

# A Machine Vision Approach for Bioreactor Foam Sensing

SLAS Technology  
2021, Vol. 26(4) 408–414  
© The Author(s) 2021



DOI: 10.1177/24726303211008861  
journals.sagepub.com/home/jla



Jonas Austerjost<sup>1\*</sup>, Robert Söldner<sup>1\*</sup>, Christoffer Edlund<sup>2</sup>,  
Johan Trygg<sup>2</sup>, David Pollard<sup>3</sup>, and Rickard Sjögren<sup>2</sup>

## Abstract

Machine vision is a powerful technology that has become increasingly popular and accurate during the last decade due to rapid advances in the field of machine learning. The majority of machine vision applications are currently found in consumer electronics, automotive applications, and quality control, yet the potential for bioprocessing applications is tremendous. For instance, detecting and controlling foam emergence is important for all upstream bioprocesses, but the lack of robust foam sensing often leads to batch failures from foam-outs or overaddition of antifoam agents. Here, we report a new low-cost, flexible, and reliable foam sensor concept for bioreactor applications. The concept applies convolutional neural networks (CNNs), a state-of-the-art machine learning system for image processing. The implemented method shows high accuracy for both binary foam detection (foam/no foam) and fine-grained classification of foam levels.

## Keywords

foam sensor, machine vision, deep learning, process analytical technology, bioprocessing

## Introduction

Foam emergence is a commonly observed phenomenon in bioprocess upstream applications.<sup>1</sup> Foam, which is mainly provoked by the combination of gassing needed to support the cell culture and the release of lipids and proteins from cells, has an adverse effect on the operation of the bioreactor and the cell culture productivity.<sup>2</sup> Loss of cell viability leading to cell rupture and foaming can then occur due to lack of specific nutrients or mechanical stresses from bursting gas bubbles and agitation. The generated foam can then rapidly develop, in some cases during the course of minutes, and block exhaust gas filters, resulting in reactor overpressure, reduced sterility integrity, and ultimately batch failure, stressing the need to detect and prevent foam emergence.<sup>3</sup>

Established strategies to prevent and eliminate foaming within a bioprocess rely predominantly on chemical methods, but mechanical and physical methods are used as well. Mechanical and physical methods are only able to destroy existing foam, whereas chemical methods are capable of preventing the emergence of foam as well as eliminating present foam.<sup>4</sup> Typical mechanical strategies to break foam include liquid sprayers, centrifugal foam breakers, or orifice foam breakers, while physical strategies include the application of ultrasonic or thermal probes to break existing foam.<sup>4–6</sup> Chemical strategies rely on the addition of so-called anti-foam agents to the cell culture broth. These antifoam agents

are surface active substances, which influence the surface properties of the medium toward a decreased foaming ability, including commonly used agents such as silicone oils, polypropylene glycol, and glycerol esters.<sup>7</sup> Due to being relatively inexpensive and easy to handle and add into bioprocessing equipment, chemical foam prevention and elimination strategies are most often used. Nevertheless, antifoam agents influence mass transport and high concentrations may negatively affect the volumetric mass oxygen transfer coefficient as well as the dissolved oxygen concentration, which are important parameters for aerobic cell culture processes.<sup>8</sup> This can result in decreased cell growth and reduced product titers.<sup>9,10</sup> Furthermore, antifoam agents may

<sup>1</sup>Sartorius Corporate Research, Sartorius Stedim Biotech GmbH, Göttingen, Germany

<sup>2</sup>Sartorius Corporate Research, Sartorius Stedim Data Analytics AB, Umea, Sweden

<sup>3</sup>Sartorius Corporate Research, Sartorius Stedim North America Inc., Boston, USA

