E3S Web of Conferences **675**, 03008 (2025) https://doi.org/10.1051/e3sconf/202567503008
GeoME'5.5 2025

Semantic Segmentation of Flotation Froth Using a Hybrid Method: Combination of U-Net and Watershed

Widad Saraji[1*], *Youssef Zerradi*[1], *Nouha El Malki*[1], *Latifa Ouadif*

[1]Laboratory of Applied Geophysics, Geotechnics, Engineering Geology, and Environmental (L3GIE), Mohammadia School of Engineers, Mohammed V University in Rabat, Morocco.

Abstract. Physical characteristics, particularly those related to bubbles, being a widely extracted feature from flotation froth images, whose main purpose is to evaluate the efficiency of the process. The introduction of computer vision into flotation processes has seen significant progress, ranging from segmentation by the watershed algorithm to more advanced techniques primarily based on artificial intelligence. The present study will show the impact of the application of AI in the system for improving the observation of flotation froth images in order to overcome the typical limitations of traditional methods. In this sense, a method that combines a deep learning model (U-Net) for semantic segmentation with watershed post-processing was adopted. According to the results obtained, this methodology is effective in the semantic segmentation of flotation froth images. The use of U-Net initially improved the accuracy of this segmentation and subsequently overcame the typical limitations of watershed. The metric calculations (IOU = 70%, Dice coefficient = 80%) also indicate the model's effectiveness. Moreover, the visual results demonstrated the model's ability to accurately recognize the boundaries of the bubbles. In conclusion, the introduction of artificial intelligence tools in mining processes represents a promising step toward a more modern mining industry.

1 Introduction

Froth flotation, the mineral separation method widely known for its extensive application in mineral processing, is based on exploiting the differences in the physicochemical properties of mineral surfaces [1]. According to these surface properties, the particles can be divided into two categories: hydrophobic and hydrophilic. The hydrophobicity of particles can be natural (examples: coal and molybdenite) [2], as it can be acquired following the action of collectors that act on their surface, making it hydrophobic, and consequently, facilitating attachment with air bubbles. As for the frothers, they play the role of a froth stabilizer, which guarantees a better recovery of valuable minerals. The efficiency of the flotation process is primarily indicated by the flotation froth [3]. Thus, its observation was primarily carried out by the operators. This latter faces many limitations, primarily its subjective character, which

* Corresponding author: sarajiwidad@gmail.com

affects the precision of the observation. To overcome these limitations, computer vision was introduced as an alternative to human vision, in order to obtain accurate information about the froth.

The main characteristics that can be extracted from flotation froth images are of three types: physical, statistical, and dynamic, with the physical ones being primarily related to the bubbles (size distribution and shape) [4].

In this sense, several segmentation algorithms have been developed, of which the main ones are summarized in Table 1.

Table 1. The main segmentation algorithms.

Algorithms used	References
Watershed	[5–8]
Valley edge detection	[9–11]
Texture spectrum analysis	[13, 14]

Segmentation based on the Watershed algorithm is an effective solution widely used to measure the size of bubbles in froth images [6, 7, 9].

However, despite its key strengths in terms of speed and power in image segmentation, this algorithm presents some limitations, notably over-segmentation in the case of large bubbles and under-segmentation for small ones [3].

Due to these limitations, as well as the complexity of froth images caused by the high number of bubbles, and their poorly distinguished contours, research has oriented, in the last decades, to potential and automatic solutions that integrate the artificial intelligence (AI) techniques, and in particular deep learning, because of their influence in different fields, extending from daily life to advanced scientific applications. Among the applications of deep learning in computer vision is semantic segmentation, which the image is dividing into segments and each pixel of it is assigned to a specific class or object [14], which ensures a more detailed and accurate comprehension of its content. Several deep learning models have been developed to carry out this task, among which the most popular are: FCN, SegNet, and U-Net.

A hybrid method combining the U-Net model and the Watershed algorithm was chosen for the semantic segmentation of flotation froth images.

