10.24425/acs.2025.157147

Archives of Control Sciences
Volume 35(LXXI), 2025
No. 4*, pages 771–792*

Convolutional neural networks for froth imaging in a flotation machine with electromagnetic regrinding impeller

Szymon OGONOWSKI and Paweł WÓJTOWICZ

This study investigates a vision-based supervisory framework for a prototype flotation machine equipped with an in-line electromagnetic regrinding impeller, a configuration intended to enhance mineral liberation and recovery for fine particles and tailings. Building on recent advances in machine vision for process industries, convolutional neural networks (CNNs) were trained to infer supervisory variables directly from froth images: (i) bubble population metrics (bubble count/density) and (ii) overall concentrate yield. An image corpus of $>$ 1000 froth frames was collected across three trials under deliberately varied operating conditions (airflow, electromagnetic mill frequency, and feed mass), yielding substantial covariate shift. The results provide preliminary evidence that CNN-based froth imaging can supply actionable supervisory signals in electromagnetic-assisted flotation. Future work will expand datasets, harden illumination robustness, incorporate spatiotemporal modeling and sensor fusion, and evaluate closed-loop control in prospective trials.

Key words: flotation, machine vision, electromagnetic mill, machine learning, supervisory control, process automation, CNN

1. Introduction

Flotation is a fundamental mineral separation process that exploits physicochemical contrasts between valuable mineral surfaces and gangue. In industrial practice, it is preceded by comminution – most critically grinding – which governs mineral liberation and establishes the particle size distribution (PSD) feeding the cells. Because both under- and over-grinding impair attachment and transport (through inadequate liberation in the former and excessive entrainment or slime

---

S. Ogonowski (corresponding author, e-mail: szymon.ogonowski@polsl.pl) and P. Wójtowicz are with Department of Measurements and Control Systems, Faculty of Automatic Control, Electronics and Computer Science, Silesian University of Technology, Gliwice, Poland.

Received 21.04.2025. Revised 12.10.2025.

coating in the latter), precise control of PSD is essential for stable operation and product quality. Equally important is the rigorous dosing and control of reagents (collectors, frothers, and modifiers) as well as air delivery and gas-liquid hydrodynamics (air rate, bubble size distribution, gas holdup, froth depth). Small deviations in these variables alter bubble-particle collision, attachment, and froth stability, leading to pronounced, nonlinear effects on recovery and concentrate grade. Traditional control systems rely on sparse physical sensors and indirect proxies, which provide limited visibility into the separation mechanisms that ultimately manifest at the froth interface.

Recent advances in computer vision and machine learning have enabled a new class of monitoring systems that extract quantitative descriptors from froth images in real time. The visual state of the froth – bubble size and count, texture, color, mobility, and stability – encodes the efficiency of bubble-particle interactions and entrainment, making it a natural target for image-based assessment and control.

The motivation for this research arises from integrating electromagnetic regrinding technology with flotation. Electromagnetic mills enable simultaneous liberation and selective breakage, improving selectivity and energy efficiency, but they also couple grinding intensity to downstream flotation hydrodynamics in ways that complicate supervision and control.

The aim of this paper is to propose and evaluate a vision-based supervisory control framework that fuses real-time froth imaging with process control of a flotation machine equipped with an electromagnetic impeller [17], with specific models for predicting bubble population metrics and overall yield implemented in TensorFlow (Python) [2].

2. Background and literature review

2.1. Fundamentals of froth flotation

Froth flotation separates hydrophobic valuable minerals from hydrophilic gangue by dispersing air into a mineral pulp and promoting selective bubble-particle attachment. At the micro-scale, the process is commonly decomposed into three rate-determining steps: (i) collision between particles and bubbles, (ii) attachment via thin-film drainage and rupture, and (iii) detachment under hydrodynamic stresses; macroscopic performance emerges from the interplay of these phenomena across the collection and froth zones of the cell [49].

In industrial circuits, flotation selectivity is highly sensitive to controllable operating variables – including gas rate (air flow), bubble size distribution (BSD), slurry level, frother/collector/modifier dosage, pulp pH and Eh, and impeller speed – which jointly determine bubble surface area flux S_b and froth stability. Over broad regimes, flotation rate has been shown to scale approximately linearly with S_b, underscoring the centrality of gas dispersion control in plant optimization [20, 49].

A conventional flotation circuit comprises rougher, scavenger, and cleaner stages arranged to maximize recovery at acceptable concentrate grade. Roughers perform the initial separation at high mass pull; scavengers recover residual values from rougher tailings; cleaners upgrade the rougher/scavenger concentrates to specification. Reagent suites and gas-liquid hydrodynamics are tuned at each stage to balance recovery and selectivity, maintain stable recycle streams, and accommodate feed variability [49]. These canonical flowsheets carry distinct control implications for residence time distribution, froth transport, and entrainment, which must be reflected in supervisory strategies [49].

Mechanical (tank) cells use a rotor-stator to suspend solids and disperse air; principal handles include impeller speed, air rate, froth depth, and launder configuration. Column cells separate collection and cleaning zones vertically (counter-current flow with sparged fine bubbles and froth washing), frequently achieving higher grades at lower entrainment for fine and ultrafine feeds [13]. Jameson cells produce intense aeration and very small bubbles in a plunging downcomer, enabling high collision rates and short residence times; industrial retrofits often report improvements in both recovery and circuit stability [24]. Comparative reviews outline mixing, carrying capacity, maximum gas rates, and design features directly relevant to controller design (e.g., gas holdup dynamics, froth residence time) [13, 24]. From a control perspective, these designs differ in input-output sensitivity, time constants, and disturbance propagation – for example, columns exhibit strong interactions between wash water and froth stability, whereas mechanical cells tightly couple impeller speed, S_b, and short-term mass-pull oscillations [13].

2.2. The role of particle size and liberation

Particle size distribution (PSD) and liberation – set upstream by crushing/grinding and often refined by regrinding – critically condition flotation kinetics and selectivity. Classic analyses show that (i) very coarse particles suffer low collision/attachment probability and are prone to detachment, (ii) fines/ultrafines ($<\sim 20{-}30$ µm) exhibit slow kinetics and high water/gangue entrainment, and (iii) an intermediate PSD maximizes recovery for a given hydrophobicity and hydrodynamic regime [47]. The required hydrophobicity for floatability increases toward the extremes of size, producing the characteristic recovery-size curve that motivates PSD control for stable operation and predictable product quality [47, 49].

Recent reviews revisit fine-particle flotation, surveying strategies such as micro/nano-bubble generation, selective (polymeric) collectors, de-sliming, shear conditioning, and the use of columns or Jameson cells to enhance fine-particle collection while mitigating entrainment [13,14]. Emerging results on nanobubble-assisted flotation indicate improved fine-particle attachment and froth stability, though industrial-scale robustness and integration with control frameworks remain active research topics [8].

As ore grades decline and sustainability pressures rise, reprocessing of historical and current tailings has gained momentum. Surveys across sulfide systems identify significant residual metal values – often in fine or poorly liberated size fractions – and report that modern flotation, frequently in conjunction with regrinding and advanced reagents, can economically recover secondary concentrates while reducing environmental liabilities [35, 49]. Control challenges in tailings reprocessing include highly variable feed mineralogy, complex PSDs, and froth regimes sensitive to minor dosage drifts; these favor robust, adaptive supervisory control using inferential sensing and disturbance-rejection strategies [35].