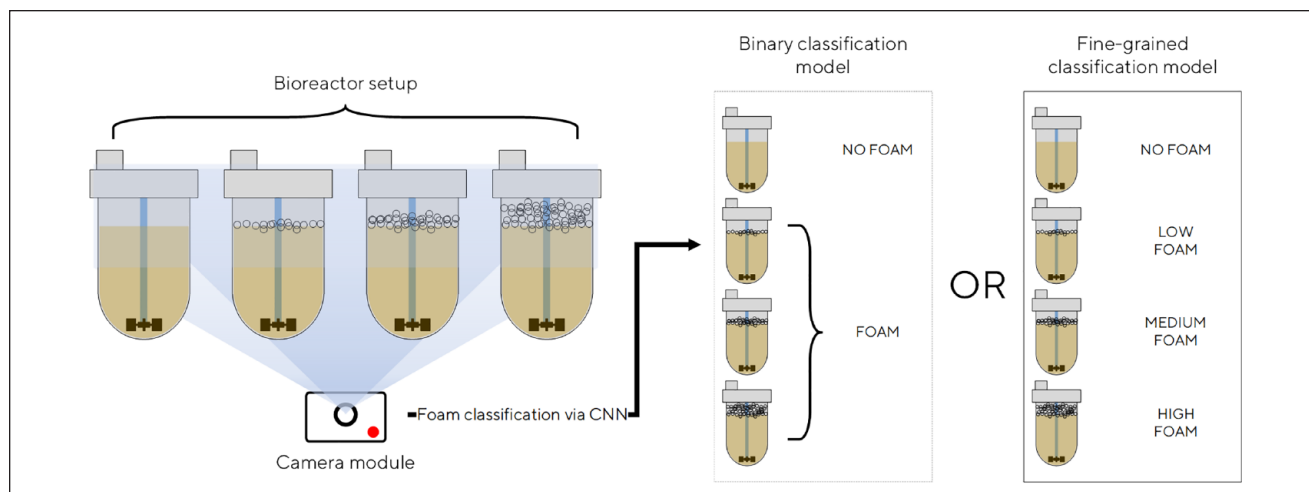
\*Both authors contributed equally to this work.

Received Dec 4, 2020, and in revised form March 11, 2021. Accepted for publication March 18, 2021.

Supplemental material is available online with this article.

## Corresponding Author:

Robert Söldner, Sartorius Corporate Research, Sartorius Stedim Biotech GmbH, August-Spindler-Str. 11, Göttingen, 37079, Germany.  
Email: robert.soeldner@sartorius.com



**Figure 1.** Schematic diagram depicting the implemented machine vision-based foam detection in a single-use bioreactor setup. First, an image is acquired by a camera module, which is then classified by a CNN. The classification can be performed by either the implemented binary classification model or the developed fine-grained classification model.

cause fouling of filters and membranes in subsequent downstream processing applications and therefore accelerate material fatigue and decrease purification efficiency.<sup>11,12</sup> This underlines the need for a well-considered antifoam agent feeding strategy based on reliable sensor data. Established foam sensors in bioprocessing include conductivity or capacitance probes placed within the bioreactor. The disadvantages of these contact-based sensing strategies are the fouling and coating of probes, which typically results in false-positive signals.<sup>4</sup> The outcome of this is high maintenance costs and possible overaddition of chemical antifoams, which can ultimately lead to batch failure. Contactless foam sensing can be reached via ultrasound sensors, but is prone to temperature shifts, humidity, and false-positive signals from splashing caused by agitation.<sup>13</sup>

As foam is a key bioprocess parameter that can be visually identified, machine vision-based approaches are promising candidates to detect foam emergence. Traditional machine vision workflows rely on extensive feature engineering, where complex algorithms are hand-engineered to achieve the task at hand.<sup>14,15</sup> Achieving good predictive performance from such a system is difficult, and they suffer from low robustness to changes in imaging conditions. However, in the past decade deep convolutional neural networks (CNNs) as well as openly available large-scale annotated data sets have contributed to exceptional progress in machine vision.<sup>16</sup> By training CNNs end to end on large data sets, CNN-based machine vision has outperformed traditional methods for a wide variety of vision tasks and now completely dominates the field.<sup>17</sup>

In this study, we present a machine vision-based strategy to detect foam within a small-scale (250 mL), single-use bioreactor setup (Fig. 1). This concept was implemented

using off-the-shelf hardware components and open-source machine learning software libraries. The established system showed high accuracy in both binary foam detection and fine-grained classification and has been identified as a promising approach to overcome the drawbacks of conventional foam sensor systems like fouling and coating, as well as limited single-level functionalities. The noninvasive system shows proof of concept for the application to bioreactor formats for both single-use and stainless steel formats.