Since its development, the U-Net has been primarily used for biomedical image analysis. Due to its promising results, its application has expanded to other fields, notably in flotation image segmentation. The aim of this study is to combine the classical image processing approach (Watershed) with an artificial intelligence tool (U-Net) to achieve accurate bubble segmentation while overcoming the typical limitations of Watershed.

2 Methodology

In this section, we describe the methodology illustrated in Fig.1 for the semantic segmentation of flotation froth images, based on segmentation by U-Net, followed by post-processing using Watershed.

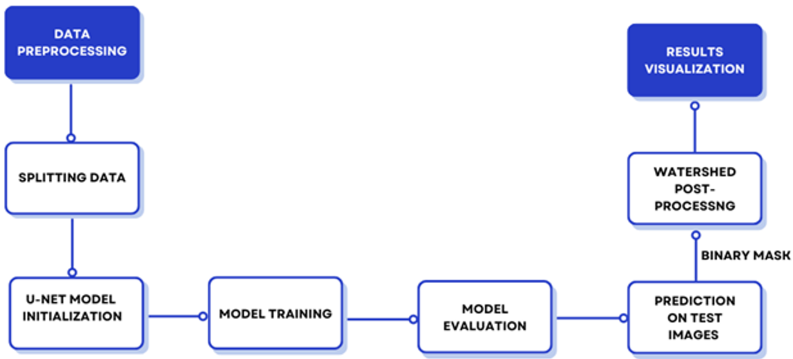

Fig. 1. Scheme of the proposed semantic segmentation methodology combining U-Net and Watershed for flotation froth bubble analysis.

The first step involves data preprocessing; the data, consisting of raw images and their corresponding annotated masks (Fig.2), are loaded and then resized uniformly to a standard size (256 × 256 pixels) to adapt them to the model's input. Followed by a normalization of pixel values with the aim of optimizing the application of the model.

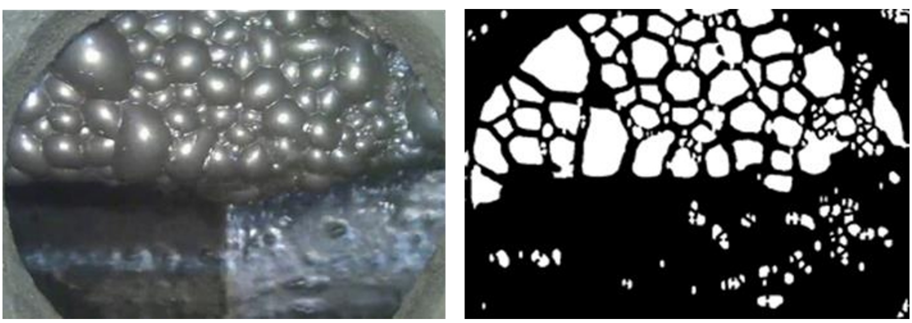

Fig. 2. Image from the training dataset: on the left, the raw image; on the right, its corresponding annotation mask.

The second step aims to divide the data into two sets: a training set (90%), as its name suggests, is used for training the model, and a test set (10%) to evaluate its performance.

Subsequently, the model initialization step is carried out with a classic architecture suitable for binary segmentation (Fig.3). This step starts with introducing an input image (input layer) in grayscale of size 256×256×1, followed by the encoder block with four convolutional levels. At each level, three steps apply: Firstly, a 2D convolution, followed by dropout in order to prevent overfitting. Thirdly, a max pooling, which gradually reduces the spatial dimensions while increasing the number of channels (32, 64, 128).

The central block, or the bottleneck of the U structure, contains two 2D convolutions using 256 filters. In the decoder part, four levels symmetrical to those of the encoder are also found. Each level begins with a transposed convolutional layer, followed by concatenation with the corresponding output from the encoder (skip connection), a characteristic of the U-Net architecture. This concatenation is followed by two 2D convolutions, interleaved with dropout.

Thus, at each decoding phase this time the size doubles (from 16×16 to 256×256), while the number of channels decreases (from 128 to 1), in order to obtain the final output image of size 256×256×1.

