks across different domains. Initially conceived for separating individual cells in biomedical images, its applicability has expanded beyond this context. In the field of grain detection, where factors such as size, appearance, and shape variations, along with potential partial or complete occlusion and inter-granular noise, present significant challenges, the use of U-net-based models can be advantageous. This choice is supported by other image segmentation research, such as the work by [Shi et al. 2022], which demonstrates an improved method of U-net image segmentation and its application to metallic grain size statistics.

Similarly, U-net has been effectively used in geological image segmentation by [Wang and Zai 2023] and [Hassan et al. 2024]. The first study focused on segmenting sandstone computed tomography images to evaluate petrophysical properties, such as permeability and flow velocity, employing metrics like accuracy, IOU, and MAE. In contrast, [Hassan et al. 2024] applied U-net with a ResNet 34 backbone for automatic mineral detection in rock thin-section images, achieving high precision, recall, and F1 scores across different datasets, specifically in the context of the Arabian-Nubian Shield. Despite the shared focus on image segmentation, these studies differ from our work, which specifically addresses granulometry, focusing on the identification and analysis of particle sizes in aggregate materials. Nonetheless, the successful use of U-Net in these related geological and material science contexts highlights its versatility and reinforces its potential for solving complex segmentation tasks in granulometric analysis, further validating its application in our study.

3. Methodology

The methodology employed in this study begins with collecting a sample for each of the products indicated in Figure 1, which illustrates the classification circuit of an aggregate plant from where the samples were taken. Analysis of Figure 1 reveals that the classification mesh separating Gravel 1 from Gravel 2 has an opening of 22 mm, resulting in distinct granulometry products, with Gravel 1 being finer and Gravel 2 coarser.

Initially, from the collected sample of Gravel 1, the direct grain evaluation method [Silva and Geyer 2018] was used, which consists of measuring the three principal axes of the grain, as shown in Figure 2a, to determine an estimated average area of the Gravel 1 grain in a two-dimensional plane. Measurements were taken on 180 randomly selected grains using a caliper to calculate the average area in square millimeters. The histogram of the estimated grain areas is presented in Figure 2b, indicating that all Gravel 1 grains have average areas between 200 and 600 square millimeters. This information will be subsequently used in the evaluation of the predictions in this study.

3.1. Unit conversion for grain analysis

The experiment was carried out in a controlled environment, allowing for the use of an object with a known size to determine a conversion factor from grain pixel parameters to
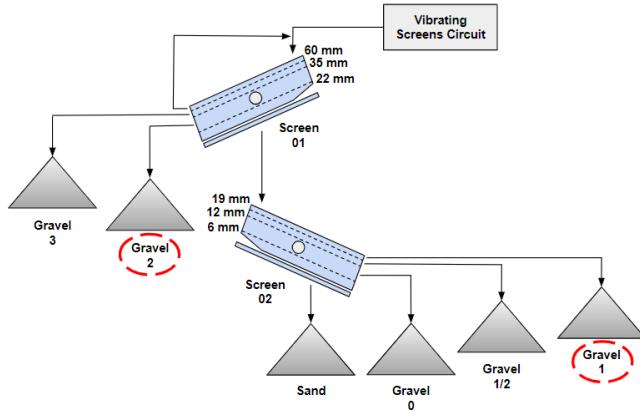
Figure 1. Representation of the aggregate classification circuit.

(a) General representation of the three main axes of a gravel grain (Source: [Silva and Geyer 2018]).