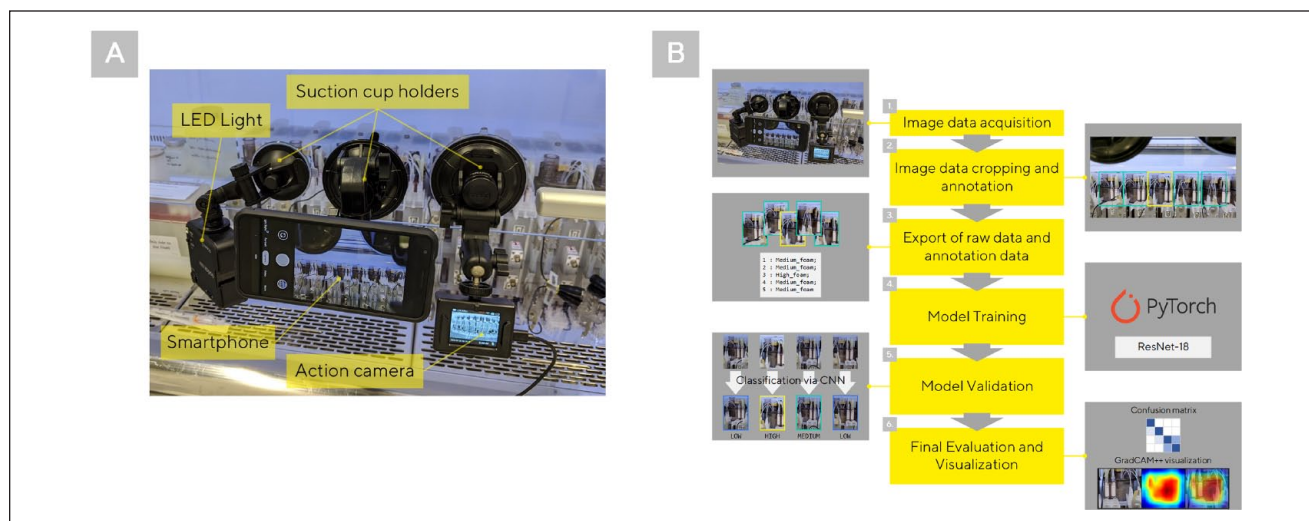
## Materials and Methods

### Experimental Setup

All bioreactor experiments were performed using an Ambr250 high-throughput multiparallel bioreactor system that has become the biotech/biopharma industry state of the art for bioprocessing research and development (The Automation Partnership [Cambridge] Ltd., Cambridge, UK, part of the Sartorius Stedim Biotech Group).<sup>18</sup> The system comprises 12 or 24 disposable vessels (250 mL each) integrated into a liquid handling system for fully automated bioprocessing operation. The system is housed inside a biosafety cabinet to enable aseptic automated sample removal and collection. Two different camera modules have been placed in front of the Ambr250 system and were used to acquire the image material for the model development. A smartphone (Google Pixel 3a XL, Google LLC, Menlo Park, CA), as well as an action camera (apeman A79, Apeman International Co., Ltd., Shenzhen, China), were used for image data acquisition. An additional light-emitting diode (LED) light source (Godox LED64 LED, GODOX Photo Equipment Co., Ltd., Shenzhen, China) was used to

**Table 1.** Experimental Plan, Which Resulted from a Full-Factorial Design DoE with 2 Levels for Each Factor (Volume, Dye Addition, Clean Bench Light).

Experiment No.	1	2	3	4	5	6	7	8
Run order	2	6	3	7	4	1	5	8
Volume (mL)	200	240	200	240	200	240	200	240
Dye addition	No	No	Yes	Yes	No	No	Yes	Yes
Clean bench light of biosafety cabinet	Off	Off	Off	Off	On	On	On	On

**Figure 2.** (A) Experimental setup used for image acquisition. A smartphone, an action camera, and an LED light source have been mounted on a clean bench glass in front of a multiparallel small-scale bioreactor system via suction cup holders. (B) Performed six-step workflow to acquire a CNN, which is able to distinguish between different levels of foam in single-use, small-scale, bioreactors. The shown workflow is an example for the fine-grained classification model.

introduce lighting varieties to the image data set during image data acquisition, in addition to the standard Ambr250 clean bench lighting modifications (clean bench light of biosafety cabinet on/off) (see **Table 1** for performed experiments). Each device was fixated on the clean bench window using a dedicated suction cup holder (**Fig. 2A**).