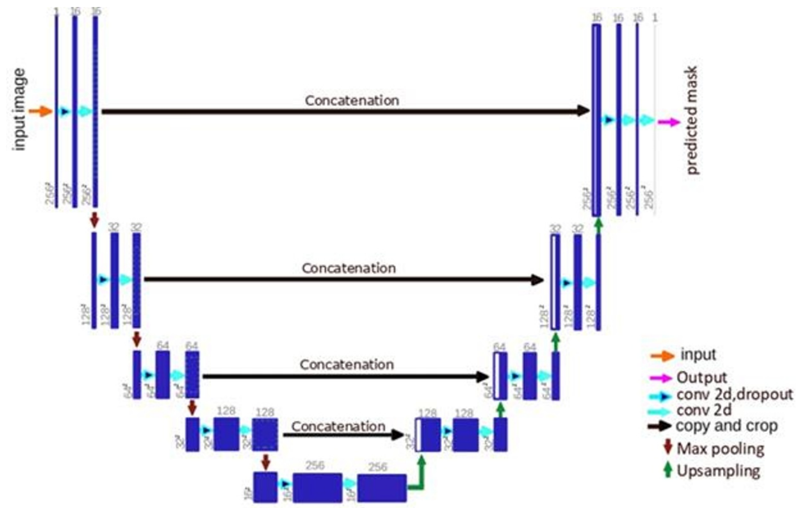

Fig. 3. U-Net Architecture used.

Once the model is defined, the training process is started, and the model learns from the training data using 50 epochs and a batch size of 32. After that, the model passes through a validation step on the test set.

Then, a model performance evaluation step is carried out based on the loss and accuracy curves, as well as on the calculation of metrics such as Intersection over Union (IoU), Dice coefficient, and F1 score, which are necessary in the case of binary segmentation to judge the model's capacity and performance to distinguish between positive and negative classes. Predictions on new internal images (issues from the same dataset) and external ones (from other samples) are then generated in the form of a binary mask (0 = background,1= bubble).

As a starting point for watershed post-processing, the predicted mask obtained by U-Net is used. This processing is structured in two steps: one to extract the large bubbles starting with a cleaning by eliminating bubbles that are below 100 pixels while keeping the others, followed by four more steps: applying the distance transformation, Gaussian smoothing, local minima suppression, and finally watershed segmentation based on the bubble centers. The second step concerns the small bubbles, which will be extracted from the residual mask by following the same steps as for the large bubbles. Finally, the results of the two steps are then merged to create the outlines of the bubbles, which will then be superimposed on the original image.

3 Results

The results obtained following the application of the proposed methodology will be presented in the following section. The masks predicted by the trained model are evaluated starting with an analysis of the metric calculation results, followed by a visual comparison with the result of applying watershed alone to demonstrate the robustness of the segmentation.

 https://doi.org/10.1051/e3sconf/202567503008

3.1 Semantic segmentation by the trained U-Net model

3.1.1 Model evaluation

To evaluate the robustness of the model for semantic segmentation of flotation froth images, we utilized a set of metrics specifically designed for this task. The results of accuracy and loss function during training, as well as the segmentation metrics on the test set, were satisfactory. We present the evaluation measures below.

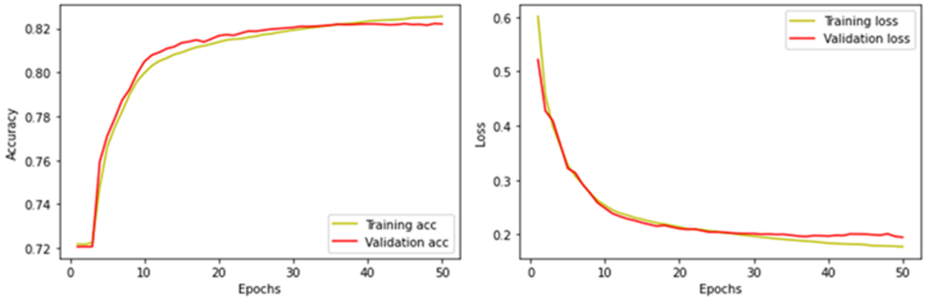

Fig. 4. Training and validation curves: accuracy (left) and loss (right).