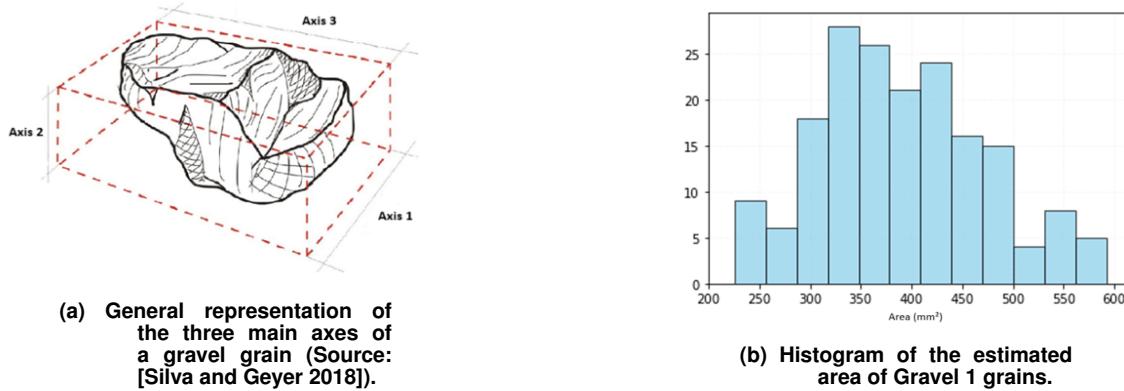(b) Histogram of the estimated area of Gravel 1 grains.

Figure 2. Characteristics of grains.

millimeters. Therefore, considering that the coin occupies an area of 285x285 pixels in the image set, as shown in Figure 3b, and has a diameter of 26.6 millimeters, as illustrated in Figure 3a, it is possible to calculate the area conversion rate as follows:

$$\text{Area Conversion Rate} = \frac{26.6\,\text{mm} \times 26.6\,\text{mm}}{285\,\text{pixels} \times 285\,\text{pixels}} = 0.0087\,\text{mm}^2/\text{pixel}^2 \qquad (1)$$

Once the unit conversion was performed, areas could be analyzed in millimeters, facilitating comparison to direct measurements.

3.2. Dataset collection and preparation

In order to create a set[1] of images that accurately present the grains to be analyzed, a prototype was employed. The prototype was based on the concept of a camera positioned above a conveyor belt, and it was designed to mimic the specialized solution offered by WEIR[2] for real-time particle analysis, as shown in Figure 4a. To ensure consistency in the images produced, the distance between the camera lens and the material was fixed at 300 millimeters, as seen in Figure 4b. This ensured that the scale of the grains in all of the images produced remained fixed.

---

[1] The dataset used for the experiments is publicly available at `https://github.com/renatosvmor/rock-grains`.

[2] The Weir Group is a multinational engineering company headquartered in Glasgow, Scotland.

(a) Measurement in millimeters.

(b) Measurement in pixels.

Figure 3. Scale transformation.

(a) WEIR company's real-time particle analysis solution (Source: [Weir Motion Metrics 2024]).
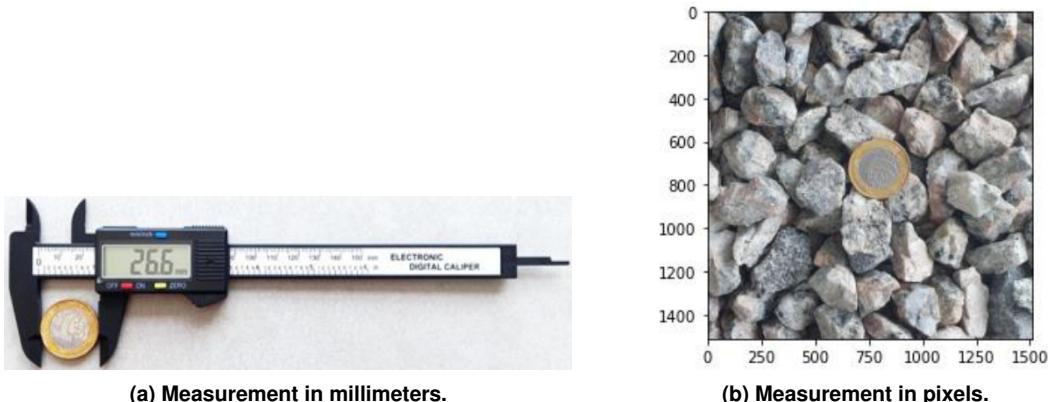