To include and identify important process and environmental parameters with respect to model quality, a design of experiments (DoE) using the Software MODDE (Sartorius Stedim Data Analytics AB, Umeå, Sweden) was performed. The experimental plan is shown in **Table 1** and was generated by using a full-factorial design (FFD) with two levels for each factor. The “Volume (mL)” entry within the experimental plan corresponds to the filling volume of the cultivation vessel (200 mL/240 mL). The “Dye addition” entry indicates if 50  $\mu$ L of food dye (Orange Red, Suchuangyi Technology Co., Ltd., Shenzhen, China) were added to the media or not (yes/no). The “Clean bench light” entry specifies if the clean bench light (which is part of the safety cabinet in which the Ambr250 system is placed) was turned on or off (on/off). To prevent any experimenter bias, the order of execution was assigned at random (“Run order” row). In addition, the external light source (“LED light”) (**Fig. 2A**)

was arbitrarily turned on and off to further introduce diversity into the acquired image material. To provoke variously strong characteristics of foam levels, different levels of air supply (5–50 mL/min) and different additions (100  $\mu$ L to 1 mL) of a 0.5 g/mL bovine serum albumin (BSA) solution (BSA acquired from Sigma-Aldrich Chemie GmbH, Taufkirchen, Germany) have been added to the used medium (4Cell XtraCHO Stock & Adaptation, Sartorius Stedim Cellca GmbH, Ulm, Germany). Furthermore, stirrer speed adjustments were performed during the image acquisition phase (500–1500 rpm). Video material was initially recorded, and the video frames were subsequently extracted. A video of every experiment was recorded for 20 to 25 min to collect a diverse data set of different foam quantities. The resolution of the acquired images was  $1920 \times 1080$  pixels for the smartphone camera and  $1520 \times 2688$  pixels for the action camera, respectively.

### Model Training

After image material acquisition, regions of interest (ROIs) were manually annotated using the cloud-based image annotation platform Dataloop (Dataloop AI, Herzliya,

**Table 2.** Acquired Image Data Set and Manually Annotated Classes.

Data Set	Whole Data Set	Action Camera	Smartphone
No foam	982	17	965
Low foam	2183	142	2041
Medium foam	1542	124	1418
High foam	428	61	367
Total	5135	344	4791

Israel). The ROIs that contain the bioreactor vessel from a side view were cropped out, rescaled to  $250 \times 250$  pixels, and assigned to a class (no foam, low foam, medium foam, high foam) by a single subject matter expert trained in bio-processing scenarios to reduce the risk of introducing inconsistent labels (see Supplemental Material for example images and their assigned classes). The resulting data set, which formed the foundation for model generation and validation, is specified in **Table 2**.

The generated data shown in **Table 2** comprise the annotated image material generated by the action camera, which took an image every 10 s, and the annotated image material originating from the smartphone camera video, with an image every 90th frame (corresponding to one image every 3 s, downsampled from original acquisition at 30 frames per second). For experiments on binary classification, which distinguishes between no foam and foam, the classes low foam, medium foam, and high foam were combined into the single class foam, whereas as all classes were used for fine-grained classification.

The annotation data and the cropped raw image data were exported and used to train CNN models for image classification, using the Python programming language (version 3.6) and the deep learning framework PyTorch (version 1.4) (Facebook Research, Menlo Park, CA). For both binary foam detection and fine-grained classification, a ResNet-18 model neural network was used.<sup>19</sup> ResNet variants are widely used for image classification, and ResNet-18 is the smallest model in this family.

Both the binary and fine-grained models were trained with cross-entropy loss for 30 epochs, a batch size of 50 images (largest batch size fitting on the graphics processing unit [GPU] used for training), a learning rate of  $1.2e^{-5}$  (chosen after pilot experiments evaluating the learning rate influence on model convergence on the validation set), the Adam optimizer,<sup>20</sup> and random horizontal flips for data augmentation. For each training, the model with lowest validation loss was saved and used for evaluation (see Supplemental Material for corresponding loss plots). The validation data were created by taking 10% of the training data at random and using that split for all models doing the same task (binary or fine-grained classification). Due to class imbalance, because of more foam than no-foam

images present, class weights of 0.4, corresponding to the ratio of foam and no-foam images, were introduced for the images having foam in them when training the binary classifier but not the fine-grained one.

### Model Validation

To validate the models, a subset of the images was excluded from model training and only used to evaluate models' classification performance as a test set. For the binary classifier, all images from one smartphone video capture were excluded to constitute a test set of 672 images, where 477 images contained foam and 195 did not. Due to the low number of high-foam images in any given video capture, another test set with 512 images was designed and used for the fine-grained model containing 98 no-foam, 218 low-foam, 154 medium-foam, and 386 high-foam images. Since the classes are imbalanced, evaluation accuracy as commonly defined, that is, the ratio of correct classifications, will be biased toward the class with the most labels. Instead, the classification performance was evaluated by calculating the F1 score, defined as

$$Precision = \frac{TP}{(TP + FP)} \quad (1)$$

$$Recall = \frac{TP}{(TP + FN)} \quad (2)$$

$$F1 = \frac{(Precision * Recall)}{(Precision + Recall)} \quad (3)$$

Here, the TP, TN, FN, and FP values depict the true-positive/true-negative and false-negative/false-positive predictions, respectively. "Precision" indicates, out of how many images that were classified as containing foam, what ratio was correctly predicted. Recall indicates the ratio of all images containing foam that were correctly classified as containing foam. The F1 score is the harmonic mean of the precision, and recall and is widely used to provide a single evaluation metric for classification models when classes are imbalanced.