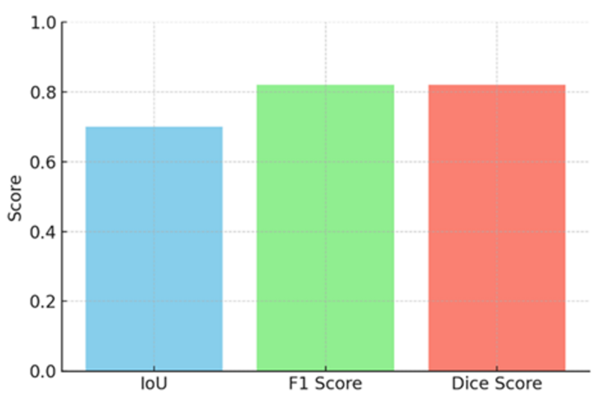

Fig. 5. Segmentation metrics on test set.

According to Fig. 4, representative of the precision and loss curves, a trend toward stability after a certain number of epochs is observable, specifically at epoch 25 and epoch 20 respectively, which indicates the convergence of the model. The accuracy values during training, as well as during validation, exceed 82%, indicating the model's capacity to learn from the training data and extrapolate its predictions to new data.

The Intersection over Union (IoU), the F1 score, and the Dice coefficient (Fig.5), presents the respective values of 70%,82% and 82%, which indicates that the predicted masks cover a large portion of the actual regions, and that a balance exists between precision and recall.

3.1.2 Predicted mask visualization

In order to show the model's performance in accurately segmenting bubbles from flotation froth images, this section presents a visual comparison between the original images, their corresponding true masks, and the predicted masks, as shown in figure below.

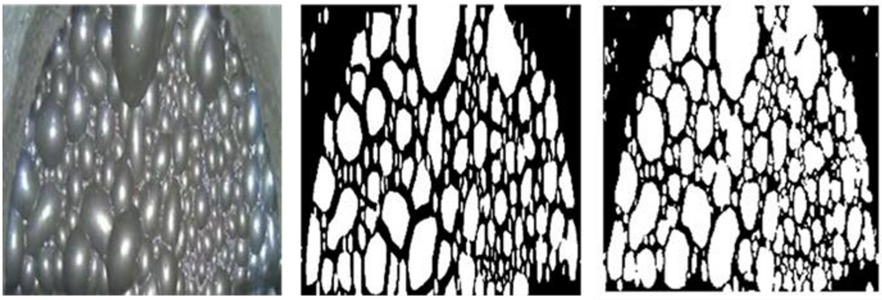

Fig. 6. Segmentation results of flotation froth bubbles using U-Net: the original image (left), true mask (middle), and model prediction (right).

The model's ability to accurately identify the majority of the bubbles was demonstrated through a significant overlap between the segmented regions and the actual mask, observed following a visual comparison between the predicted mask resulting from the application of U-Net and the annotated mask. However, some notable divergences were observed, particularly for certain adjacent small bubbles.

3.2 Post-processing by Watershed

The results of the visual comparison showed the model's ability to effectively learn the morphological features of the bubbles, while clearly indicating the necessity for a post-processing step to enhancing the identification and separation of adjacent small bubbles. Thus, the results of applying a Watershed post-processing step are presented below.

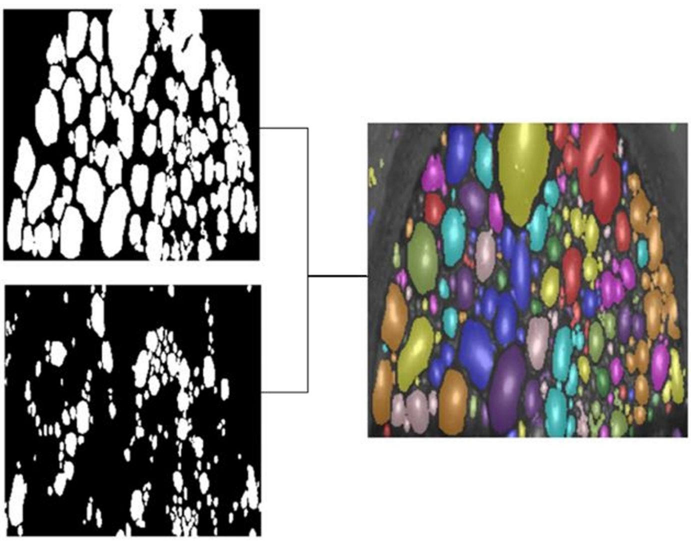

Fig. 7. Two-step segmentation of flotation bubbles from the predicted mask: first, large bubbles (at the top), then small bubbles from the residual mask (below), generating a global mask (right).

