Journal of Hazardous Materials Advances 19 (2025) 100870

Contents lists available at ScienceDirect

Journal of Hazardous Materials Advances

journal homepage: www.elsevier.com/locate/hazadv

Image-based machine learning applications for flocculation modelling in water treatment: Prospects towards automation

Abayomi O. Bankole [a,b,c,*], Rodrigo Moruzzi [a,b,f,*], Rogério G. Negri [b,d], John Bridgeman [e], Soroosh Sharifi [f]

[a] Department of Civil and Environmental Engineering, São Paulo State University (UNESP), Bauru Campus, São Paulo, Brazil
[b] Institute of Science and Technology, São Paulo State University (UNESP), São José dos Campos, São Paulo, Brazil
[c] Department of Water Resources Management and Agrometeorology, COLERM, Federal University of Agriculture, Abeokuta, Nigeria
[d] Graduate Program in Natural Disasters (UNESP/CEMADEN), São José dos Campos, São Paulo, Brazil
[e] Department of Civil and Environmental Engineering, School of Engineering, University of Liverpool, Liverpool, UK
[f] Department of Civil Engineering, School of Engineering, University of Birmingham, Birmingham, UK

ARTICLE INFO

Keywords:
Advanced vision intelligence
Deep learning
Sensitivity analysis
Pollution monitoring
Smart water treatment
System automation

ABSTRACT

Advancements in machine learning (ML) underscore the necessity for real-time monitoring in water treatment processes, particularly the flocculation phase. Digital imaging profoundly enhances the monitoring of precipitation, sedimentation, and especially coagulation and flocculation units. Flocculation is an essential phase that influences subsequent treatment stages like sedimentation and filtration, as it aids the removal of diverse pollutants. Yet, despite significant progresses in ML, and its advanced deep learning (DL) techniques, image-based flocculation modelling remains under-represented due to limited research. This review bridges that gap by critically examining image-based ML and DL techniques for flocculation modelling, charting a pathway towards automation. First, the evolution of flocculation kinetics modelling and best practices in floc image acquisition techniques are outlined. The significance of advanced vision platforms and image processing methodologies in flocculation monitoring is discussed, highlighting the benefits and limitations of various operators. State-of-the-art DL algorithms for semantic segmentation, data augmentation, weakly supervised learning, and real-time segmentation techniques are carefully examined, emphasizing their transformative potential in automated flocculation modelling and water treatment monitoring. Findings showed that by integrating sensitivity analysis with sparse sampling and weakly supervised learning, we propose innovative strategies to accelerate accurate floc mask generation for robust databases. Implementing real-time segmentation algorithms promises to revolutionize pollutant monitoring during treatment, propelling system automation towards Industry 4.0 and 5.0 standards. Our insights offer a roadmap for future research, aiming to promote the adoption of automated systems in water treatment and pollutant tracking.

1. Introduction

1.1. Background

The ultimate goal of water and wastewater treatment processes is to efficiently remove pollutants and provide safe water (Altowayti et al., 2022; Cleasby, 1984; Crittenden et al., 2012; Gregory, 2009), in addition to a parallel target of achieving climate friendly treatment process (James et al., 2023, 2025). Safe water supply systems are critical to achieving a sustainable and resilient society, especially given the interrelatedness of water with all aspects of human life (Saqr et al., 2025; Yu et al., 2023). This makes water treatment sustainability critical to meeting the 2030 Sustainable Development Goals and the 2050 projected global water scarcity (Bankole et al., 2022b, 2023; Safeer et al., 2022). The rapid increase in global population presents a major challenge for most water and wastewater treatment plants (WWTP) (Borges et al., 2022; James et al., 2023; Li et al., 2021). For instance, production pressure on drinking water facilities has increased due to the daily increase in water demand, which is exacerbated by shortage in safe water and increased pollution rate (Odjegba et al., 2021, 2024).

Flocculation is a very crucial and complex stage in water treatment, fundamental to pollutant removal, and ensuring safe water supply at

* Corresponding authors at: Department of Civil and Environmental Engineering, São Paulo State University (UNESP) Bauru Campus, São Paulo, Brazil.
E-mail addresses: abayomi.o.bankole@unesp.br (A.O. Bankole), rodrigo.moruzzi@unesp.br (R. Moruzzi).

https://doi.org/10.1016/j.hazadv.2025.100870
Received 14 March 2025; Received in revised form 12 June 2025; Accepted 24 August 2025
Available online 25 August 2025
2772-4166/© 2025 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

conventional and emerging treatment plants (Bankole et al., 2024a; Moruzzi et al., 2020b; Moruzzi et al., 2017). The flocculation phase aims to promote the formation of flocs by aggregating suspended particles and the contaminants in water through induced shear stress. Due to the complex relationships guiding the process, such as the induced shear velocity (G_f), flocculation time (T), coagulant type, and particle-particle interactions (Gregory, 2009; Moruzzi et al., 2017, 2018), existing practice relies on multiple jar test experiments to determine, inter alia, optimum coagulant dosage, flocculation time, reactor design, among others. The lag time created by the process and analyses, has resulted in problems such as chemical wastage, operator time loss, equipment overburden, high production costs, and particularly lack of early warning measures (Altowayti et al., 2022; Bankole et al., 2024a; Li et al., 2024a; Yu et al., 2010).

Therefore, flocculation and water treatment process modelling presents a practical and cost-effective way to tackle the current operational complexity (Chowdhury and Karanfil, 2024). This is evidenced through several numerical models built to evaluate the effectiveness and kinetics of flocculation in batch, pilot and large-scales water treatment (Bratby, 2016; He et al., 2012; Lu and Spielman, 1985; Smoluchowski, 1917), comprehensively reviewed by Jarvis et al. (2005), Thomas et al. (1999), Watanabe (2017) and Wei et al. (2018). Nonetheless, numerical modelling methodologies are highly technical, time consuming and complex to solve in some cases, requiring subject matter experts (Bankole et al., 2024a). These challenges collectively have created possibilities of using machine learning (ML) and its advanced subset, deep learning (DL), as a promising tools to model flocculation processes and kinetics, offering significant potential for improving treatment plant performance (Wang et al., 2022a; Yokoyama et al., 2024).