To qualitatively validate the models' predictions, the locations where the models attend for predictions were visualized using the GradCAM++ method.<sup>21,22</sup> GradCAM++ uses a weighted combination of the positive partial derivatives of the last CNN layer to produce a heat map over the image highlighting regions the model pays much attention to when making its prediction. Although GradCAM++ does not provide a full explanation behind the prediction, the heat maps provide intuition whether the model predictions are based on sensible information.



**Table 3.** CNN Classification Performance on Foam Detection.

Model	Precision (%)	Recall (%)	F1 Score (%)
Binary	97.95	96.95	97.45
Fine-grained	76.35	78.68	75.58

## Results and Discussion

Two foam classification models were generated following a six-step workflow (**Fig. 2B**). One of the resulting models is a binary model, which is able to distinguish between foam and no foam. The second model is a fine-grained model, which is able to classify foam-containing bioreactor images into the classes no foam, low foam, medium foam, and high foam.

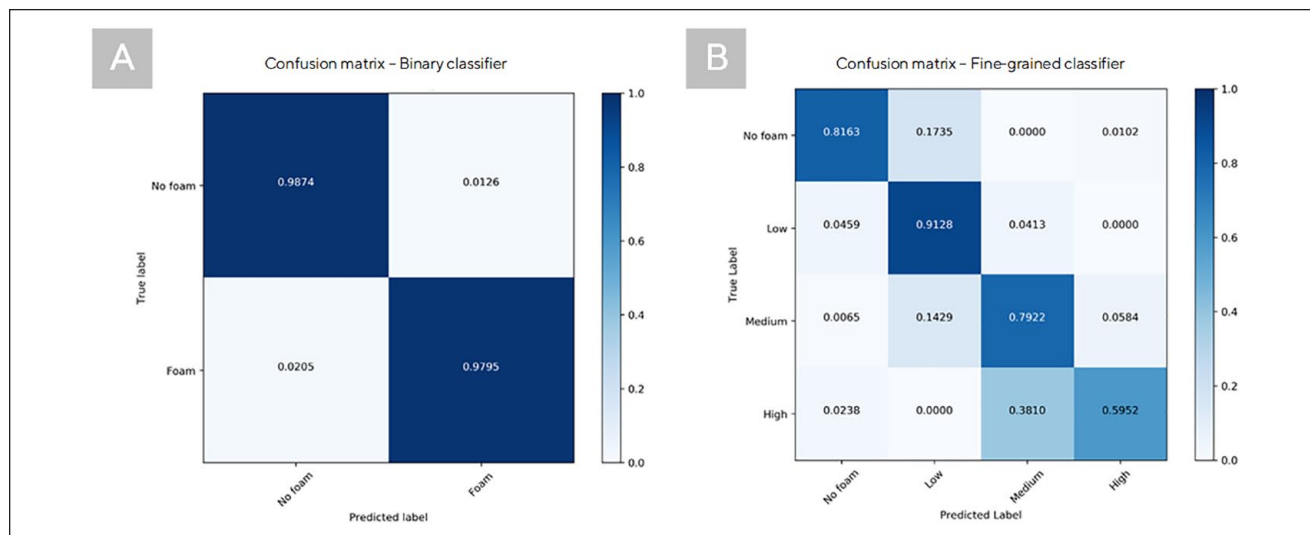
The developed binary foam classification model showed strong performance, with an F1 score higher than 97% on an independent test set (**Table 3**), indicating that the machine vision system reliably detects foam buildup. The fine-grained classifier showed promising results with around a 76% F1 score on the considerably more difficult task of fine-grained classification. Inspecting the confusion matrix for the fine-grained classifier shows that the main source of error is high-foam images mistaken as medium-foam ones and, to slightly lower degree, medium-foam images mistaken as low-foam ones (**Fig. 3B**). The models' predictions were visualized using the GradCAM++ method<sup>21</sup> (**Fig. 4**). This method provides a heat map representation that indicates regions of high decision-making allowing qualitative investigation of the models' predictions. For the binary classifier, the model correctly focuses

on to the liquid–gas interface when making a correct prediction, but focuses on other parts of the vessel when making incorrect predictions (**Fig. 4A**). A similar behavior can be observed for the fine-grained classification model (**Fig. 4B**). Here, too, the model focuses on the foam area for correct predictions and other parts of the vessel environment for incorrect predictions.