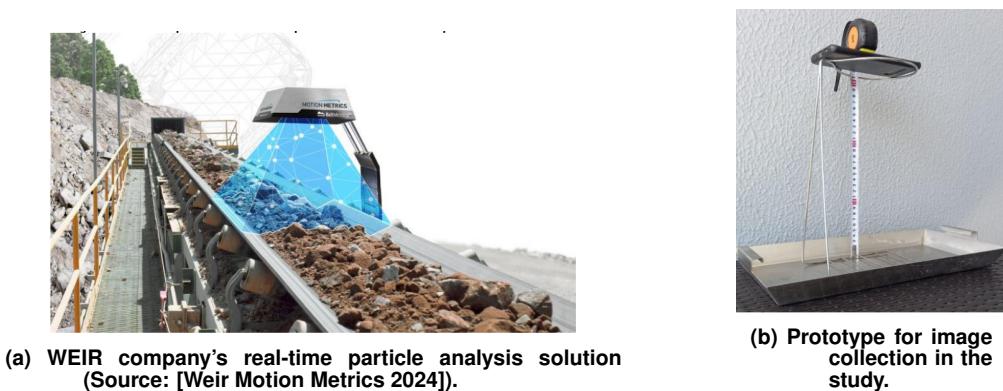**(b) Prototype for image collection in the study.**

Figure 4. Image capture system.

For this study, 49 images were produced, each with dimensions of 1280x1280 pixels. This dataset consisted of 22 images with Gravel 1 contaminated by Gravel 2, as illustrated in Figure 5b, and 27 images containing only Gravel 1 with no contamination, as illustrated in Figure 5a. Although the dataset is relatively small, it reflects the practical constraints and challenges of obtaining labeled data in a real-world industrial setting.

We performed the experiments using a holdout validation approach, allocating 60% of the data to the training set and 40% to the test set. This distribution differs from the typical 80/20 or 90/10 splits due to the small size of the dataset. The decision to use 40% for testing was made to ensure a more reliable and accurate assessment of model performance, as a smaller percentage would result in insufficient data for a robust evaluation. Additionally, 20% of the training set was reserved for validation to fine-tune the model during training.
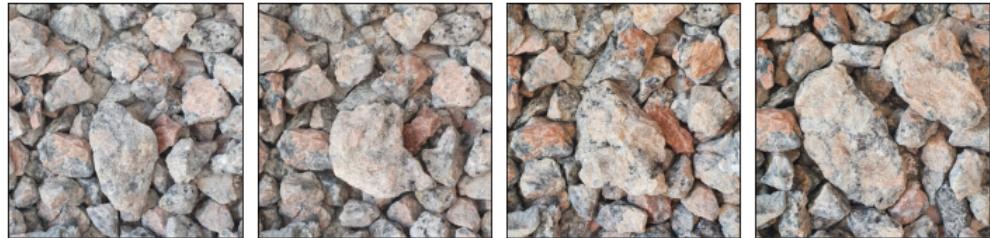
3.3. Image annotation process

We used the Robotflow platform[3] for image annotation. Each grain was meticulously delineated, resulting in masks that predominantly contained a single class (grains) with minimal separation from the background.

To enhance the visibility of grain edges, an erosion operation was applied to each annotation file. This preprocessing step was crucial because the accuracy of grain edge detection significantly impacts the study's outcomes. Consequently, the annotations were refined to emphasize these edges, as illustrated in Figure 6. This approach was adopted

---

[3]Robotflow. Available at: `https://roboflow.com/`. Acessed on: October 26, 2024

(a) Images containing only Gravel 1.

(b) Images contaminated with Gravel 2.

Figure 5. Dataset samples.

to improve the model's accuracy and effectiveness in segmenting and classifying rock grains.

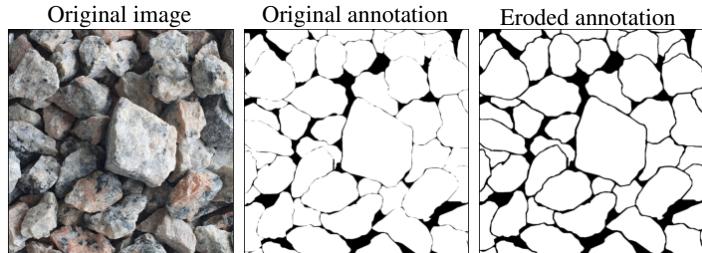Original image Original annotation Eroded annotation

Figure 6. Binary annotation showing refined grain edges.