Recently, ML has been successfully used to predict pathogen removal and treatment efficiency in large scale water works, resulting in significant improvements in treatment plants performance (Altowayti et al., 2022; Bressane et al., 2023; Hamada et al., 2018; Nasr et al., 2012; Wang et al., 2021, 2022a; Zhang et al., 2013a). Similarly, the coagulation-flocculation phase has been modelled using diverse ML algorithms, including tree-based models (Achite et al., 2022; Wang et al., 2022a), artificial neural networks – ANN (Kim and Parnichkun, 2016; Liang and Liu, 2023; Zhang et al., 2013a), different variants of neuro-fuzzy inference systems (Kim and Parnichkun, 2016; Liang and Liu, 2023; Yuan et al., 2016), recurrent neural network and deep neural network models (Farhi et al., 2021; Pisa et al., 2020), for coagulant dosage and treated water quality prediction. Bankole et al. (2024a) and Oliveira et al. (2018) also modelled aggregate floc evolution and fractal dimension evolution using ANN.

These adoptions have also led to a series of reviews on water and wastewater treatment modelling using ML and DL. However, a significant research gap persists concerning the comprehensive modelling of flocculation mechanisms and kinetics. Existing reviews on ML and DL applications in water treatment have largely overlooked these critical aspects, often focusing solely on coagulant dosage rather than the underlying forces, mixing intensity, particle concentration, or fluid mechanics governing floc formation. For instance, Chowdhury and Karanfil (2024) in their review on possibilities of artificial intelligence implementation in drinking water treatment provides detailed account of over 1200 articles on ML and DL implementation across batch-scale, laboratory-scale, and full treatment plants modelling, including the coagulation and flocculation. Similarly, Alvi et al. (2023) conducted a critical review on the prospect of DL as an alternative to biochemical models in wastewater treatment, reporting on over twenty-six variants of ML algorithms and soft computing approaches that have been successfully used in wastewater treatment. Altowayti et al. (2022) also conducted a comprehensive review on the role of ML and DL algorithms in conventional treatment, with a combination of detailed bibliographic analysis.

Safeer et al. (2022) also reviewed the recent advances in the use of ML in water and wastewater treatment processes, including coagulation and flocculation. However, none of the studies reviewed model the kinetics of flocculation, instead focusing solely on coagulant dosage. The authors also emphasized a finding on the annual drift in the trend of ML applications in water treatment. This is consistent with the conclusion of Li et al. (2021) that both flocculation mechanisms (forces, etc.) and its kinetics (mixing intensity, concentration of particles, fluid mechanics etc.) have not been covered by existing ML reviews. Notably, ML and advanced DL applications in water sciences span across Image Analysis and Pattern Recognition, Predictive Modelling and Process Optimization, Experimental Design and Parameter Estimation, Numerical Simulations and Model Calibration, Anomaly Detection and Process Control, Data Integration and Multiscale Modelling, among others. Whereas ML and DL techniques with image analysis remain an emerging application in water treatment as it aids pollutants tracking, but they are rarely reported in the existing comprehensive reviews. This notable paucity of studies specifically addressing image-based flocculation modelling using ML highlights a pressing need for a focused analysis in this domain.

Current and emerging floc monitoring techniques predominantly rely on image acquisition, highlighting the importance of image analysis in understanding flocculation dynamics (Bankole et al., 2024b; Moruzzi et al., 2024). This is especially true as advanced vision computing, driven by DL, is transforming the way images are processed and analysed. Accurate floc image analysis is crucial, as mishandling these images can introduce bias into flocculation data, potentially leading to reduced treatment efficiency (Daus et al., 2024). DL has revolutionized computer vision applications in fields such as autonomous vehicles, medicine, hydrology, and remote sensing, and its application to flocculation could similarly enhance the modelling of floc kinetics, pollutants tracking (e.g. microplastics) and other aspects of water treatment. However, seamless and consistent data acquisition throughout the treatment process, which would enable comprehensive modelling of the entire system, remains a challenge, particularly at larger scales (Bankole et al., 2024a). Developing autonomous models for water treatment using DL techniques is poised to address these challenges and will likely attract significant research attention in the coming years. Chowdhury and Karanfil (2024) highlight that a DL framework for image recognition could pave way for advanced modelling of flocculation mechanisms, integrating factors such as floc properties, coagulant dosage, and treatment quality.

This review uniquely addresses these identified gaps by critically examining the image-based ML and advanced DL applications for modelling flocculation mechanisms and kinetics in water treatment and exploring its prospect towards automation. By focusing on this underrepresented domain, particularly leveraging advanced vision by DL, this study aims to enhance the understanding of floc dynamics and pollutant tracking, ultimately exploring the prospect of achieving automated and smart water treatment processes in line with industry 4.0 and 5.0 goals.

2. Flocculation kinetic modelling

Over the past decades, flocculation monitoring and modelling has been broadly classified into three themes, based on the review by Jarvis et al. (2005) and Watanabe (2017): (i) characterization of flocs physical properties (shape, strength and fractal nature); (ii) flocculation kinetics (rate and frequency of floc formation, and floc size distribution); and (iii) engineering aspects (peripheral speed of paddle and turbulence intensity). Of these three, flocculation kinetics is the most important because the prime objective of flocculation is to facilitate aggregate formation (particle collision), which alters particle size distribution, and to facilitate treatment separation efficiency (Abreu et al., 2025; Lawler, 1997; Moruzzi et al., 2024; Yukselen and Gregory, 2002).

The Smoluchowski (1917) model on discrete particle is the first numerical model upon which subsequent work has been based. This expressed the interaction of monodispersed particles during flocculation - Eq. (1).

$$\frac{dn_k}{dt} = \frac{1}{2}\sum_{i+j=k} H(i,j)n_i n_j - \sum_{i,j=1}^{\infty} H(i,j,k)n_{i,j}n_k \quad (1)$$

where $\frac{dn_k}{dt}$ is the variation of the number of particles k per time; i, j, and k are discrete particles sizes; $H(i, j)$ is the collision rate between the particles i and j [L^3T^{-1}]; $H(i, k)$ is the collision rate between the particles i and k [L^3T^{-1}]; n_i is the number of particles of size i; n_j is the number of particles of size j; n_k is the number of particles of size k.