2.3. Froth phase, entrainment, and stability – variables of control interest

Beyond collection-zone kinetics, the froth phase strongly influences grade via entrainment (water and fine gangue carry-over) and selective drainage. Froth stability exhibits a non-monotonic relationship with performance: insufficient stability leads to valuable particle drop-back; excessive stability traps gangue and slows drainage. Practical control therefore targets an operating window of "optimum stability", manipulated by gas rate, frother type/dose, froth depth, wash water (in columns), and launder design [13,49]. Quantitative visual descriptors – froth velocity, bubble size/shape, and mobility – are increasingly used as proxies for stability and mass pull in advanced control [3, 43].

From the control point of view, two facts are of great importance. First, flotation is a multivariable, nonlinear, and time-varying process with strong interactions and disturbances in feed composition, PSD, and water chemistry [43]. Second, the froth surface encodes rich, causal information about separation mechanisms. Over three decades, machine-vision studies have shown that image features – bubble size distributions, texture statistics, and froth velocity/mobility – can serve as inferential sensors for mass pull, grade, and stability, enabling base-layer stabilization and higher-level optimization when fused with conventional measurements [3, 43]. These insights motivate the development of deep-learning models that map froth imagery to key supervisory variables such as bubble population metrics and yield within modern control architectures [3, 43].

2.4. Implications for automation and vision-based supervision

Over the last five years, froth imaging has shifted from hand-crafted texture/shape descriptors to end-to-end deep learning that infers supervisory variables (mass pull, grade, air recovery, stability regimes) directly from images or short video clips. Recent reviews synthesize this transition, documenting convolutional and spatiotemporal networks, transfer learning, and multi-task objectives that jointly estimate bubble size/velocity and metallurgical targets [9, 39]. Empirical studies show that pre-trained CNNs can accurately regress bubble size and froth surface velocity from plant images [25], while improved U-Net/PIV hybrids and attention mechanisms track froth motion fields for stability estimation and regime identification [10]. For property prediction, hybrid architectures (e.g., CNN-LSTM) have been used to forecast ash content/grade from froth video sequences, handling non-stationarity via temporal context [16, 31]. Parallel work tackles robust bubble sizing without explicit segmentation using learned descriptors to cope with occlusions and glare [48]. A broader automation review positions these vision models within intelligent flotation – linking sensors, soft sensors, and control action pathways [9, 10, 16, 25, 31, 39, 42, 48].

Modern frameworks increasingly treat froth imaging as an inferential sensor within multivariable control: deep ensembles classify operating regimes (e.g., under-/over-aerated, unstable froth), estimate bubble population metrics, and provide calibrated confidence to downstream controllers [51]. Vision-derived features (froth speed, directionality, coarseness, bubble size distribution) are fused with conventional tags (air rate, froth depth, cell level) for set-point optimization and disturbance rejection, improving stability under feed PSD and chemistry shifts [3, 43, 51]. Several mature systems demonstrate 24/7 plant-scale reliability:

- Metso VisioFroth™ / FrothSense™ / FrothSense+ – long-running industrial platforms that compute froth speed/velocity, direction, bubble size and stability from camera feeds; FrothSense+ adds neural-network analytics, multi-region analysis and froth height measurement, and can close the loop via Froth Speed Control™ [33].
- FloVis (AMEplus/KGHM) – a computer-vision system deployed in copper plants for regime recognition and supervisory control; it extracts bubble size/shape/color and mobility features and links them to control actions under variable feed conditions [4, 15].
- FLSmidth Smart Vision Systems – froth cameras leveraging deep neural networks to optimize reagent and air control; offered within broader APC suites and retrofit packages for improved froth-level control and stability [12].
- ABB Ability™ Expert Optimizer – APC/MPC platform that integrates vision-based froth metrics and traditional measurements to adjust froth levels, air flows, and reagents for stabilization and recovery/grade targets, building on earlier MPC deployments in zinc and copper circuits [1].

- Mintek FloCam – a low-cost froth camera that estimates froth velocity and height and supports froth-stability-based control (air profiling); used both for operator visualization and automatic stabilization strategies [34].

Historically, industrial datasets from VisioFroth/FrothSense deployments have underpinned air recovery estimation and links between froth transport/velocity and metallurgical outcomes – relationships now being revisited with deep models and uncertainty quantification [3, 9, 30].

2.5. Deep learning and CNN for vision systems

Modern machine learning (ML) encompasses algorithms that learn predictive mappings from data rather than relying on hand-crafted rules. Within ML, deep learning (DL) denotes representation-learning with multilayer neural networks that can hierarchically extract features of increasing abstraction directly from raw signals such as images, audio, or text [19, 29]. In computer vision, DL – especially convolutional neural networks (CNNs) – has supplanted manually engineered features by learning task-aligned visual descriptors end-to-end. This shift has enabled dramatic gains across classification, detection, segmentation, tracking, and metric regression tasks, and it has catalyzed the deployment of vision as a reliable sensor in monitoring, diagnostics, and control pipelines.

In supervised learning, a model $f_\theta\colon R^{\{H\times W\times C\}} \to R^m$ parameterized by θ maps an image x to a target y (class labels or continuous variables). Training seeks parameters that minimize the empirical risk over a dataset $\{(x_i, y_i)\}\ \{i = 1\}_{i=1}^{N}$:

$$\min_{\theta} J(\theta) = \frac{1}{N}\sum_{i=1}^{N} l\left(f_\theta\left(x_i\right), y_i\right) + \lambda\Omega(\theta),$$

where l is a task-appropriate loss (e.g., cross-entropy for classification; l_1/l_2/Huber for regression) and Ω is a regularizer (e.g., weight decay). Parameters are optimized by stochastic gradient descent (SGD) or adaptive variants (e.g., Adam), with updates of the form $\theta \leftarrow \theta - \eta\nabla_\theta l$ computed on mini-batches [19]. Generalization is promoted via data augmentation, architectural priors, regularization (dropout, weight decay), and transfer learning from large-scale pretraining corpora (e.g., ImageNet).

CNNs embed two key priors for images – local connectivity and translation equivariance – through convolutional layers whose learned kernels are shared spatially. A canonical CNN stacks blocks of convolution + nonlinearity (e.g., ReLU/SiLU) and spatial downsampling (strided conv or pooling) to progressively enlarge the receptive field while compressing spatial resolution. High-level features are aggregated by global average pooling and fed to task heads (softmax for classification; linear layers for regression). Architectural refinements such as VGG (deep, homogeneous stacks), ResNet (residual/skip connections enabling very deep networks), and EfficientNet (compound scaling of depth/width/resolution with depthwise-separable convolutions) have delivered strong accuracy – efficiency trade-offs suitable for industrial deployments [23, 45, 46]. CNNs achieve state-of-the-art performance across vision benchmarks and underpin many production systems (e.g., AlexNet's breakthrough on ImageNet; subsequent residual and efficient models) [23, 28].