By employing this detailed annotation process, we ensured that the model received high-quality training data. The precise delineation and enhanced edge visibility contributed to more accurate segmentation and classification of rock grains, ultimately improving the model's performance.

3.4. Images preprocessing

In the image preprocessing stage, both Median and Gaussian filtering were employed for denoising purposes. Median filtering, a nonlinear technique, is effective in preserving edge information while removing salt and pepper noise. On the other hand, Gaussian filtering, a smooth linear technique, is effective at removing noise with a Gaussian distribution. These preprocessing steps ensured cleaner images, which are crucial for the subsequent segmentation and classification tasks. The preprocessed images are shown in Figure 7.

3.5. Checking limitations of traditional segmentation methods

Several studies have successfully applied conventional image segmentation techniques across various domains. For example, the Canny edge detection algorithm [Canny 1986] has been widely used in medical imaging, remote sensing, and industrial inspection. Similarly, morphological operations have enhanced edge detection and noise removal in digital
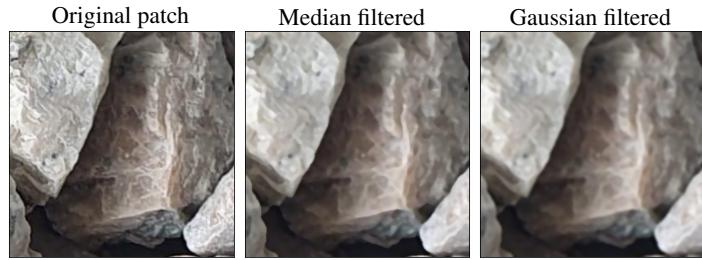
Figure 7. Example of preprocessed image with filtering techniques.

image processing [Haralick et al. 1987]. These methods have proven effective for images with well-defined edges and low noise levels.

In our study, we followed the steps below to test traditional segmentation method:

1. Grayscale conversion: converted the images to grayscale to simplify intensity analysis;
2. Binarization: applied binarization to categorize pixel intensities as foreground or background based on a threshold;
3. Median filtering: used median filtering to eliminate salt and pepper noise and improve edge detection accuracy;
4. Canny edge detection: employed the Canny edge detection algorithm to highlight significant intensity transitions, revealing edges and boundaries within the image;
5. Dilation and closing operations: performed dilation and closing operations to refine and connect detected edges, producing a final prediction;
6. Connected components analysis: found connected components to identify and label distinct objects in the image.

3.6. Evaluating U-Net architectures

In this study, we opted for the U-Net architecture as our deep learning-based approach. This model is robust to variations in image quality and can handle complex segmentation tasks involving diverse and noisy datasets. We evaluated and compared the performance of four U-Net variants: Standard U-Net (U-Net) [Ronneberger et al. 2015], Residual U-Net (RU-Net) [Alom et al. 2018], Attention U-Net (AU-Net) [Oktay et al. 2018], and Attention Residual U-Net (ARU-Net) [Chen et al. 2019]. Each model's performance was assessed using key metrics, including Mean Intersection over Union (MIoU), Intersection over Union for foreground (IoU-fg), Intersection over Union for background (IoU-bg), and accuracy (ACC).

3.7. U-Net architecture tuning

We created the U-Net model used in this study using the Keras package. This model is a variant of U-Net that incorporates attention mechanisms and residual connections to improve performance. Inspired by the work of [Wang et al. 2021a], the model accepts grayscale images of size 256x256 and comprises an encoder with four convolutional blocks. Each block has two convolutional layers followed by max-pooling and dropout layers, with attention gates applied at multiple levels to focus on relevant features. The decoder includes four transposed convolutional blocks, each combined with attention-gated features from the encoder, additional convolutional layers, and residual connections to enhance gradient flow and feature propagation. The output layer is a convolutional layer with a sigmoid activation function, generating a probability map for segmentation. The attention blocks help in emphasizing important regions, while residual connections address the vanishing gradient problem and preserve spatial information. The model is compiled with the Adam optimizer, binary cross-entropy loss, and accuracy metric, making it suitable for the binary segmentation tasks at hand.