This led to particle collision frequency model under perikinetic and orthokinetic flocculation (Eqs. (2) and (3)) as presented in Moruzzi et al. (2018), where the Boltzmann constant k is introduced; temperature (T) in Kelvin, viscosity – μ [$ML^{-1}T^{-1}$]; and $\frac{dv}{dy}$ is the velocity gradient in the fluid [T^{-1}]

$$H(i,j) = \left(\frac{2kT}{3\mu}\right)\frac{d_j}{d_i+1}(d_i+d_j) \quad (2)$$

$$H(i,j) = \frac{1}{6}\frac{dv}{dy}(d_i+d_j)^3 \quad (3)$$

The orthokinetic model (A&K model) developed by Argaman and Kaufman (1970) based on a multi-compartment pilot plant trial, established aggregation and breakage interaction of primary particles in water, under an induced shear gradient. This further established the relevance of the dynamic steady state in flocculation as the dynamic equilibrium state between aggregation and breakage. In summary, the velocity gradient's impact on floc growth as a function of flocculation time (T) was expressed according to Eq. (4).

$$\frac{n_o}{n_m} = \frac{\left(1+k_A G_f \frac{T}{m}\right)^m}{1+k_B G_f^{\,2}\frac{T}{m}\left[\sum_{i=0}^{m-1}\left(1+k_A G_f \frac{T}{m}\right)^i\right]} \quad (4)$$

where n_o is the concentration of primary particles; n_m is the concentration of primary particles in the m-th reactor in series; $\frac{n_o}{n_m}$ is the performance parameter, G_f is the velocity gradient, also known as the induced shear stress (s^{-1}); T is the total flocculation time for the pilot plant; k_A and k_B are the apparent aggregation and breakage constants; and m is the number of reactor in series. In this formulation G_f is assumed invariable throughout the m reactors in series.

The orthokinetic A&K model was modified by Bratby (2016) to accommodate instances where the turbidity (N_0) of water before flocculation (end of coagulation) is used as a surrogate measurement for primary particles in jar test batch assay, while the post-sedimentation turbidity (N_t) represents final floc concentration at the time t (Eq. (5)).

$$\frac{N_0}{N_t} = \left[\frac{k_B}{k_A}G_f + \left(1-\frac{k_B}{k_A}G_f\right)e^{-k_A G_f t}\right]^{-1} \quad (5)$$

where N_o is primary particle concentration; N_t is the particle concentration during flocculation. This equation is valid for jar tests only (batch essays), and the kinetic constants can be extrapolated to continuous flow using Eq. (4).

Marques and Filho (2017) and Marques and Filho (2022) incorporated irreversible floc breakage into the A&K model, asserting that post-sedimentation turbidity is caused by the inability of broken flocs to settle, while residual turbidity increased after certain T (considered as equilibrium by the authors). While using turbidity as a surrogate method for particle concentration is simple, it lacks precision because the turbidity of water after inertia and sedimentation transport mechanisms does not represent the kinetic energy at work during shear induced agitation, which depicts ideal flocculation conditions. Hence, the use of particle analysis during flocculation is considered as an improved technique by several authors (Bankole et al., 2024a; Chakraborti et al., 2000, 2003; Moruzzi et al., 2018, 2020a; Moruzzi and de Oliveira, 2013). The particle analysis approach also led to the development of empirical findings such as the particle–particle, and particle–cluster interactions established in the work of Moruzzi et al. (2017) and Bankole et al. (2024a). Through the particle analysis, other aggregate feature models were developed, such as the fractal evolution (He et al., 2012; Moruzzi et al., 2018; Zhang et al., 2021), fractal dimension (Liang et al., 2015; Moruzzi et al., 2017) and particle size distribution. For instance, the physical properties (area, volume and mass) and geometry of flocs are expressed by its fractal dimension, obtained by the scaling power law (Eq. (6)). This shows that aggregates with similar size may have different fractal dimensions (Moruzzi et al., 2018).

$$A \sim l^{D_f} \quad (6)$$

where A is the aggregate area (μm^2 or mm^2); l is the aggregate length (μm or mm); and D_f is the aggregate fractal dimension (dimensionless).

Moruzzi et al. (2024) introduced a single indicator β for monitoring aggregate heterodispersity and flocculation kinetics based on the aggregation and breakage constants (k_A and k_B) of Argaman and Kaufman. The implementation of the power law (β) displayed stability in modelling flocculation kinetics in batch and pilot reactors in series under varying gradient velocities, displaying a drift and simplification of the complex kinetics. Zhu (2022) also introduced the Budyko-type model's concept from catchment hydrology to model flocculation kinetics under turbulent conditions, leveraging on the mass balance principle. The study showed that Budyko-type model adoption for flocculation kinetic is reliable based on multiple data testing and calibration, with a very high accuracy (~ 97 %). Nonetheless, the A&K model is one of the most explored and remains relevant as it captures the mechanism of flocculation in a relatable manner (Moruzzi et al., 2024; Yukselen and Gregory, 2002). This assertion is reinforced through the summarized review of empirical contribution to flocculation modelling, as presented in Table 1.

Bankole et al. (2025a) developed an open-source tool that used swarm intelligence to automatically model flocculation kinetics using images acquired from NiDIA system. The main goal projected by the authors is to address the existing pitfall through inaccurate initial guesses of the least-squares method of the aggregation and breakage constants emphasised by Haarhoff et al. (1996) as a major concern in the A&K model implementation. The tool has been made available as a web application (flocxion-application) and extended to also simulate the total hydraulic retention time (THRT) for desired treatment efficiency, representing a recent significant contribution towards the transition to an autonomous treatment monitoring system, although the current version could only model retention under stable G_f conditions. The authors further open-sourced their code to promote further improvement on the modelling framework by future studies, as documented in Bankole et al. (2022a).

Based on the modelling concept described above, aggregate (floc) size and structure during flocculation process is the most important factor when considering treatment effectiveness (Bridgeman and Dapelo, 2025; Chakraborti et al., 2003; Gregory, 2009; Moruzzi et al., 2024), particularly when flocculation is combined with filtration (Park et al., 2006), sedimentation (Jin et al., 2003; Vahedi and Gorczyca, 2012) or membrane technique (Amjad et al., 2015; Park et al., 2006). This is to be expected as flocs with good settling capability enhance separation and reduce cake formation on membranes (Lopez-Exposito et al., 2019). Furthermore, the ability to determine the flocs' properties enhances the flocculation monitoring process and determination of optimal conditions for floc formation.

2.1. Flocculation image acquisition techniques

The flocs' physical properties are essential for flocculation kinetic modelling, and measurement accuracy influences subsequent analyses of the entire treatment process. Industrial applications of flocculation,

Table 1
Summary of empirical findings on flocculation kinetic modelling in water treatment system.