Effective training requires careful dataset design (balanced classes, representative operating regimes), split hygiene (avoiding temporal or spatial leakage), and augmentation tailored to physical invariances (e.g., small rotations, brightness/contrast jitter) [44]. Transfer learning – initializing from ImageNet-pretrained weights and fine-tuning on the target domain – yields faster convergence and improved generalization in small- to medium-scale industrial datasets. For noisy labels and heavy-tailed errors, robust regressors (e.g., Huber loss) help stabilize learning. Uncertainty estimation (e.g., Gaussian output heads with negative log-likelihood; calibration metrics such as ECE) is increasingly reported to qualify predictions for downstream decision-making [22]. When dynamics matter (videos), temporal extensions – frame-wise encoders with feature pooling, temporal 1-D convolutions, ConvLSTM/GRU layers, or 3-D CNNs – capture motion cues that correlate with process states [11].

In process industries, cameras now function as soft sensors, providing high-bandwidth, spatially resolved information that complements sparse physical measurements. Vision-derived descriptors – texture statistics, bubble size/shape distributions, froth velocity/mobility – can be mapped to hidden process variables (mass pull, product grade, stability regimes) and integrated into monitoring, fault detection, and supervisory control layers [3, 43]. Recent work increasingly replaces hand-crafted features with CNNs that infer key performance indicators directly from images or short clips, enabling robust inferential sensing and opening pathways to closed-loop control when fused with conventional tags within MPC or rule-based frameworks. In domains beyond mineral processing, vision-based feedback has matured in robotics and manufacturing inspection, where CNNs provide perception modules for visual servoing, adaptive quality control, and anomaly detection – illustrating the broader feasibility of using learned vision models in control applications [19, 29].

In summary, deep learning – and CNNs in particular – offers a principled, data-driven route to transform raw images into actionable process variables. The combination of inductive biases (convolution and weight sharing), scalable architectures (ResNet/EfficientNet), and sound training practices (transfer learning, augmentation, calibrated uncertainty) makes CNNs a compelling choice for vision-enabled supervision and control, especially where conventional sensors provide limited observability of key mechanisms.

3. Flotation machine with EM regrinding impeller

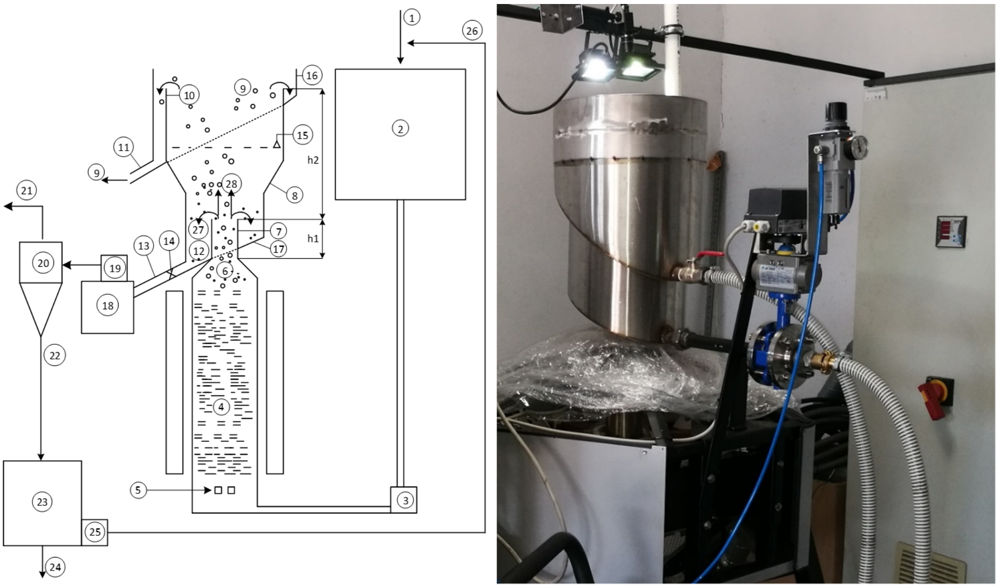
This paper reports experiments on a flotation machine equipped with an in-line regrinding impeller that "activates" particles immediately upstream of bubble-particle contacting, with the goal of enhancing liberation and improving overall recovery. The apparatus is explicitly targeted at fine particles and tailings, where conventional circuits often struggle. While the hydrodynamics and mechanical design of this prototype are still being optimized, work is proceeding in parallel on the automation layer. In contrast to the encouraging body of results available for vision-assisted monitoring and control in classic tank and column cells, the present study is preliminary and focuses on establishing whether a vision-based supervisory control paradigm is feasible for this class of machine.

3.1. Design of the FM

A flotation machine integrated with an electromagnetic mill (EMM) [50] enables a coupled comminution-flotation operation, in which particle liberation and surface conditioning proceed concurrently within a single apparatus. The EMM serves as the comminution module: instead of mechanically driving the grinding chamber, a rotating electromagnetic field directly accelerates ferromagnetic rods that act as grinding media. The working chamber itself is stationary and consists of a non-ferromagnetic tube encircled by six electromagnets that generate the rotating field [50]. When energized, the field induces controlled motion of the rods, producing intensive, stochastic interactions – impact, abrasion, and shear – with the feed material. This contact mechanics yields efficient size reduction and enhanced liberation, while avoiding the mechanical losses associated with moving-vessel mills.

From a design standpoint, electromagnetic and geometric parameters (coil configuration, field amplitude and frequency, rod geometry, and chamber dimensions) are selected to maximize magnetic driving forces on the media while minimizing electrical energy consumption [36]. The direct actuation of the grinding media, the absence of rotating mechanical assemblies, and the ability to tune field parameters provide a compact, responsive platform that can be integrated upstream of or coupled with flotation, facilitating tight control of particle size distribution and surface properties prior to bubble–particle attachment. A schematic diagram of the flotation machine is presented in Fig. 1a, while Fig. 1b illustrates the semi-industrial installation of the flotation machine prototype with direct control system elements (e.g. pulp level sensors, discharge valve) and vision system elements (camera, LED lights) above the tank.

The feed material (1), along with process water and flotation reagents, is introduced into the slurry tank (2), from which it is subsequently pumped under con-trolled pressure into the working chamber of the electromagnetic mill (4),



(a) (b)

Figure 1: Schematic diagram of the system (a) and semi-industrial installation with the prototype of the flotation machine (b)

where particle size reduction is performed. Concurrently, air (5) is injected into the slurry stream, and the generated air bubbles are dispersed by the motion of the grinding media within the active zone of the mill. The resultant dispersion, comprising ground mineral particles and entrained air bubbles (6), is directed through the outlet manifold (7) into the flotation tank (8), wherein phase separation between valuable minerals and gangue occurs.

The froth product (9), enriched with hydrophobic mineral particles, is collected via the overflow sill (10) and extracted through an overflow nozzle (11), while the remaining chamber product (12) is discharged through an outlet nozzle (13) regulated by a control valve (14). In certain applications, such as copper ore flotation, the froth product represents the final output of the integrated grinding-flotation process.