3.8. Model training parameters

The model was configured to run for 40 epochs with a batch size of 30, incorporating early stopping and a custom learning rate schedule. Early stopping monitored the validation loss with a patience of 10 epochs, restoring the best weights if training was halted prematurely. The custom learning rate (lr) schedule function adjusted the learning rate over epochs using a power decay function, as described in Equation 2.

$$ lr = lr \times \left(1 - \frac{\text{iterations}}{\text{total\_iterations}} \right)^{0.9} \tag{2} $$

3.9. Refining model predictions: post-processing strategies

Each prediction was subjected to a series of filters applied to exclude non-representative grains, ensuring a more accurate grain size analysis.

1. Watershed technique: grains with slight contact were separated to ensure distinct grain boundaries.
2. Clear border technique: grains touching the image border were removed since their actual sizes could not be determined accurately.
3. Area thresholding: grains with an area smaller than 200 square millimeters were excluded, as they fall below the inferior size limit for Gravel 1.
4. Shape analysis: grains with a ratio of convex hull area to polygon area greater than 1.2 were removed because they represent grains with shapes that do not conform to the expected cubic structure.

After the post-processing steps presented above, the grain size is calculated as explained in Section 3.1. This calculation is then used to classify each segmented object as either Gravel 1 or Gravel 2, which in turn allows for the final classification of each image as either containing only Gravel 1 or being contaminated with Gravel 2.

4. Results

This section presents the evaluation of the final customized model's performance in distinguishing between images that contain only Gravel 1 and those contaminated with Gravel 2. The criteria for this differentiation is based on the size of the grains: images containing only Gravel 1 must not have grains exceeding 600 square millimeters, whereas images contaminated with Gravel 2 can include such grains.

4.1. Evaluation of traditional segmentation in gravel classification

As discussed in Section 3, traditional segmentation methods proved inadequate for our study due to their sensitivity to noise and the necessity for manual parameter tuning. This inadequacy is evident in the high level of noise present in the segmentation outputs, as shown in Figure 8. The figure highlights the limitations of traditional methods, reinforcing our decision to adopt a deep-learning approach, such as U-net, for more accurate and reliable segmentation results.

4.2. Comparative analysis of different U-Net architectures

ARU-Net demonstrated superior performance across all evaluated metrics. Specifically, it achieved the highest MIoU of 0.740, indicating the best overall segmentation accuracy. The IoU-fg of 0.884 and IoU-bg of 0.595 further highlight its balanced performance in segmenting both foreground and background elements effectively. Additionally, the overall ACC of 0.891 was slightly higher compared to U-Net and RU-Net, reinforcing its robustness and reliability. The results are summarized in Table 1 as a grayscale heat map, where darker cell colors indicate better scores.
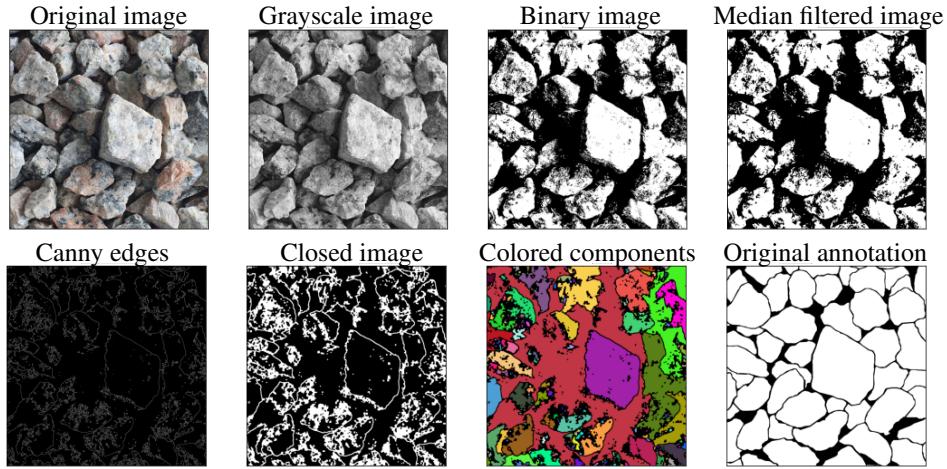
Figure 8. Output from conventional image segmentation techniques.