Source	Indicator	Type of Reactor system	Proposed / Evaluated Model
Argaman and Kaufman (1970)	Primary particles	Continuous stirred single compartment reactor.	Orthokinetic model for aggregation (K_A) and breakage (K_B) in continuous system.
Argaman (1971)	Water turbidity	Pilot plant reactor chambers in series.	Orthokinetic model.
Bratby et al. (1977)	Water turbidity	Continuous stirred single compartment reactor.	Implementation of batch Orthokinetic model.
Haarhoff and Joubert (1997)	Water turbidity	Continuous stirred single compartment reactor.	Sensitivity of orthokinetic constants to operational variables.
Bratby (1981)	Water turbidity	CSTR tanks in series.	Escalation of A&K model for pilot scale from batch mode assay.
Haarhoff et al. (1996)	Water turbidity	Continuous stirred tanks in series.	Experimented A&K model on full-scale plant and error susceptibility.
Moruzzi et al. (2017)	Primary particles measured by image analysis	Continuous stirred single compartment reactor.	Established particle – particle and particle-cluster interaction as dominant factor for dynamic equilibrium in A&K model.
Marques and Filho (2017, 2022)	Water turbidity	Continuous stirred single compartment reactor.	Incorporate irreversible breakage of flocs to A&K model.
Moruzzi et al. (2024)	Aggregate Size Distribution (ASD)	Continuous stirred single compartment reactor and Pilot plant chambers in series.	Implemented Power Law single parameter (β) into A&K model.
Bankole et al. (2025a)	Primary particles	Continuous stirred single tank reactor with simulation for multi-chambers reactor.	Developed open-source application for automatic modelling of A&K from images.

Note: The key results of aggregation and breakage constants for each finding can be accessed from the referred studies.

such as found in materials and pulp making industries, as well as other water treatment techniques besides the conventional treatment, facilitated the advancements of particle measurement technology. The general advancement in particle tracking influenced the development of a reliable and user-friendly strategy for flocculation monitoring. In the context of water and wastewater treatment, floc measurement primarily focuses on physical attributes (size, shape, density, porosity, etc.). Approaches to measuring floc size are broadly classified into intrusive and non-intrusive methods (Liang et al., 2015; Moruzzi et al., 2018). This includes (i) Laser Particle Size Analyser (LPSA); (ii) Focus Beam Reflectance Measurement (FRBM); (iii) Microscopy; (iv) Dynamic Image Analysis (DIA); and (v) Particle Vision Measurement (PVM).

The intrusive method is classified as measurement procedures that require interaction with the sample/reactor tank. While the non-intrusive method attempts to preserve aggregate forms through passive measurements. Generally, non-intrusive and some intrusive methods, such as DIA and FBRM, are considered *in-situ* techniques. Table 2 summarizes the distinct features, advantages, and limitations of these measuring techniques, while Fig. 1 depicts their pictorial configuration. Section S1 (1.1 – 1.5) of the supplementary files provide further description of these techniques in flocculation image acquisition.

Of all the techniques, DIA is the most adapted and widely used floc monitoring concept. Moruzzi et al. (2018) developed a non-intrusive dynamic image analysis (NiDIA) methodology to monitor aggregate morphological evolution during the flocculation and settling phases of water treatment (Moruzzi et al., 2020b). The proposed NiDIA framework is shown in Fig. 2. To establish the extent to which image analysis methods have been innovatively adopted, Pablo et al. (2022) used a Huawei P30 mobile phone with a 40 Megapixel camera to capture flocs images from a low turbidity and alkalinity batch assay, confirming that image-based floc monitoring process is reliable and provides low-cost approach to measuring aggregate properties.

Generally, floc measurements are acquired via images and analysed using image processing software or computing platforms such as MATLAB and Python. Observations with microscopy, FBRM, PVM and variants of DIA are made by acquiring images or videos and processing them using image processing software.

These tools are excellent for exploring aggregate features in two dimensions (2D), such as area, perimeter, diameter, elongation ratio, texture, and so on. Further manipulations are also explored during image preprocessing in order to obtain desired floc features. Fiji ImageJ (Pablo et al., 2022), National Institute of Health (NIH-image – Bethseda, USA) (Chakraborti et al., 2003) and Image ProPlus (Moruzzi et al., 2018; 2020a, 2020b; Bankole et al., 2024a) are among the commonly used flocculation image analysis software. Besides these, applicability of the floc data acquisition techniques, scale of operations and compatibility with the ML and DL models are highly imperative. A summary of these critical examination is presented in Table 3, to further elucidate the applicability, cost implications, type of extractable data type, the applicable scale of observations and feasibility of implementation with image-based ML and DL models.

The inability to account for the aggregates' third dimension is a major drawback in 2D image analysis for flocculation modelling, particularly in terms of fractal dimensions. Flocs are 3D objects, but image analysis only provides direct 2D measurements. Liu et al. (2022) critically reviewed the evolution of 2D and 3D imaging in particle size estimation and tracking, recognizing that recent advances in 3D estimations are limited. Researchers have often used the scaling power law or mass models, such as the Population Balance Model (PBM), to transform floc size data to mass and volume.

Among numerous studies, Maiti et al. (2017) developed a mass model for estimating mass-based particle size distribution using image analysis that achieved 99 % accuracy when compared to the traditional particle sieving method. Studies have reviewed mass-based particle size distribution models derived from image analysis (Wang, 2006; Zhang et al., 2012). Similarly, researchers have investigated aggregate features based on mass-volume relationships with PBM (Hounslow et al., 1988; Nazemzadeh et al., 2021; Seghir et al., 2022; Singh et al., 2022). PBM and related mass models are numerical concepts that are outside the scope of this current review. The next section discusses floc image processing and underlying factors that limit the performance effectiveness of these techniques in floc image analysis and how they influence analysis errors during floc feature extraction.

2.2. *Floc image processing and analysis error*

Image preprocessing includes, inter alia, procedures such as denoising, filtering, and thresholding, feature extraction, among others. Recent studies on flocculation models now process images using low-level computing platforms such as MATLAB and Python, utilizing computer vision features. This may be considered necessary because software-based image analysis often demand lots of time. Pablo et al. (2022) reported 30 – 40 min to process images captured from 15-minute flocculation experiments (5 images per observation) using the Image J software. This shows that the processing duration doubles the time required to acquire the images. In an instance where numerous images are acquired, such as the 21,500 images reported by Moruzzi et al. (2018), processing time will undoubtedly increase. While the rendering is fast, most of the time is spent on post-processing tasks such as sorting

Table 2
Summary of flocculation particle size measuring techniques. Adapted after Liang et al. (2015).