The flow rate of the flotation slurry is managed by adjusting the slurry pump (3) in combination with the control valve (14), with real-time monitoring provided by a slurry level sensor (15). The chamber product is conveyed to a secondary slurry tank (18) and subsequently pumped (19) to a hydrocyclone (20) for classification or thickening. Within the hydrocyclone, the slurry is separated into an overflow fraction (21) and an underflow fraction (22), which is collected in a dedicated tank (23). Based on its particle size distribution and mineralogical content, the underflow may either be rejected as tailings (24) or, if sufficiently coarse and mineral-rich, recycled (26) to the grinding circuit via a pump (25) for further comminution and beneficiation.

3.2. Control system with IoT-based component

The flotation process is governed by a multilayered control system implemented through Programmable Logic Controllers (PLC) and Supervisory Control and Data Acquisition (SCADA) systems [38]. The PLC-based direct control layer enables direct regulation of key operational parameters, including the flow rates of the preprocessed feed, air, and chemical reagents, as well as the flotation pulp level within the flotation tank. The SCADA layer facilitates real-time visualization of process variables, integrates data from diverse monitoring, measurement, and control subsystems, and provides comprehensive database functionalities. Its user interface additionally supports manual set-point adjustments, the scheduling of experiments, and the execution of automatic supervisory algorithms.

In Figure 2, a schematic of the flotation-machine installation is presented with measurements, control loops, and final control elements indicated, where: TLI1 denotes the level measurement in the flotation tank; TLI2, the level measurement in the feed tank; TFI1, the feed-flow measurement to the mill; TFI2, the air-flow measurement to the mill; TTI1 and TTI2, the winding-temperature measurements of the mill; TTI3, the temperature measurement of the material entering the mill; TTI4, the temperature measurement at the mill outlet. The loop FC1 regulates the feed flow to the mill based on TFI1, with the manipulated variable being

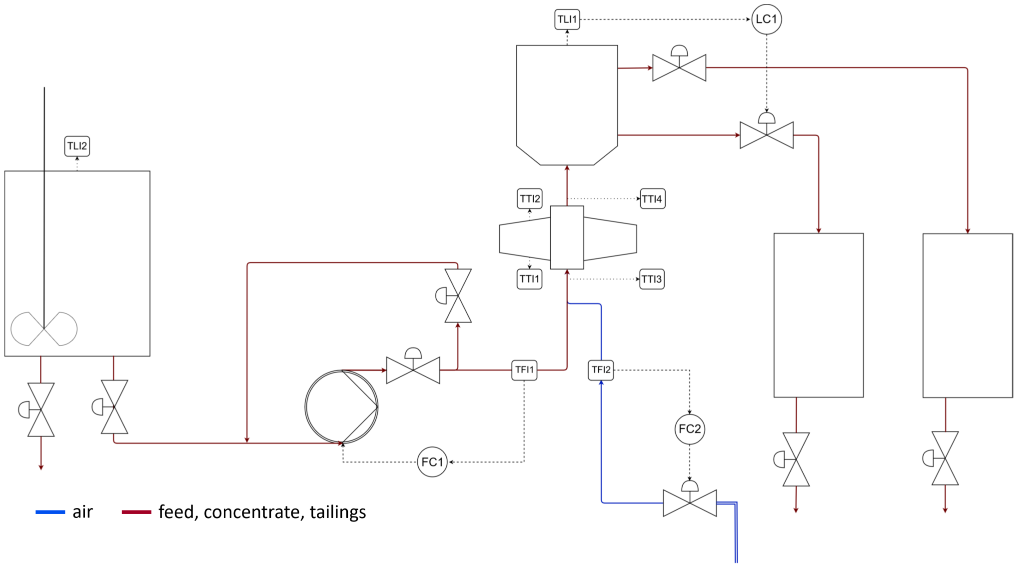


Figure 2: Direct control structure

the pump inverter frequency; FC2 regulates the air flow to the mill based on TFI2, with the manipulated variable being the opening degree of the solenoid valve; LC1 regulates the liquid level in the flotation tank based on TLI1, with the manipulated variable being the throttle angle of the discharge (waste) valve. The figure also identifies the actuators and the type of conveyed medium. Within the SCADA application, the operator may select a manual-control mode for factorial experiments or a standard PID controller whose parameters are editable directly from the SCADA interface. Each control loop can be switched independently between manual and automatic operation. Owing to the specific geometry of the flotation-machine tank (visible in Figure 1), the LC1 level-control loop additionally allows selection of a gain-scheduling mode in which the controller settings are adjusted automatically as a function of the current level measurement. These settings were determined in prior experiments on the basis of an identified non-linear tank model. Dedicated inverters are responsible for the operation of all pumps as well as the electromagnetic mill itself. In addition to standard monitoring functions, the SCADA system is designed to enable the development and online testing of supervisory and optimization control algorithms [36, 37].

To enhance the system's monitoring capabilities, an Internet of Things (IoT)-based vision module was integrated [5, 18]. This module consists of a high-speed LUCID Triton camera [32], a Raspberry Pi 4 computational unit [40], and a Google Coral Artificial Neural Network (ANN) accelerator [21]. The vision system enables continuous acquisition of froth images at rates up to 220 frames per second under varying technological conditions for subsequent analysis and machine learning (ML) applications (see Fig. 3).

(a) (b)

Figure 3: Vision system elements mounted above the flotation machine tank (a) and exemplary images of flotation froth registered during experiments (b)

4. Materials and methods

4.1. Machine learning

Froth-image-based estimation of flotation variables was formulated as a supervised regression task. Given an RGB image of the froth surface $x \in R^{\{H\times W\times C\}}$, a neural network f_θ was trained to predict one or more continuous targets $y{\in}R^m$ (e.g., bubble population metrics and instantaneous yield), such that $\widehat{y} = f_\theta(x;\theta)$. When targets were physically coupled (for example, bubble statistics and mass pull), a multi-task formulation with a shared visual backbone and task-specific heads was adopted to improve sample efficiency and generalization [7, 41].

Froth images were acquired by a fixed industrial camera observing the lip zone under controlled illumination. Ground-truth labels consisted of (i) bubble count density, derived from a reference vision pipeline, and (ii) concentrate yield, obtained from laboratory or plant assays. Because image timestamps and plant tags were not perfectly synchronized, time series were aligned using the lag that maximized cross-correlation between froth-motion proxies and mass pull within a bounded window, thereby reducing label latency [3, 43].

Images were cropped to a froth region of interest, resized, and per-channel normalized using training-set statistics. To improve robustness to variability in lighting, surface texture, and camera pose, bounded photometric and geometric augmentations (random cropping, small rotations, horizontal flips, contrast/brightness jitter, and additive noise) were applied, with ranges selected to preserve the physical meaning of bubble scale and morphology [44].

Convolutional encoders with ImageNet pretraining – specifically ResNet and EfficientNet backbones – were employed, followed by global average pooling and shallow regression heads [23, 46]. For temporal cues, two strategies were examined: (a) temporal pooling of per-frame embeddings and (b) lightweight sequence modeling (temporal 1-D convolution or GRU) atop framewise features [11]. Pretraining was used to mitigate overfitting and to improve convergence on modest industrial datasets [19].