Table 1. Performance metrics of different U-Net models.

Model	MIoU	IoU-fg	IoU-bg	ACC
U-Net	0.729	0.881	0.587	0.890
RU-Net	0.713	0.866	0.561	0.886
AU-Net	0.722	0.873	0.570	0.891
ARU-Net	0.740	0.884	0.595	0.891

The superior performance of ARU-Net can be attributed to its ability to better capture and integrate multi-scale contextual information and refine feature maps, leading to more accurate and detailed segmentation. Therefore, this model was selected for further experiments and applications in this study due to its consistent and robust performance across all metrics.

4.3. Attention Residual U-Net Training Performance

Figure 9 shows the training and validation performance of the model over 20 epochs. The left plot demonstrates the steady decrease in loss for both the training and validation datasets, indicating effective learning and good generalization. The right plot shows increasing accuracy for both training and validation towards 0.9, reflecting the model's improved predictive performance. Overall, the model maintains stable performance on unseen data without significant overfitting.
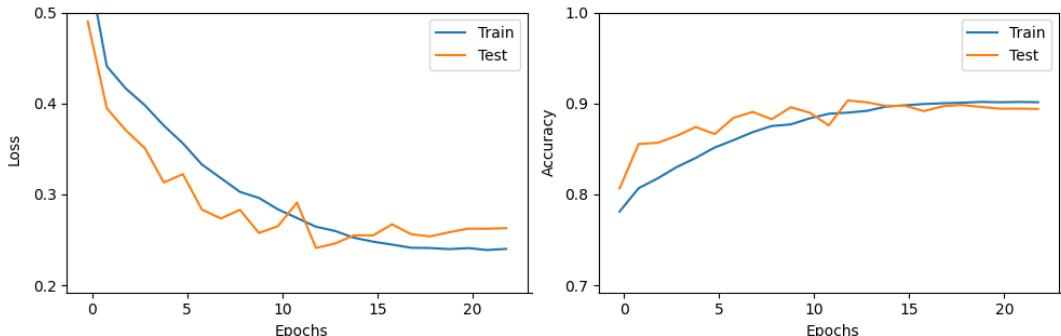

Figure 9. Loss and ACC curves for ARU-Net.

As explained in Section 3, each prediction from ARU-Net was subjected to a series of filters applied to exclude non-representative grains. Figure 10 illustrates the results obtained from the filtering process. This process is crucial for distinguishing the images containing only Gravel 1 from those contaminated with Gravel 2.
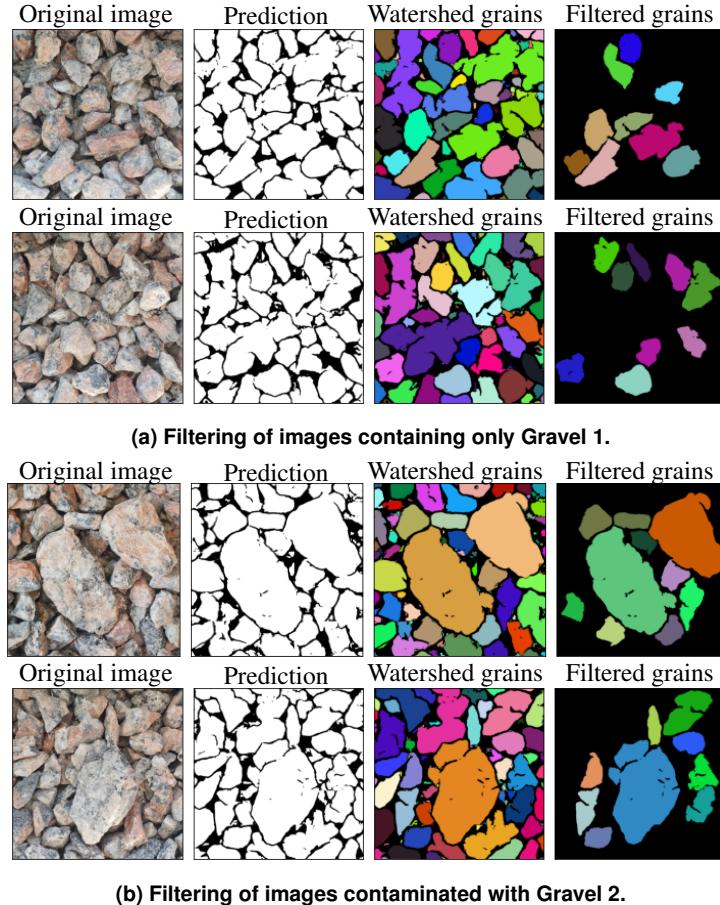