Measurement technique	Particle size definition	Operational range (μm)	% Concentration	Advantages	Disadvantages
LPSA	Spherical equivalent diameter	0.01 – 3500	< 0.1	• Rapid PSD measurement and analytical procedures.	• Breakage of fragile floc by ultrasonic and circulating or suction pump. • Provides bias floc size of irregular shaped aggregates.
FBRM	Chord length	1 – 2000	< 20	• Real-time *in-situ* floc evaluation with high precision. • Particle evolution monitoring. • Irregular particle size measurement.	• Unreliable under high turbidity with low shear stress conditions. • Probe fouling (sapphire window blockade) due to large and dense aggregates, leading to measurement error. • Energy distribution alteration around the probe, causing particle fragmentation.
Microscopy	Geometric equivalent diameter	0.001 – 10,000	< 0.5	• Particle size measurement up to nano scale. • Offers hybrid implementation with high-speed camera. • Offers comprehensive aggregate analysis including taxonomical characterization	• Most microscopes require sampling, leading to alteration in floc features. • Process requires multiple sample collection to obtain representative data.
DIA	Geometric equivalent diameter	1 - 5000	< 0.5	• Offers automatic and continuous aggregate feature measurement. • Highly scalable, providing low-cost particle monitoring system. • High precision and accuracy. • Capacity to acquire high volume image under limited time.	• Aggregate measurements are restricted by camera lens properties (pixel size, resolution etc.).
PVM	Geometric equivalent diameter	0.5 – 2000	< 0.5	• Real-time aggregate measurement and particle tracking capability via video recording.	• Low resolution of aggregates smaller than 20 μm. • Poor performance under high concentration suspension.

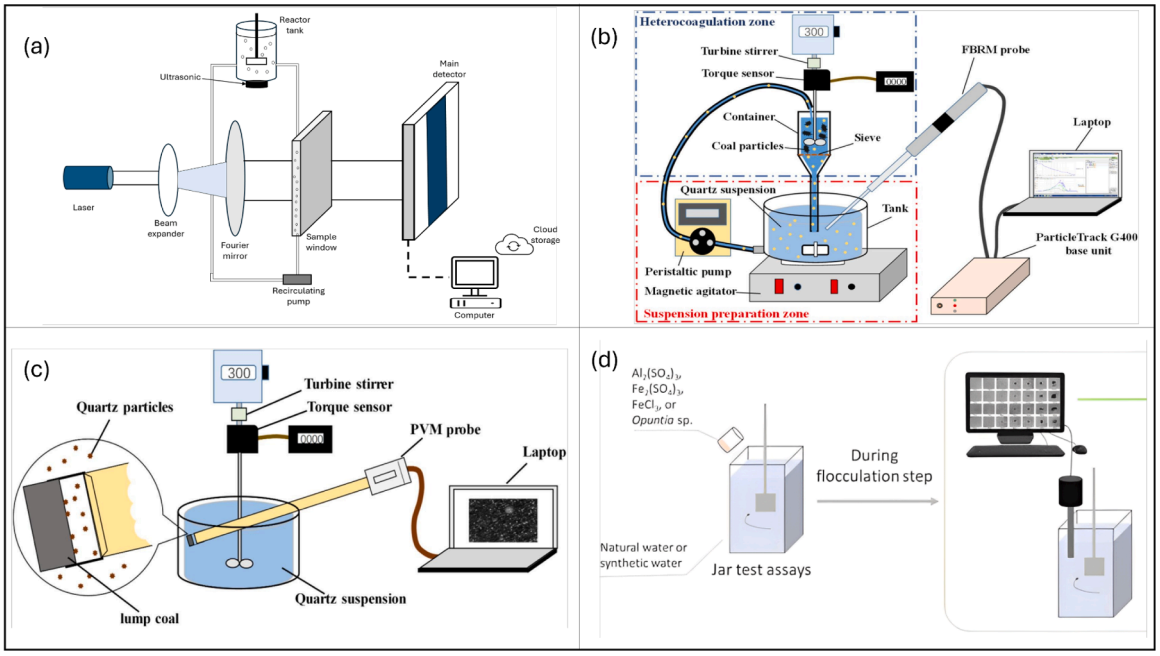


Fig. 1. Schematic diagram of particle size measurement using (a) laser particle size analyser; (b) FBRM (Hu et al., 2021); (c) PVM (Hu et al., 2021); (d) *In-situ* Microscopy set up (Santos Nunes et al., 2022).

and data verification. This limits the use of software for real-time flocculation modelling, demanding the need for seamless image analysis approaches, especially when ML is being implemented for efficient treatment optimisation.

Image processing on computing platforms is advantageous in many ways. Most software operations can be performed with a few lines of code, providing the ability to model directly from image acquisition process, and allowing for trial and error in the case of sensitivity analysis. Flocculation image processed with software are often performed through thresholding technique (Chakraborti et al., 2000, 2003; Moruzzi and Reali, 2010; Moruzzi et al., 2017, 2018, 2019, 2020b), a conventional semantic segmentation method. Meanwhile, modelling platforms allow access to open-source libraries containing enhanced semantic segmentation algorithms such as those based on ML and DL concepts. Chong et al. (2023) argued that the capacity to build a pipeline of computing algorithms for automatic image analysis and modelling is responsible for the recent adoption of intelligent vision and ML across different fields.

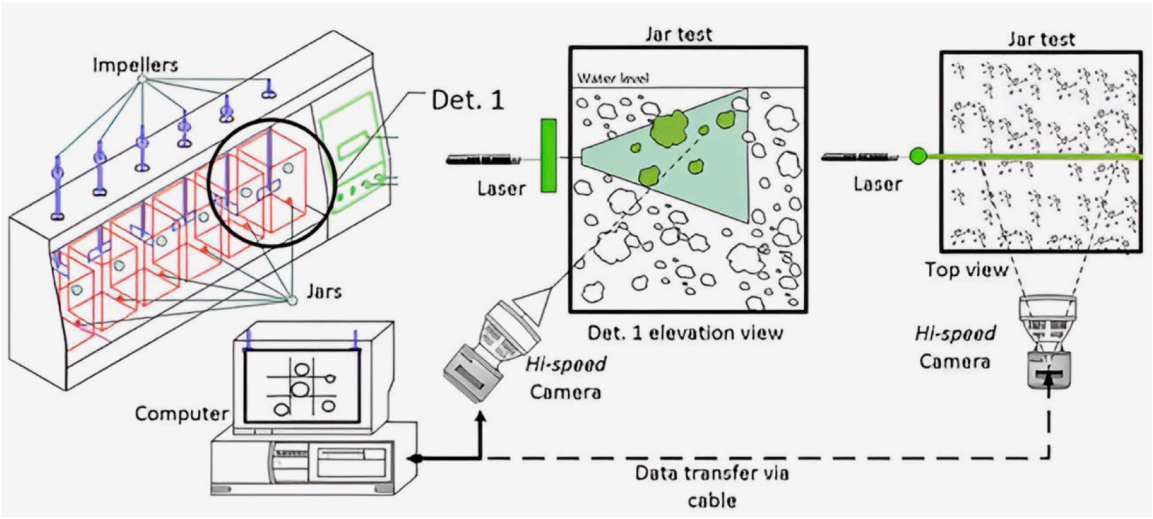


Fig. 2. Schematic framework of non-intrusive dynamic image analysis (Moruzzi et al., 2017, 2020b).