For scalar regression, the Huber loss L_δ was minimized to reduce sensitivity to label noise:

$$L_\delta\left(y,\widehat{y}\right) = \begin{cases} \dfrac{1}{2}\left(y-\widehat{y}\right)^2, & |y-\widehat{y}| \leqslant \delta, \\ \delta\left(|y-\widehat{y}| - \dfrac{1}{2}\delta\right), & \text{otherwise.} \end{cases}$$

In multi-task settings with targets $y^{(k)}$, the objective was defined as the sum of per-task losses with uncertainty-based weights $\lambda_k = 1/\sigma_k^2$ learned during training [26]. Parameters were optimized using Adam with cosine learning-rate decay and L_2 regularization [27].

Data were split by production run/campaign to prevent temporal leakage, producing train/validation/test partitions; generalization was further assessed with grouped (k)-fold cross-validation at the campaign level. Primary metrics included RMSE, MAE, and the coefficient of determination R^2. Spearman's ρ and normalized RMSE (relative to the operating range) were additionally reported. Where probabilistic heads were used, negative log-likelihood and Expected Calibration Error were computed [22].

All experiments were conducted in Python using TensorFlow/Keras for modeling [2], NumPy and Pandas for data handling, scikit-learn for splitting and metrics, OpenCV for image preprocessing, and Matplotlib for visualization. Training and inference were executed on NVIDIA GPUs via the CUDA toolkit, reflecting prevailing best practices for industrial machine-learning workflows.

4.2. Experiments design

All images analyzed in this study were acquired during pilot-scale flotation trials conducted within the CuBR IV Project, Phase I: *Technical feasibility study for the recovery of valuable components from flotation tailings using combined beneficiation methods*. The primary objective was to assess the feasibility of conducting flotation in the proposed prototype apparatus, with particular emphasis on the formation and persistence of bubble–mineral aggregates [49]. Tests were performed as batch flotation in the dispersion chamber only. Given the limited prior operational experience with this device and the ongoing engineering work required to enable continuous operation, the campaign was intentionally exploratory in nature. Experimental runs were designed as periodic trials with controlled parameter variations and included both a fine feed fraction (< 0.020 mm) and a coarse fraction (> 0.100 mm). The fine-fraction tests targeted the efficiency of separating ultrafine particles by exploiting air dispersion and stabilizing durable bubble–mineral attachments in the manufactured prototype, consistent with known constraints and strategies for fine-particle flotation [47]. The coarse-fraction tests focused on verifying the device's capacity for regrinding to achieve liberation, followed by effective flotation of the released minerals [49].

Reagent practice followed industry conventions with adjustments for the batch configuration. A mixed xanthate collector – industrial blend of ethyl and isobutyl sodium xanthogenate with Hostaflot (sodium salt of diethyldithiophosphoric acid) at a mass ratio of 7:3 – was employed as the collector suite, and Nasfroth 245B served as the frother. Owing to the batch mode of operation, frother dosages per unit dry mass of tailings were higher than those typically applied in pneumo-mechanical machines. Collector dosage was systematically varied across runs to quantify its effect on separation performance in the new device [6]. On the equipment side, the principal manipulated variables were the air dosage and the electromagnet (mill) operating frequency.

Across the experimental matrix, the following process variables were deliberately changed: (i) feed characteristics, including particle-size class (fine/coarse), pulp density, and batch volume; (ii) reagent dosages normalized to feed volume or dry mass; (iii) air delivery, encompassing both injection location and volumetric flow rate; and (iv) electromagnetic mill operating frequency.

In summary, the experimental campaign intentionally perturbed operating conditions and, unavoidably, influenced froth characteristics as the team iteratively learned how to stabilize the process across combinations of feed properties, reagent dosages, feed flowrate, and air flowrate. Consequently, the dataset comprises numerous short sequences of images captured under quasi-steady conditions interspersed with transitions, yielding substantial variability in froth appearance (e.g., bubble size distributions, color/brightness), which is advantageous for training robust visual features. However, the number of images obtained from extended, fully stable operating points is comparatively limited, which constrains the statistical power for evaluating model performance under steady-state regimes.

5. Model design, training and results

An image corpus exceeding one thousand froth frames was assembled across three experimental trials, each comprising multiple concentrates and deliberately varied operating conditions (airflow, electromagnetic regrinding frequency, and feed mass). Natural covariate shift was thus introduced by design. Each image was accompanied by reference annotations from a vision system (bubble count, density, displacement magnitude/angle, mean RGB) and, at batch level, laboratory yield. Because the dataset was small by deep-learning standards, transfer learning, augmentation, and leakage-aware validation were emphasized [23, 44].

Hold-out splits (approximately 80/20 within trial) were performed and merged to form global train/test sets while preserving campaign boundaries; grouped k-fold cross-validation provided variance estimates. Early stopping on validation loss (with a fixed patience) was employed. Unless otherwise stated, Adam with an initial learning rate of 3×10^{-4}, mini-batches constrained by GPU memory, cosine decay, and weight decay 10^{-4} were used [27].

5.1. CNN architecture design and ablation outcomes

Initially, a baseline model was constructed to serve as a reference point. Such a model facilitates understanding of the fundamental characteristics of the task and provides a starting point for subsequent experiments and model improvements. The initial architecture comprised six hidden layers, including two convolutional layers and two interleaved pooling layers, followed by two fully connected layers. All layers employed the Rectified Linear Unit (ReLU) activation function. Parameter optimization was performed using the Adam algorithm with its default learning rate. The model accepts as input an image with a resolution of 365 × 310 pixels and produces a single scalar output corresponding to the predicted number of air bubbles in the image. The described neural network architecture is shown in Figure 4.

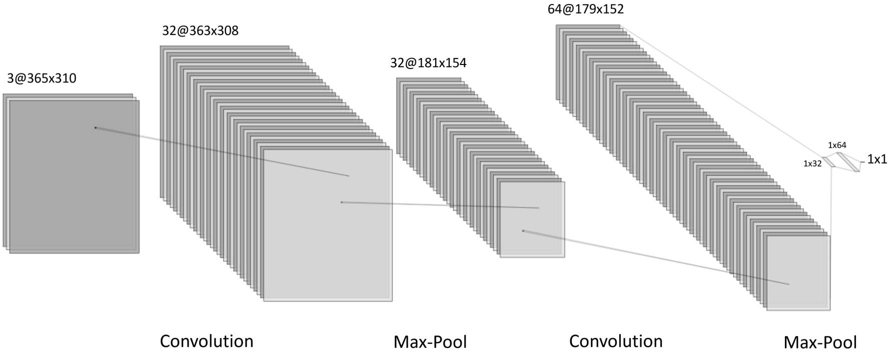


Figure 4: General structure of the baseline model

The final architecture was selected through an ablation-driven process that began with progressively deeper CNN baselines trained from scratch and proceeded to pretrained encoders. Initial baselines with one convolutional block were found to overfit severely, exhibiting low training error and high test error, whereas two-block configurations provided the best trade-off among scratch models; further depth degraded generalization due to over-parameterization. On this basis, transfer learning was introduced, and pretrained ResNet/EfficientNet backbones with global average pooling and a shallow regression head were adopted, as these models consistently improved test RMSE and training stability relative to scratch networks [23, 46].