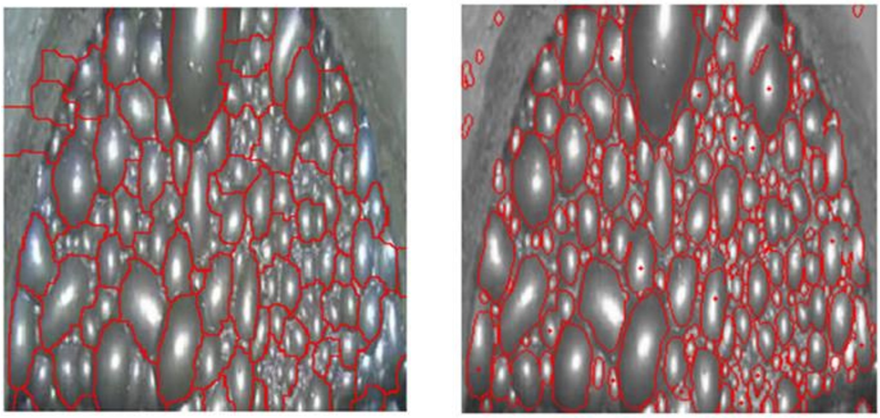

Fig. 8. Segmentation results of Watershed only (left) versus Watershed segmentation based on U-Net mask (right).

By structuring the process into two stages, the method first allowed for the efficient extraction of large bubbles by eliminating the small ones, and then enabled the segmentation of small bubbles from the residual mask (Fig.7).

After that, both results were combined into a global mask, from which the extracted bubble contours are overlapping on the original image, showing excellent match between the segmented and real bubbles, thus attesting to Watershed's ability to refine the initial prediction generated by U-Net, exceeding the over-segmentation and under-segmentation problems observed when using Watershed alone (Fig.8).

4 Discussion

The hybrid methodology applied in this article has enabled semantic segmentation of flotation froth images, with the primary objective of developing an effective segmentation method to mitigate the limitations of conventional methods, particularly the watershed algorithm, by integrating artificial intelligence (AI) tools.

The obtained results indicate the performance of the combined method in segmenting bubbles in flotation froth images. The model allowed for precise identification of the appearance of the bubbles, as demonstrated by the evaluation results, notably the IoU score reaching 70% and the Dice coefficient of 82%. Moreover, the visual analysis of the mask predicted by the model demonstrates its ability to effectively learn the overall morphological characteristics of the bubbles. However, in the case of overlapping and small bubbles, discrepancies between the predicted mask and the actual mask have been observed, leading to a decrease in segmentation accuracy, which is well documented in the literature.

Recently [15] have highlighted the challenge of segmenting images containing numerous small objects such as flotation bubbles. This is due to their similar characteristics in terms of texture, shape, and color. Which involved a watershed post-processing step comprising two sub-steps: one dedicated to large bubbles, using as a starting point the predicted mask resulting from the application of U-Net, and the second to small ones, starting from the residual mask. And consequently, overcome the problems of under-segmentation and over-segmentation.

5 Conclusions

In this article, a combined methodology for semantic segmentation of flotation froth images was adopted, with the aim of overcoming the typical problems associated with the watershed segmentation method. The combination of the U-Net model with the classical Watershed segmentation method to identify and delineate the bubbles significantly improved the accuracy of bubble delineation, including small and overlapping bubbles. This is demonstrated both by the results of the metric calculations and by the visual comparison of the predicted mask and the ground truth mask. To conclude, the study highlights both the crucial role of AI in mining applications, while also paving the way for the exploitation of these results in future work, particularly regarding bubble classification based on morphology, whose importance lies in ensuring understanding and control during the flotation process, mainly in reagent dosage.