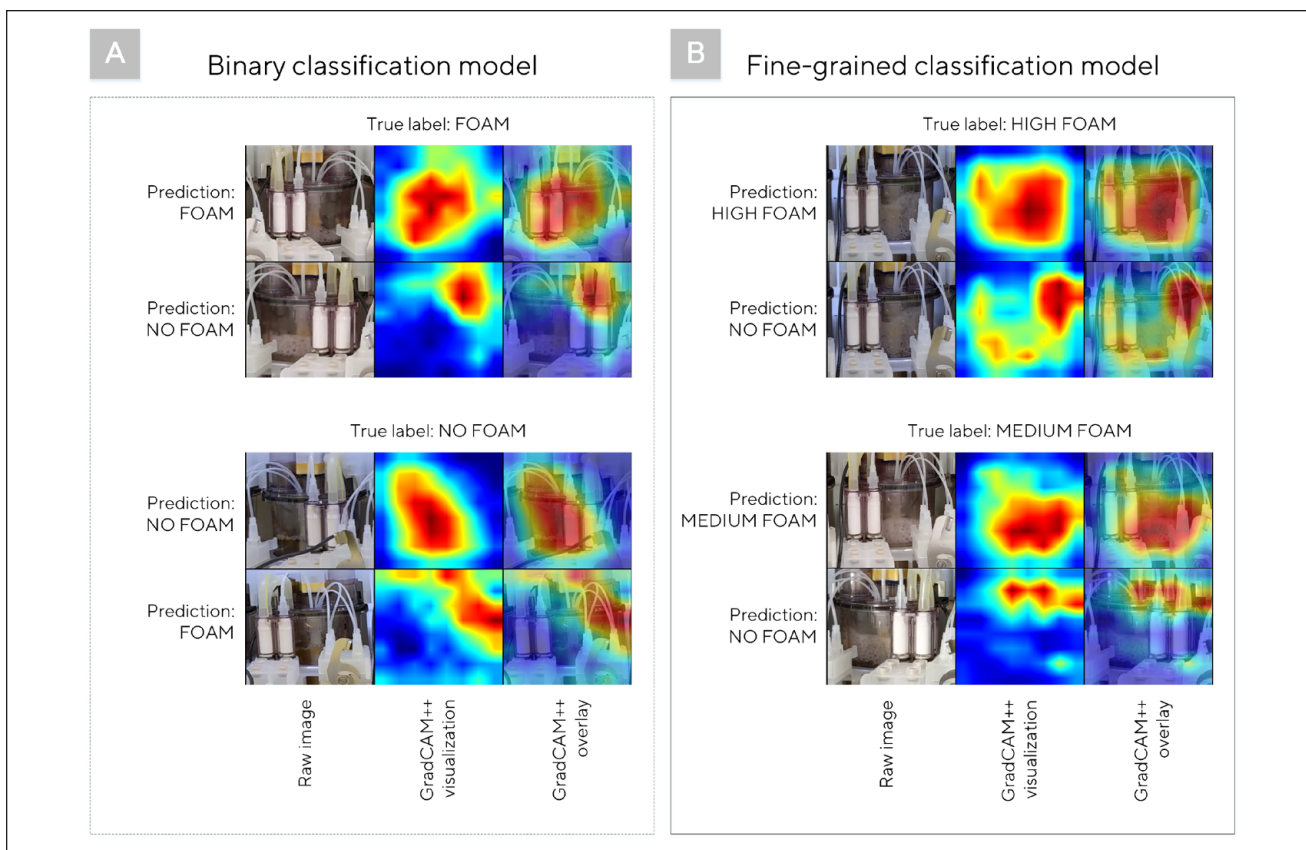
These visualizations allow interpretation of the CNN classifiers' behavior and indicate that the developed models are capable of recognizing the foam region within the vessel area and using it for the classification tasks. Incorrect classifications have been observed, which are the result of the CNNs not focusing on the foam area of the vessel or edge cases introduced by the manual annotation of images. However, these failures may be avoided by averaging the predictions over time-consecutive sequences of images, instead of relying on only a single image, to receive a more robust signal. For example, the frame rate of video acquisition is 30 frames per second, and the worst-case wrong prediction with the binary foam detection is every 48 frames, on average, assuming uniformly randomly distributed failures over time. In this case, averaging the prediction over 30 frames per second may drastically reduce the impact of the misclassifications. This approach is applicable to actual cultivation setups, where foam emergence usually takes several seconds to minutes.