Alternative input pathways were then examined. A second image branch carrying handcrafted, preprocessed inputs (binary thresholding or Scharr edge maps) was evaluated to encourage explicit sensitivity to bubble boundaries. This design did not yield generalization gains and, in several cases, reduced robustness, which was attributed to the brittleness of such transforms under varying illumination and froth brightness [44]. Consequently, the dual-branch image pathway was discarded in favor of a single, learned feature extractor.

Auxiliary non-visual inputs were also considered. A compact multilayer perceptron (MLP) branch for control tags (airflow and electromagnetic-mill frequency) was concatenated with visual embeddings. While average gains were modest, occasional benefits were observed in low-contrast regimes; therefore, the tag branch was retained as an optional module for scenarios where auxiliary measurements are available, while the default configuration remained vision-only to maximize portability.

Error analyses identified a failure mode on very bright froth images, consistent with diminished contrast and reduced detectability of bubble boundaries [3]. To mitigate this, augmentation ranges for brightness and contrast were widened (within physically plausible limits), and the final model selection emphasized architectures whose validation performance degraded least under such perturbations. Taken together, these results led to a final design comprising a pretrained EfficientNet/ResNet backbone, global average pooling, and a lightweight regression head (with dropout for regularization), optionally fused with a small tag-MLP when control signals are present. This design was preferred for its balance of generalization, robustness to illumination changes, and computational efficiency on GPU hardware. Inference latency was shown to scale with model depth and input resolution. Substantial speedups were obtained on GPUs relative to CPUs, particularly for deeper or multi-input models. Practical acceleration options included moderate down-scaling of inputs, the use of depthwise-separable backbones (e.g., EfficientNet-B0), and standard compression techniques (pruning/quantization) for deployment [46].

5.2. Task A – Bubble count prediction

Prediction of the number of bubbles per image was pursued as a proxy for the underlying bubble population. Progressive CNN baselines with one to three convolutional blocks were evaluated and compared against off-line analyses produced by the AMEplus FloVis system [4]. The FloVis bubble counts served as a reference benchmark during training and evaluation. It should be emphasized that FloVis employs entirely different, hand-crafted image-analysis procedures for froth characterization and does not rely on machine-learning methods.

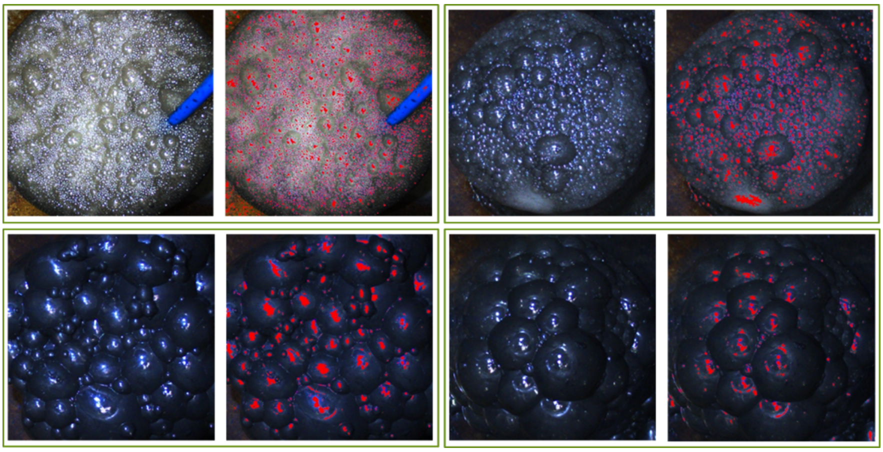


Figure 5: Four examples of froth images: raw (left side) and after FloVis system analysis for bubbles recognition (right side), showcasing differences of froth structure, brightness and color

It was found that shallow models trained from scratch tended to overfit – exhibiting low training error but high test error – indicating insufficient inductive bias for the available data volume. Among the scratch configurations, two-block CNNs offered the best trade-off, whereas deeper stacks degraded due to over-parameterization. In contrast, transfer-learned backbones improved test RMSE and stability. Adding a second input branch with handcrafted preprocessing (thresholded or Scharr-filtered images) did not improve generalization, which was attributed to brittleness under varying froth brightness [44] (compare with Fig. 3b). Incorporation of auxiliary control tags (airflow and electromagnetic-mill frequency) yielded only modest average gains, though occasional benefits were observed in regimes with weak visual contrast. The poorest performance arose for very bright froth images from a specific trial and concentrate, consistent with diminished contrast and reduced detectability of bubble boundaries [3]. However, results obtained in this task support the feasibility of vision-based inference for bubble population metrics even in the preliminary phase of the new flotation machine development. Figure 6 represent exemplary results of the bubbles count

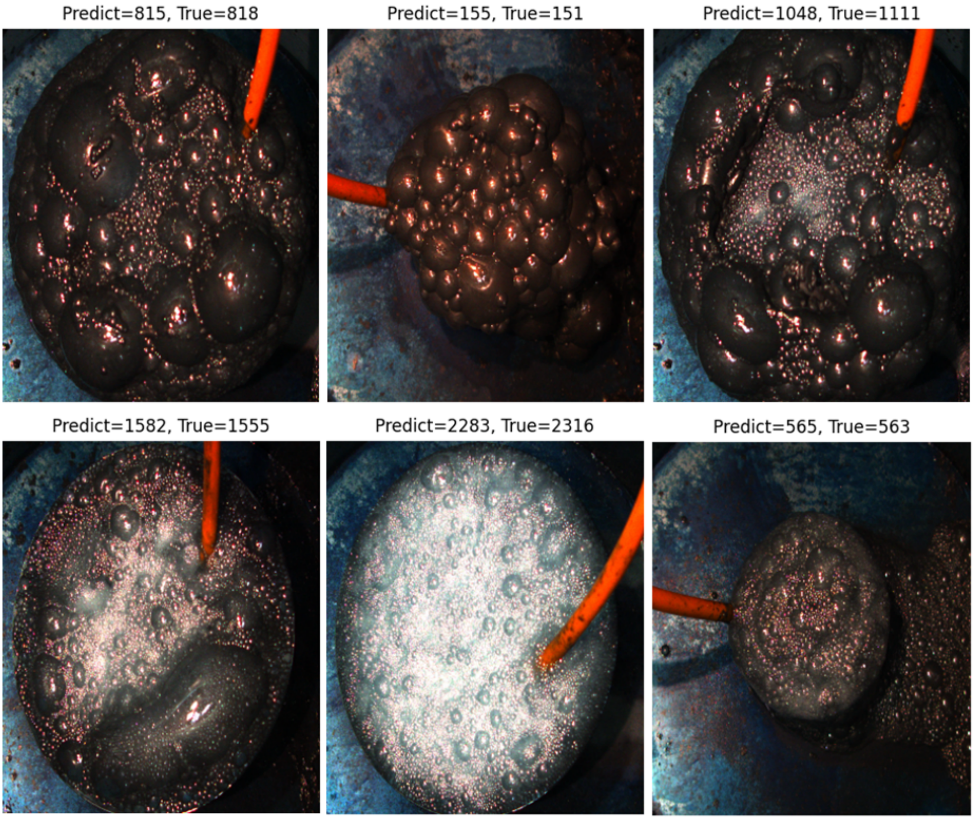


Figure 6: Examples of froth images with detection results comparison

prediction in the froth image compared to the "true" results obtained with FloVis system. The presented set showcase the best results with detection error < 0.5% and the worst cases with detection error > 6%.