Table 3
Summary of compatibility analysis of floc data acquisition techniques with image-based ML and DL models, and real-time data acquisition and potential for modelling.

Flocs particle size measurement technique	Can the method generate floc image data alone?	Cost	Scale of observation (Pilot, Lab, or large-scale) aside from full-scale)	Compatibility with real-time monitoring and modelling	Compatibility with image-based ML and DL model and model type?	Source
LPSA	1. Generate only particle size, shape etc. in numerical values when used alone. 2. Technique can only acquire images when coupled with camera device.	$80,000 – $250,000. Expensive. Require maintains.	Suitable for controlled laboratory-based studies.	Not compatible	Not compatible.	Bao et al. (2018); Govoreanu et al. (2009)
FBRM	1. Produced chord length/floc size, counts, and shape. 2. Offers image acquisition when combined with PVM.	$50,000 – $150,000.	Suitable for controlled laboratory-based studies.	Suitable for real-time and *in-situ* monitoring and modelling.	Not compatible.	Hu et al. (2021); Muhaimin et al. (2021); Sankaranarayanan et al. (2019)
Microscopy	Very suitable image acquisition.	$20,000 – $50,000. Require maintenance and relatively expensive depending on the technique, floc size of interest and scale.	Full-scale (Pilot and Laboratory scale) treatment	Suitable for real-time, *in-situ*, and automated monitoring and modelling	Compatible with both types of models.	Santos Nunes et al. (2022); Xu et al. (2024)
DIA	Highly suitable for image data acquisition.	$100,000 – $250,000+Adaptable for simple camera devices like smartphones.	Suitable for all scales flocculation treatment	Suitable for real-time modelling and automated modelling	Compatible with both types of models.	Fukasawa et al. (2024); Pablo et al. (2022); Sun et al. (2016)
PVM	Very suitable for acquiring floc images.	$70,000–$180,000	Most suitable for laboratory and pilot scale.	Suitable for real-time, on-line, and *in-situ* monitoring.	Compatible with both types of models.	Hu et al. (2021)

NB: The cost included are subjective to variations caused by market value, model preference, and buyer-seller agreements. The quoted prices are verifiable at: LPSA; FBRM; Microscopy; DIA and PVM.

Currently, there are limited studies on flocculation modelling that incorporate ML imaging techniques. Yuan et al. (2016) presented an intelligent vision and Fuzzy Control system for flocculation monitoring in waterworks, with the goal of achieving an autonomous dosing system. During the flocculation process, the author captured floc images with an underwater camera. The images were grayscale and transformed before segmenting with an automatic thresholding algorithm. Baum et al. (2023) pre-processed images generated by an automatic microscope with morphological reconstruction by erosion before thresholding. The flocs' features (size and fractal geometry) were analysed using a multinomial logistic regression solver and compared to the performance of a fully connected convolutional neural network (CNN). Li et al. (2020) employed a charge-coupled device (CCD camera) to capture flocculation images from a batch experiment of papermaking effluent with poly-aluminium flocculants. Prior to binarization, the images were filtered, followed by edge detection. The study also showed that images captured in the sedimentation column and analysed for floc properties can be used to automatically estimate aggregate settling velocity. This also

demonstrates that the use of a computing platform in floc monitoring transcends the flocculation process and may enable the automation of the entire water treatment process. It is also important to note that the highlighted gaps in computing image analysis are mainly attributable to seamless workflow and analysis that computing platforms offer, compared to software.

Despite the highlighted benefits of flocculation image processing, the importance of capturing quality images during acquisition is of high necessity. Factors such as lighting condition, shutter speed, image resolution, camera positioning and depth of field etc., all determine image quality and suitable preprocessing techniques. Although these factors are often overlooked/poorly discussed or associated with analysis errors. Nonetheless, they remain a major consideration in image-based floc measurement and monitoring in water treatment. For instance, bright-field illumination is essential for capturing images with uniform light intensity (Sun et al., 2016), which could aid the extraction of floc image textural features. It is important to note that studies has recently found floc image texture to be resourceful in modelling flocculation kinetics, particularly particle aggregation behaviour (Ma et al., 2023). But no study has established the sensitivity of image processing to floc texture measurements. Also, determination of optimal shutter speed to ensure sufficient light exposure before capturing image is essential to minimize motion blur that could cause the flocs to have fuzzy edges, further leading to overestimation of floc physical features. Sun et al. (2016) in their flocculation experiment found 300 μs to be optimal shutter speed for aggregate travel distance of 1 μm and reported that kernel filtering aid to avoid out of focus particles captured during imaging. Higher shutter speed have been reported in recent studies, such as Moruzzi et al. (2024) that reported 800 μs with 10 Hz frequency (10 images captured per second) and Li et al. (2024b) that reported 4000 μs and frequency of 100 Hz.

The camera lens' depth of field is another essential factor that determines the effectiveness of imaging technique in flocculation and water treatment. Flocculation images contain flocs within and outside the camera lens depth of field (focus), which often makes segmentation error analysis essential. This essential feature (depth of field) is affected by the turbidity of water (Nelli et al., 2024), particularly in high turbid water or assay with high concentration of flocs. For instance, ability to penetrate the water surface to satisfactory depth by the lens and the illumination sources can be impacted in high turbid water compared to lower turbidity. Sun et al. (2016) analysed the impact of turbidity on camera lens depth of field and the field of view, and its effect on floc image analysis. A linear relationship was found between concentrations of clay particles and turbidity of water, while depth of field value of 500 $\pm$ 90 μm was found appropriate for up to 60 NTU turbidity.

Another essential factor is the reduction of artefacts or bubbles in the reactor, particularly those attached to the wall of the reactors (Pablo et al., 2022). Although limited control exists on the avoidance of bubbles in large-scale real-time treatment monitoring. Still, imaging technique has been used even in flotation treatment techniques that require the presence of bubbles for efficient treatment (Moruzzi and Reali, 2010). Regardless, the adverse effect of noise or artefacts created by bubbles on floc image analysis cannot be overruled. Image resolution is essential as it determines the probable minimum floc size that can be captured by imaging device. Low resolution images are more susceptible to analysis error for smaller flocs due to over estimation (particularly floc size) as argued by Chakraborti et al. (2003). Analysis error is recognized as a significant constraint of image processing, resulting in over or under estimation of floc features. Moruzzi et al. (2018, 2020a) among other authors also argued that estimation of image analysis error offer moderation of image preprocessing technique accuracy. Although Li et al. (2024a) argued that analysis of smaller aggregates attracts error irrespective of the image resolution. However, there is currently no study that has established the sensitivity of high-resolution floc images to error analysis in recent times, particularly with very high-resolution images. The image analysis error estimates give recommendations on the extent to which preprocessing can be done during segmentation, hence controlling the disparities between the ideal flocs and estimated error (Moruzzi et al., 2018).