(a) Filtering of images containing only Gravel 1.

(b) Filtering of images contaminated with Gravel 2.

Figure 10. Grain selection for validation analysis.

After this process, we calculated the confusion matrix, as seen in Figure 11, which illustrates the model's effectiveness in distinguishing between images containing only Gravel 1 and those contaminated with Gravel 2. The model achieved an accuracy rate of 67.5%. More specifically, the confusion matrix reveals the following results:

- True Positives (TP): the model accurately identified 35% of the instances as contaminated with Gravel 2.
- True Negatives (TN): the model correctly recognized the absence of Gravel 2 in 32.5% of the instances .
- False Positives (FP): the model incorrectly identified 17.5% of the instances as contaminated with Gravel 2.
- False Negatives (FN): the model missed 15% of the instances that were actually contaminated with Gravel 2.

These results indicate that while the ARU-Net model demonstrates reasonable proficiency in distinguishing between Gravel 1 and Gravel 2, further refinement is necessary to enhance its accuracy and reduce misclassifications. By incorporating additional training data and applying more advanced techniques, we aim to achieve higher accuracy in future research.

	Predicted	
	Gravel 1 with Gravel 2	Gravel 1 without Gravel 2
Gravel 1 with Gravel 2	35%	15%
Gravel 1 without Gravel 2	17.5%	32.5%

(Actual)

True positive	
True negative	
False positive	
False negative	

Figure 11. Confusion matrix illustrating the performance of the ARU-Net model.

5. Conclusions

In this study, we developed a customized ARU-Net model to detect oversized contaminated aggregate rock grains, even with a limited dataset, showcasing its potential to disrupt and innovate within the aggregate industry. The model demonstrated strong generalization capabilities, accurately identifying 70% of positive contamination cases. While the current rate of false positives (FP) needs to be addressed, the false negatives (FN) rate of 30% is considered manageable given the frequent collection of images in practical applications. Future efforts will focus on eliminating false positives to enhance the model's reliability for deployment.

One aspect that needs improvement in our work is the collection of a larger amount of data. Although it is a costly task, we intend to invest in this area in future work to gather more information for both training and evaluating the model. Moreover, to further improve the model's accuracy, potential enhancements could involve incorporating more diverse training data, fine-tuning model parameters, and exploring additional preprocessing techniques to better handle edge cases. These enhancements aim to make the model more robust and effective, creating opportunities for its practical application in the aggregate industry. By continually refining and optimizing the model, it can become a valuable tool for real-time detection and quality control in aggregate production processes.

References

Alom, M. Z., Hasan, M., Yakopcic, C., Taha, T. M., and Asari, V. K. (2018). Recurrent residual convolutional neural network based on u-net (r2u-net) for medical image segmentation. *arXiv preprint arXiv:1802.06955*.

Associação Brasileira de Normas Técnicas (2005). *NBR 7211: Agregados para concreto - Especificação*. Rio de Janeiro. Acesso em: 27 abril 2024.