5.3. Task B – Yield prediction

Two formulations were investigated for yield prediction. First, classification into three yield bands (< 10%, 10–20%, > 20%) yielded seemingly perfect accuracy on both training and test sets; however, this outcome was invalidated by temporal redundancy – near-duplicate frames within a run – revealing split leakage [19]. Second, numeric regression was attempted; because a single batch yield corresponded to many frames, small per-batch perturbations were injected during training to prevent degenerate memorization. The best regressor achieved an RMSE of approximately 0.76 on a held-out test split but showed degraded performance on entirely unseen runs, indicating sensitivity to distribution shift and underscoring the need for more diverse data and stricter leakage controls [41].

5.4. Summary of the results

Across three trials and > 1000 froth images with deliberate covariate shift, the ablation study converged on a transfer-learned CNN (EfficientNet/ResNet backbone with global average pooling and a lightweight regression head) as the most reliable architecture, outperforming shallow scratch models that overfit and deeper scratch models that under-generalized. Hand-crafted preprocessed inputs (thresholding/Scharr) did not confer benefits and reduced robustness under brightness variation, whereas standard augmentation and leakage-aware validation proved essential for stability on this small industrial dataset. For Task A (bubble count), the final model achieved improved test RMSE and qualitative robustness, with the principal failure mode arising on very bright froth images where bubble boundaries were weakly contrasted. For Task B (yield), band-wise classification appeared perfect but was invalidated by temporal leakage, while numeric regression attained RMSE ≈ 0.76 on held-out splits yet degraded on unseen campaigns, indicating sensitivity to distribution shift and the need for broader data and stricter split hygiene. Runtime profiling confirmed that GPU inference scaled favorably with model complexity, enabling practical deployment envelopes, and the chosen optimizer and schedule (Adam with cosine decay and weight decay) supported stable convergence across experiments. Overall, the results support the feasibility of vision-based inference for bubble population metrics and provide preliminary evidence for yield prediction in flotation machine with electromagnetic regrinding impeller, while highlighting data diversity, illumination robustness, and leakage control as the critical levers for continued improvement.

6. Conclusions and future work

The study undertaken should be regarded as preliminary, as it has been conducted in parallel with ongoing optimization of both the mechanics and hydrodynamics of a prototype flotation machine (integrated with an in-line regrinding impeller) and with the development of an accompanying supervisory control layer. Within this evolving context, convolutional neural networks were shown to provide plausible image-based estimates of bubble population metrics and, to a more limited extent, of concentrate yield. Transfer learning, augmentation, and uncertainty-aware multi-task training were found to be beneficial under the constraints of a modest dataset; however, susceptibility to temporal leakage and distribution shift was observed, particularly for yield prediction, where performance deteriorated on previously unseen operating campaigns.

These observations suggest several directions for continued research. First, expansion of the dataset across more operating regimes, illumination conditions, and ore types is required to reduce variance and to harden models against distribution shift. Second, improvements to the prototype machine – including air dispersion, froth management, and regrinding intensity – are being pursued, as these mechanical refinements directly influence froth morphology and thus the observability of process states in images. Third, migration from static image regression to spatiotemporal modeling and sensor fusion (combining vision with control tags and flow/chemistry measurements within a state-space framework) is expected to improve identifiability and stability, enabling closed-loop applications. Finally, rigorous leakage-aware evaluation protocols and prospective trials on unseen campaigns are planned to establish external validity, while real-time deployment studies will be undertaken to quantify the value of vision-based supervisory signals within advanced process control. In sum, although encouraging evidence of feasibility has been obtained, substantial work remains before reliable, plant-hardened, vision-based supervision can be asserted for this specific class of flotation machinery.

References

[1] ABB: Ability™ *Expert Optimizer for grinding & flotation* (*APC/MPC integration*). ABB, (n.d.).

[2] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, et al.: TensorFlow: A System for Large-Scale Machine Learning. *Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI*), (2016), 265–283.

[3] C. Aldrich, C. Marais, B.J. Shean and J.J. Cilliers: Online monitoring and control of froth flotation systems with machine vision: a review. *International Journal of Mineral Processing*, **96**(1-4), (2010), 1–13. DOI: 10.1016/j.minpro.2010.04.005

[4] AMEplus: *FloVis brochure*. AMEplus, https://www.ameplus.pl/_pdf/ameplus_flovis.pdf

[5] K. Ashton: That “Internet of Things” Thing. *RFID Journal*, (2009).

[6] S.M. Bulatovic: *Handbook of Flotation Reagents: Chemistry, Theory and Practice*. Elsevier, Amsterdam 2007. DOI: 10.1023/A:1007379606734

[7] R. Caruana: Multitask learning. *Machine Learning*, 28(1), (1997), 41-75.

[8] J. Chen, J. Chen and Y. Cheng: Current status of research on nanobubbles in particle flotation. *Physicochemical Problems of Mineral Processing*, 60(1), (2024). DOI: 10.37190/ppmp/183613

[9] X. Chen, D. Liu, L. Yu, P. Shao, M. An and S. Wen: Recent advances in flotation froth image analysis via deep learning. *Engineering Applications of Artificial Intelligence*, 147 (2025). DOI: 10.1016/j.engappai.2025.110283

[10] Y. Chen, D. Xu and K. Wan: A froth velocity measurement method based on improved U-Net++ semantic segmentation in flotation process. *International Journal of Minerals, Metallurgy and Materials*, **31** (2024), 1816–1827. DOI: 10.1007/s12613-023-2787-2

[11] J. Donahue, L.A. Hendricks, S. Guadarrama, S. Venugopalan and T. Darrell: Long-term recurrent convolutional networks for visual recognition and Description. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA, (2015). DOI: 10.1109/CVPR.2015.7298878

[12] FLSmidth: *Sensors & Vision Systems (DNN-based froth cameras); Froth Recovery upgrades*. FLSmidth.

[13] J.A. Finch and G.S. Dobby: Column flotation: a selected review. Part I. *International Journal of Mineral Processing*, **33**(1-4) (1991), 343–354. DOI: 10.1016/0301-7516(91)90062-N

[14] M.H.G. Flavarjani, A. Sobouti, B. Rezai and A. Ahmadi: A review of fine particle flotation: mechanistic insights and strategies. *Minerals Engineering*, **234** (2025). DOI: 10.1016/j.mineng.2025.109685

[15] J. Galas and D. Litwin: Machine learning technique for recognition of flotation froth images in a nonstable flotation process. *Minerals*, **12**(8), (2022). DOI: 10.3390/min12081052

[16] Y. Gan, G. Zhang, F. Lu and X. Wang: Predicting future velocity of mineral flotation froth using SMA-LSTM. *Measurement*, **229** (2024). DOI: 10.1016/j.measurement.2024.114436

[17] T. Gawenda, D. Foszcz, D. Saramak and D. Krawczykowski: *Crushing and flotation machine. European Patent EP3560598 B1*. 2019.