## Concluding Remarks and Outlook

Conventional foam sensor probes show several disadvantages, as they are prone to fouling and coating, and sensitive



**Figure 3.** Confusion matrices of the developed classification models. **(A)** Confusion matrix for the binary classifier mode. **(B)** Confusion matrix for the fine-grained classifier model. Each row indicates the true image labels; columns indicate the respective model's prediction. Figures indicate the proportion of model predictions for images of the labels within each row.



**Figure 4.** Exemplary GradCAM++ visualizations of the developed classification models. **(A)** Visualizations of the binary classification model. **(B)** Visualizations of the fine-grained classification model. Image boxes from left to right: Raw input image; respective GradCAM++ heat map, where blue means low attention and red means high; and the input image with the corresponding heat map overlay.

to reactor conditions such as humidity, agitation splashing, and temperature. This can result in false-negative sensing and overdosing of antifoam, leading to batch failure. Furthermore, the expensive cost and the lack of robustness do not justify their application in single-use bioprocessing setups. The proposed combination of commodity camera modules and the developed CNN approach for real-time foam identification and quantification overcome these drawbacks. The initialized models can be implemented into real-life setups either by hardcoding the ROI into the image acquisition modules or, in the case of flexible module/equipment positions, via preceding object detection tasks that deliver the appropriate ROI, such as deep cropping approaches.<sup>23</sup> The developed algorithm demonstrated high performance regarding foam identification (binary classification, foam/no foam) with an F1 score of 98% on an independent test set. Furthermore, fine-grained classification of foam levels (no foam, low foam, medium foam, high foam) showed good results; the model achieved an F1 score of 76% over all classes, indicating great promise for image-based foam quantification. The main source of error here

was distinguishing between medium-foam and high-foam images, a task difficult even for a subject matter expert if no metric scale for orientation is provided. Furthermore, the foam height is usually not distributed equally along the foam surface, which adds further complexity to this machine vision task. However, it has been shown that the established system provides an inexpensive, accurate, and flexible alternative to traditional foam-sensing systems.

Going forward, implementing an antifoam agent feed strategy based on the resulting fine-grained sensor signal to ensure an antifoam agent addition based on demand would minimize negative effects on bioprocessing equipment features as well as cell behavior. A resulting reduction of batch failures based on antifoam overdosing would improve the efficiency of process development as well as manufacturing processes. Other useful additions to the concept system include implementation of outlier detection capabilities to reduce the impact of process artifacts, for example, the accidental blocking of the camera view to the bioreactor with an object or operator hand during routine operation. Potentially, further accuracy could be added to the model by introducing

exact foam metrics (via surface or volume measurements) as annotation data. Additionally, to further reduce the risk of biased or inconsistent labels, a diverse data set labeled by multiple subject matter experts whose assessments are then aggregated, for instance by majority voting, is preferable, especially in difficult-to-judge edge cases. Furthermore, both models presented in this work can definitely be optimized for higher performance by tuning the model architecture, learning rate, loss function, and so on, which is something we leave for future work.

To conclude, the presented concept results show great promise for the application of machine vision to implement cheap, flexible, and robust monitoring for foam control for upstream bioprocessing. It is anticipated that this machine vision methodology will be further expanded to other areas of bioprocessing.

### Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

### ORCID iDs

Jonas Austerjost  <https://orcid.org/0000-0002-2080-1556>

Christoffer Edlund  <https://orcid.org/0000-0003-0003-3681>

### References

1. Etoc, A.; Delvigne, F.; Lecomte, J. P.; et al. Foam Control in Fermentation Bioprocess: From Simple Aeration Tests to Bioreactor. *Appl. Biochem. Biotechnol.* **2006**, *130*, 392–404.
2. Routledge, S. J. Beyond De-Foaming: The Effects of Antifoams on Bioprocess Productivity. *Comput. Struct. Biotechnol. J.* **2012**, *3*, e201210001.
3. Vardar-Sukan, F. Foaming: Consequences, Prevention and Destruction. *Biotechnol. Adv.* **1998**, *16*, 913–948.
4. Flickinger, M. C.; Delvigne, F.; Lecomte, J. Foam Formation and Control in Bioreactors. In *Encyclopedia of Industrial Biotechnology*; John Wiley & Sons: Hoboken, NJ, 2010; pp 1–13.
5. Vardar-Sukan, F. Foaming and Its Control in Bioprocesses. In *Recent Advances in Biotechnology*; Springer Netherlands: Dordrecht, 1992; pp 113–146.
6. Goldberg, M.; Rubin, E. Mechanical Foam Breaking. *Ind. Eng. Chem. Process Des. Dev.* **1967**, *6*, 195–200.
7. Junker, B. Foam and Its Mitigation in Fermentation Systems. *Biotechnol. Prog.* **2007**, *23*, 767–784.
8. Kawase, Y.; Moo-Young, M. The Effect of Antifoam Agents on Mass Transfer in Bioreactors. *Bioprocess Eng.* **1990**, *5*, 169–173.
9. Wang, J.; Honda, H.; Park, Y. S.; et al. Effect of Dissolved Oxygen Concentration on Growth and Production of Biomaterials by Animal Cell Culture. In *Animal Cell Technology: Basic & Applied Aspects*; Springer Netherlands: Dordrecht, 1994; pp 191–195.
10. Restelli, V.; Wang, M. D.; Huzel, N.; et al. The Effect of Dissolved Oxygen on the Production and the Glycosylation Profile of Recombinant Human Erythropoietin Produced from CHO Cells. *Biotechnol. Bioeng.* **2006**, *94*, 481–494.
11. Liew, M. K. H.; Fane, A. G.; Rogers, P. L. Fouling Effects of Yeast Culture with Antifoam Agents on Microfilters. *Biotechnol. Bioeng.* **1997**, *53*, 10–16.
12. Mohamad Pauzi, S.; Anak Halbert, D.; Azizi, S.; et al. Effect of Organic Antifoam's Concentrations on Filtration Performance. In *Journal of Physics: Conference Series*; Institute of Physics Publishing: Bristol, UK, 2019; Vol. 1349, p 12141.
13. Rod, R. L. Ultrasonic Liquid Level Sensor. In *1957 IRE National Convention Record (IRECON 1957)*; Institute of Electrical and Electronics Engineers: Piscataway, NJ, 1957; pp 36–38.
14. Condé, B. C.; Fuentes, S.; Caron, M.; et al. Development of a Robotic and Computer Vision Method to Assess Foam Quality in Sparkling Wines. *Food Control* **2017**, *71*, 383–392.
15. Cimini, A.; Pallottino, F.; Menesatti, P.; et al. A Low-Cost Image Analysis System to Upgrade the Rudin Beer Foam Head Retention Meter. *Food Bioprocess. Technol.* **2016**, *9*, 1587–1597.
16. Wahab, N.; Khan, A.; Lee, Y. S. Transfer Learning Based Deep CNN for Segmentation and Detection of Mitoses in Breast Cancer Histopathological Images. *Microscopy* **2019**, *68*, 216–233.
17. Hussain, M.; Bird, J. J.; Faria, D. R. A Study on CNN Transfer Learning for Image Classification. In *Advances in Intelligent Systems and Computing*; Springer Verlag: Berlin, 2019; Vol. 840, pp 191–202.
18. Sandner, V.; Pybus, L. P.; McCreath, G.; et al. Scale-Down Model Development in Ambr Systems: An Industrial Perspective. *Biotechnol. J.* **2019**, *14*, 1700766.
19. He, K.; Zhang, X.; Ren, S.; et al. Deep Residual Learning for Image Recognition. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*; IEEE Computer Society: Washington, DC, 2016; Vol. 2016-Decem, pp 770–778.
20. Kingma, D. P.; Ba, J. L. Adam: A Method for Stochastic Optimization. In *3rd International Conference on Learning Representations (ICLR 2015)—Conference Track Proceedings*; San Diego, CA, May 7–9, 2015.
21. Selvaraju, R. R.; Cogswell, M.; Das, A.; et al. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. *Int. J. Comput. Vis.* **2020**, *128*, 336–359.
22. GitHub. stefannc/GradCAM-Pytorch: A Pytorch Implementation of GradCAM, GradCAM++, and Smooth-GradCAM++. <https://github.com/stefannc/GradCAM-Pytorch> (accessed Oct 16, 2020).
23. Wang, W.; Shen, J. Deep Cropping via Attention Box Prediction and Aesthetics Assessment. In *Proceedings of the IEEE International Conference on Computer Vision*; Institute of Electrical and Electronics Engineers: Piscataway, NJ, 2017; Vol. 2017-October, pp 2205–2213.