Bamford, T., Esmaeili, K., and Schoellig, A. P. (2021). A deep learning approach for rock fragmentation analysis. *International Journal of Rock Mechanics and Mining Sciences*, 145:104839.

Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6):679–698.

Chen, L., Zhou, X., Wang, M., Qiu, J., Cao, M., and Mao, K. (2019). ARU-Net: Research and application for wrist reference bone segmentation. *IEEE Access*, 7:166930–166938.

Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep Learning*. MIT Press. http://www.deeplearningbook.org.

Haralick, R. M., Sternberg, S. R., and Zhuang, X. (1987). Image analysis using mathematical morphology. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(4):532–550.

Hassan, S. M., Laban, N., Abo Khashaba, S. M., El-Shibiny, N. H., Bashir, B., Azer, M. K., Drüppel, K., and Keshk, H. M. (2024). Semantic segmentation of some rock-

forming mineral thin sections using deep learning algorithms: A case study from the nikeiba area, south eastern desert, egypt. *Remote Sensing*, 16(13).

LeCun, Y., Jackel, L., Bottou, L., Brunot, A., Cortes, C., Denker, J., Drucker, H., Guyon, I., Muller, U., Sackinger, E., et al. (1995). Comparison of learning algorithms for hand-written digit recognition. In *International conference on artificial neural networks*, volume 60, pages 53–60. Perth, Australia.

Leiva, C., Acuña, C., and Castillo, D. (2021). Development and validation of an online analyzer for particle size distribution in conveyor belts. *Minerals*, 11(6):581.

Maitre, J., Bouchard, K., and Bédard, L. P. (2019). Mineral grains recognition using computer vision and machine learning. *Computers & Geosciences*, 130:84–93.

MT Expo (2022). Demanda global por areia e agregados tende a subir 45% até 2060. Accessed: 2024-05-17.

Oktay, O., Schlemper, J., Le Folgoc, L., Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, N. Y., Kainz, B., et al. (2018). Attention u-net: Learning where to look for the pancreas. In *Medical Imaging with Deep Learning*, pages 63–72.

Přikryl, R. (2021). Geomaterials as construction aggregates: a state-of-the-art. *Bulletin of Engineering Geology and the Environment*, 80(12):8831–8845.

Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*, pages 234–241. Springer.

Shi, P., Duan, M., Yang, L., Feng, W., Ding, L., and Jiang, L. (2022). An improved u-net image segmentation method and its application for metallic grain size statistics. *Materials*, 15(13):4417.

Silva, D. d. A. and Geyer, A. L. B. (2018). Análise e classificação da forma do agregado graúdo britado para concreto. *Revista Científica Multidisciplinar Núcleo do Conhecimento ISSN*, 2448:0959.

Wang, F. and Zai, Y. (2023). Image segmentation and flow prediction of digital rock with u-net network. *Advances in Water Resources*, 172:104384.

Wang, Y., He, Z., Xie, P., Yang, C., Zhang, Y., Li, F., Chen, X., Lu, K., Li, T., Zhou, J., and Zuo, K. (2021a). Attention recurrent residual convolutional network for multi-organ segmentation in medical imaging. *Security and Communication Networks*, 2021:1–9.

Wang, Y., Ma, G., Mei, J., Zou, Y., Zhang, D., Zhou, W., and Cao, X. (2021b). Machine learning reveals the influences of grain morphology on grain crushing strength. *Acta Geotechnica*, 16(11):3617–3630.

Weir Motion Metrics (2024). Solução da empresa weir para análise de partículas em tempo real. `https://www.motionmetrics.com/belt-metrics/`. Acesso em: 27 abril 2024.

Yang, D., Wang, X., Zhang, H., yu Yin, Z., Su, D., and Xu, J. (2021). A mask r-cnn based particle identification for quantitative shape evaluation of granular materials. *Powder Technology*, 392:296–305.

Zhong, X., Deetman, S., Tukker, A., and Behrens, P. (2022). Increasing material efficiencies of buildings to address the global sand crisis. *Nature Sustainability*, 5(5):389–392.