This work was supported by the National Center for Scientific and Technical Research (CNRST). The dataset (flotation froth images) used in this study is publicly available at https://www.kaggle.com/datasets/obobojk/froth-bubbles (Kaggle).

References

1. S. A. Khoso, Z. Gao, M. Tian, Y. Hu, & W. Sun, The synergistic depression phenomenon of an organic and inorganic reagent on FeS2 in Cu[sbnd]S flotation scheme. Journal of Molecular Liquids, **299** (2020). https://doi.org/10.1016/j.molliq.2019.112198.
2. Barry. Wills & James. Finch, Wills' Mineral Processing Technology: An Introduction to the Practical Aspects of Ore Treatment and Mineral Recovery, Eighth Edition (Elsevier Science and Technology Books, Inc., 2016).
3. A. Jahedsaravani, M. Massinaei, & M. Zarie, Measurement of bubble size and froth velocity using convolutional neural networks. Minerals Engineering, **204** (2023). https://doi.org/10.1016/j.mineng.2023.108400.
4. C. Aldrich, C. Marais, B. J. Shean, & J. J. Cilliers, Online monitoring and control of froth flotation systems with machine vision: A review. International Journal of Mineral Processing, **96** (2010) 1–13. https://doi.org/10.1016/j.minpro.2010.04.005.
5. A. Jahedsaravani, M. Massinaei, & M. H. Marhaban, An image segmentation algorithm for measurement of flotation froth bubble size distributions. Measurement: Journal of the International Measurement Confederation, **111** (2017) 29–37. https://doi.org/10.1016/j.measurement.2017.07.023.
6. A. Jahedsaravani, M. H. Marhaban, M. Massinaei, M. I. Saripan, & N. Mehrshad, Development of a new algorithm for segmentation of flotation froth images (2014).
7. M. L. Tian, J. M. Yang, & M. Q. Tian, Improved extracting algorithm of internal markers for froth image watershed segmentation of coal flotation. International Journal of Simulation: Systems, Science and Technology, **17** (2016). https://doi.org/10.5013/IJSSST.a.17.26.16.
8. H. Zhang, Z. Tang, Y. Xie, X. Gao, & Q. Chen, A watershed segmentation algorithm based on an optimal marker for bubble size measurement. Measurement: Journal of the International Measurement Confederation, **138** (2019) 182–193. https://doi.org/10.1016/j.measurement.2019.02.005.
9. A. W. Banford & Z. Aktas, The effect of reagent addition strategy on the performance of coal flotation. Minerals Engineering, **17** (2004) 745–760. https://doi.org/10.1016/j.mineng.2004.01.005.
10. G. Forbes & G. De Jager, Bubble size distributions for froth classification (2005).
11. W. Wang, F. Bergholm, & B. Yang, Froth delineation based on image classification. Minerals Engineering, **16** (2003) 1183–1192. https://doi.org/10.1016/j.mineng.2003.07.014.

12. B. Lin, B. Recke, J. K. H. Knudsen, & S. B. Jørgensen, Bubble size estimation for flotation processes. Minerals Engineering, **21** (2008) 539–548. https://doi.org/10.1016/j.mineng.2007.11.004.
13. P. N. Holtham & K. K. Nguyen, On-line analysis of froth surface in coal and mineral flotation using JKFrothCam (2002).
14. T. Boyadzhiev & K. Ivanova, Unification of Semantic and Instance Segmentation with BoundaryX. Cybernetics and Information Technologies, **25** (2025) 54–67. https://doi.org/10.2478/cait-2025-0022.
15. E. Prokopov, D. Usacheva, M. Rumiantceva, & V. Efimova, Weak Segmentation and Unsupervised Evaluation: Application to Froth Flotation Images. Proceedings of the International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (Science and Technology Publications, Lda, 2025), pp. 500–507. https://doi.org/10.5220/0013181100003912.