[18] A. Gilchrist: *Industry 4.0: The Industrial Internet of Things*. Apress, New York 2016.

[19] I. Goodfellow, Y. Bengio and A. Courville: *Deep Learning*. MIT Press, Cambridge, MA (2016).

[20] B.K. Gorain, J.P. Franzidis and E.V. Manlapig: Studies on impeller type, impeller speed and air flow rate in an industrial scale flotation cell. Part 1: Effect on bubble size distribution. *Minerals Engineering*, **8**(6), (1995), 615-635. DOI: 10.1016/0892-6875(95)00025-L

[21] Google: *Coral USB Accelerator*. https://coral.ai/products/accelerator

[22] C. Guo, G. Pleiss, Y. Sun and K.Q. Weinberger: On calibration of modern neural networks. *34th International Conference on Machine Learning (ICML)*, Sydney, Australia, (2017). DOI: 10.48550/arXiv.1706.04599

[23] K. He, X. Zhang, S. Ren and J. Sun: Deep residual learning for image recognition. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (2016). Las Vegas, NV, USA, DOI: 10.1109/CVPR.2016.90

[24] G.J. Jameson: The Jameson cell. In: M.C. Fuerstenau, G.J. Jameson, R.-H. Yoon (Eds): *Froth Flotation: A Century of Innovation*. Society for Mining, Metallurgy, and Exploitation, Inc. (SME), 2010.

[25] A. Jahedsaravani, M. Massinaei and M. Zarie: Measurement of bubble size and froth velocity using convolutional neural networks. *Minerals Engineering*, 204 (2023). DOI: 10.1016/j.mineng.2023.108400

[26] A. Kendall, Y. Gal and R. Cipolla: Multi-task learning using uncertainty to weigh losses for scene geometry and semantics. *IEEE Conference on Computer Vision and Pattern Recognition (CVF)*, Salt Lake City, UT, USA (2018). DOI: 10.1109/CVPR.2018.00781

[27] D.P. Kingma and J. Ba: Adam: A method for stochastic optimization. *International Conference on Learning Representations (ICLR)*, San Diego, USA (2015). DOI: 10.48550/arXiv.1412.6980

[28] A. Krizhevsky, I. Sutskever and G.E. Hinton: ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, **60**(6), (2017), 84-90. DOI: 10.1145/3065386

[29] Y. LeCun, Y. Bengio and G. Hinton: Deep learning. *Nature*, **521** (2015), 436–444. DOI: 10.1038/nature14539

[30] J. Leiva, L. Vinnett and J. Yianatos: Estimating air recovery by measuring froth transport over the lip in a bi-dimensional flotation cell. *Minerals Engineering*, **36-38** (2012), 303–308. DOI: 10.1016/j.mineng.2012.05.023

[31] Y. Li, H. Liu and F. Lu: Research on prediction of ash content in flotation-recovered clean coal based on NRBO-CNN-LSTM. *Minerals*, **14**(9), (2024). DOI: 10.3390/min14090894

[32] LUCID Vision Labs: Triton GigE machine-vision camera. https://thinklucid.com/triton-gige-machine-vision

[33] Metso: VisioFroth™ / FrothSense™ / FrothSense+ product pages and application notes. Metso, (2024–2025).

[34] Mintek: FloCam froth stability monitoring and control. Mintek, 2019.

[35] H. Na, Y. Eom, K. Yoo and R.D. Alorro: A review on the reprocessing of sulfide tailings for resource recovery and AMD prevention using mineral processing methods. *Minerals Engineering*, **218** (2024). DOI: 10.1016/j.mineng.2024.109025

[36] S. Ogonowski: On-line optimization of energy consumption in electromagnetic mill installation. *Energies*, **14**(9), (2021). DOI: 10.3390/en14092380

[37] S. Ogonowski, Z. Ogonowski and M. Świerzy: Power optimizing control of grinding process in electromagnetic mill. *21th International Conference on Process Control (PC)*, Strbske Pleso, Slovakia, (2017), 370–375. DOI: 10.1109/PC.2017.7976242

[38] S. Ogonowski, Z. Ogonowski and M. Pawełczyk: Multi-objective and multi-rate control of the grinding and classification circuit with electromagnetic mill. *Applied Sciences (Switzerland)*, **8**(4), (2018). DOI: 10.3390/app8040506

[39] J. Qu, S. Luukkanen and H. Wan: Machine vision-driven analysis of flotation froth properties: A comprehensive review of multidimensional feature extraction and intelligent optimization. *Minerals Engineering*, 233 (2025). DOI: 10.1016/j.mineng.2025.109591

[40] Raspberry Pi Foundation: Raspberry Pi 4 Model B. https://www.raspberrypi.com/products/raspberry-pi-4-model-b/

[41] S. Ruder: An overview of multi-task learning in deep neural networks. arXiv:1706.05098, (2017). DOI: 10.48550/arXiv.1706.05098

[42] H. Dong, F. Wang, D. He and Y. Liu: Flotation equipment automation and intelligent froth feature extraction in flotation process: a review. *Reviews in Chemical Engineering*, **41**(3), (2025), 225–239. DOI: 10.1515/revce-2024-0023

[43] B.J. Shean and J.J. Cilliers: A review of froth flotation control. *International Journal of Mineral Processing*, **100**(3-4), (2011), 57–71. DOI: 10.1016/j.minpro.2011.05.002

[44] C. Shorten and T.M. Khoshgoftaar: A survey on image data augmentation for deep learning. *Journal of Big Data*, **6** (2019). DOI: 10.1186/s40537-019-0197-0

[45] K. Simonyan and A. Zisserman: Very deep convolutional networks for large-scale image recognition. *International Conference on Learning Representations (ICLR)*, San Diego, CA, USA, (2015).

[46] M. Tan and Q.V. Le: EfficientNet: Rethinking model scaling for convolutional neural networks. *36th International Conference on Machine Learning (ICML)*, Long Beach, California (2019).

[47] W.J. Trahar: A rational interpretation of the role of particle size in flotation. *International Journal of Mineral Processing*, **8**(4), (1981), 289–327. DOI: 10.1016/0301-7516(81)90019-3

[48] L. Vinnett, R. Leon and D. Mesa: Artificial neural network (ANN) modelling to estimate bubble size from macroscopic image and object features. *Physicochemical Problems of Mineral Processing*, **59**(5), (2023). DOI: 10.37190/ppmp/209962

[49] B.A. Wills and T. Napier-Munn: *Wills' Mineral Processing Technology*. (8th ed.). Elsevier, Amsterdam 2015.

[50] M. Wołosiewicz-Głąb, D. Foszcz and S. Ogonowski: Design of the electromagnetic mill and the air stream ratio model. *IFAC-PapersOnLine*, **50**(1), (2017), 14964–14969. DOI: 10.1016/j.ifacol.2017.08.2554

[51] X. Zhou and Y. He: Deep ensemble learning-based sensor for froth image recognition. *Sensors*, **24**(15), (2024). DOI: 10.3390/s24155048