The influence of floc image acquisition considerations (lighting, shutter speed, resolution etc.) on modelling performance is a delicate discussion as limited studies exists on these. Segmenting floc images with blurry or fuzzy boundaries has been reported in studies with an outstanding accuracy, but report has not been given on how such modelling influences the flocculation kinetics indicators such as the K_A and K_B values per conditions or other kinetic indicators. For instance, high segmentation results can be observed by the model but further evaluation on how they affect the accuracy of the ideal kinetics need to be estimated. While the detection of out of focus objects can easily be identified during ground truthing, responses of DL models in segmenting such floc images have not been reported, whereas such scenarios are envisaged in full-scale implementation where an underwater camera may face constrains from illumination in an outdoor environment.

Concerning the impact of image resolution on model performance, Al-Ani et al. (2025) reported in their study that evaluated the effect of sample size and resolution of microscopic microalgal images on DL model performance. The simple fully connected CNN architecture showed that lower image resolution affects segmentation accuracy, while images with an average resolution of 788 $\times$ 530 recorded an outstanding performance, based on their evaluation. This, among other studies, further showed that pixel density influences model performance, which is envisaged since DL models learn better at the presence of quality data. Although, these studies are not on flocculation kinetics and nor did they focus on water treatment floc images, our aim here is to draw correlation on the potential impact of image resolution on the efficient segmentation of floc images as argued by Li et al. (2024a), and downstream model performance, as recently proposed by Bankole et al. (2025a). Therefore, future studies are strongly encouraged to explore impact of image resolution on DL model segmentation performance and the sensitivity analysis of downstream flocculation kinetic computation (preferably with ML as well) to bridge this significant gap.

The following section discusses the ideal image preprocessing and image reconstruction functions available on most image processing computing platforms, using the Python open-source library (Open CV (Bradski, 2000) and Scikit Image (Pedregosa et al., 2018)) as an example.

3. Semantic segmentation in flocculation analysis

The ultimate purpose of image processing is to improve the quality of the image and the accuracy of the retrieved information (Maheshan and Prasanna Kumar, 2019) or to assist algorithms in learning from the identifiable pattern embedded in the image, in the case of ML and DL. Semantic segmentation is the pixel-based classification of images to identify the target (object) and background. Raw images are often associated with disturbances that contain redundant information, i.e. noise. They may also include missing or incorrect values (Chatterjee et al., 2016). These images are often processed at the pixel level with vision intelligence in order to preserve the embedded important information via the semantic segmentation process. Image enhancement techniques include denoising, convolving, reconstruction, feature extraction, and tracking. The following section summarizes the primary image analysis approaches based on open-source intelligent vision found in the literature. Gray-scaling converting techniques are not discussed in this section because most image-based floc investigation techniques used in water treatment often acquire images in grayscale. Nonetheless, readers can consult Chong et al. (2023) for in-depth description of color images and image gray scaling.

3.1. Image denoising

Different images contain various types of noise, including Gaussian,

salt-and-pepper, Poison and Perlin noise (Tawfik et al., 2022). Floc images from water and wastewater treatment systems frequently include noises created by shear induced bubbles or the flow of water itself. Furthermore, any out of focus flocs captured within the image windows are unneeded information that must be precisely removed in order to improve the image quality and its correctness, i.e. the information embedded in each pixel. This makes the image denoising a critical step in floc image analysis. Smoothing, or the application of filters, is commonly used to eliminate or minimize noise. Image filters are an incredibly powerful tool for handling noise, but they must be used cautiously. Filters are often characterized as high-pass or low-pass filters based on the intensity with which they affect image pixels (Bradski, 2000; Pedregosa et al., 2018). In addition, filters frequently smooth the image at the expense of the embedded object's edges, or floc edges in this case (Tawfik et al., 2022). Mean, median, gaussian, bilateral, and non-local means filters are some of the most common image filters accessible on computing platforms. Additionally, the filters could be classed as linear or non-linear. Mean and gaussian filters are often used as linear filters, although median and bilateral filters are non-linear in principle.

Gaussian filters use a gaussian function to convolve an image, resulting in a blurrier image. This is because the kernel only analyses the space beneath it, without regard for pixel intensity. The gaussian filter is a low-pass filter that applies to images having gaussian noise. Similarly, the median filter is another low pass filter that smooths images by replacing the gray scale values of each pixel in the kernel region with the median value of neighbouring pixels. Median blurring is highly efficient at filtering salt-and-pepper noise. The mean filter, (or average smoothing), uses a definitive normalized kernel to average pixels in the same region. The bilateral filter is similar to the gaussian filter in that it uses a kernel, but it outperforms the gaussian filter in terms of pixel intensity beneath the kernel area. This increases its uniqueness and usefulness for denoising images with objects' edge preservation. Peng et al. (2024) recently studied the impact of five types of filters (Line transform, Histogram equalization, mean filter, median, and bilateral filter) on camera images obtained from a flocculation–precipitation experiment. The study found the bilateral filter to outperform others that fail to effectively denoise the images. It also emphasized the importance of carefully inspecting the impact of filters on flocs' texture and analysis accuracy, to aid proper usage of the filters on floc images. Buades et al. (2011) introduced the non-local mean filter, which likewise assumes that the image noise is a white noise. The target pixels' value is estimated using the weighted average of neighbouring pixel. Fig. 3a shows images of different filters on a floc image acquired by NiDIA under flocculation condition ($G_f = 20\ s^{-1}$ and $T = 15$ min) with pixel size of 270 μm.

3.2. Edge detection analysis

Edges of flocs or objects in an image are continuous pixels with comparable intensity/values that separate two different parts of the image. The ultimate idea of edge segmentation is to define the object's boundaries and segment the regions contained within them. Particles in the image may have an extremely blur boundary, as evidenced by multiple pixels at the object's border with nearly identical pixel values. The edge detection technique is also known as edge-based segmentation (Zhao et al., 2024). Generally, every grayscale-based image segmentation method begins with plotting the histogram distribution of the image pixels in order to gain insight into both the pixel distribution and prospective processing approaches. Edge detector operators include Sobel, Canny, Prewitt, Scharr, and Roberts (Liu et al., 2022; Rogowska 2009). These operators use differentiation operations to recognize the edges of objects, taking into account any discontinuity in pixel values (Chaki and Dey, 2018). In the case of Sobel, a gaussian filter is often applied prior to edge detection to reduce noise and produce smoother, more defined edges. Fig. 3b presents the output of various edge-based segmentations of floc images. It is worth noting that edge segmentation can sometimes overestimate the size of smaller particles, particularly when edges are blurred and noise is present, making it sometimes unsuitable for floc image analysis (Giering et al., 2020).

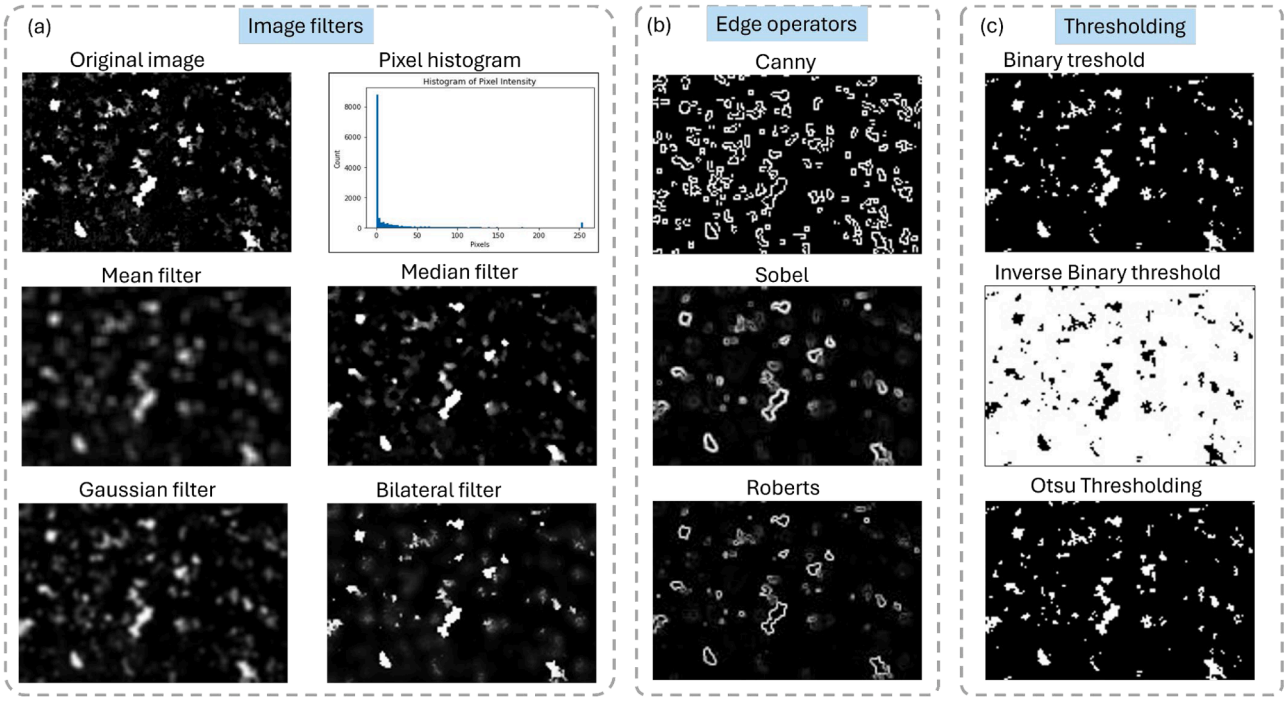


Fig. 3. NiDIA floc image from Moruzzi et al. (2018) under flocculation conditions ($G_f = 20$ s-1; $T = 15$ min; pixel size = 270 μm) processed with (a) different filter operators; (b) edge operator; (c) different threshold operations. The filters and threshold were carried out using OpenCV, edge operators on Scikit image were utilized for Roberts and Sobel edge operators.

3.3. Thresholding

Thresholding is the most basic classical semantic segmentation method. There are various thresholding approaches that follow similar concepts and has been well-explored in flocculation analysis. The goal of thresholding is to divide an image into two distinct classes: background and foreground (flocs) depending on pixel values (Rogowska, 2009). Some techniques use the image pixel histogram, or the pixel value intensity, while others use the pixels' local mean and standard deviation. These techniques are generally classified as local and global thresholding, but their operations include the histogram-based, mean, adaptive, and global thresholding. The ease of computing and the fact that thresholding is less computationally expensive has facilitated its use in digital image processing, including flocculation and sedimentation image processing.

Global thresholding is a technique for handling images that have a bimodal pixel histogram, where the foreground and background are distinct. The global threshold divides the object (foreground) from the background by considering the complete image (Rogowska, 2009). It is a very basic and fast thresholding technique that transforms the input image using a predefined threshold value based on the pixel histogram or knowledge expert. Eq. (7) provides a mathematical expression for global thresholding, in which each pixel (x,y) in the image $f(x,y)$ are classed as 1 and the pixels for the background in the image are classified as 0, resulting the binary image $g(x,y)$.

$$g(x,y) = \begin{cases} 1, f(x,y) \geq T \\ 0, f(x,y) < T \end{cases} \tag{7}$$

The Otsu algorithm is a popular type of global thresholding used in digital image processing to automatically determine the appropriate threshold value to segment/binarize the image. Notably, recent flocculation research such as those by Yuan et al. (2016), Baum et al. (2023), and Peng et al. (2024), have explored and compared the performance of Otsu algorithm in flocs images segmentation to other operators. Baum et al. (2023) reported that the Otsu algorithm performed poorly when handling floc images that had been pre-processed using a 5-pixel erosion disk, citing the images' low dynamic range. Peng et al. (2024) discovered that the Otsu algorithm outperformed fixed thresholding in their experiment; however, the authors emphasized that such a technique may be unable to cope with certain floc images, particularly when a cloudy water is assessed, which could result in significant overlap between the image foreground and background. Additionally, Chong et al. (2023) found that the Otsu algorithm underperforms when thresholding images with poor contrast, noise and uneven light intensity. This demonstrates that Otsu or global thresholding algorithms may not effectively handle certain floc images, particularly those with a high number of primary particles. It also emphasized the need to carefully select the appropriate segmentation strategy when analysing flocculation images.

The local thresholding segment images based on the sub-regional technique. Before computing the optimal threshold per sub-image, the image is partitioned into overlapping sub-image or the image intensities within the neighbourhood of individual pixel per sub-images (Rogowska, 2009). Local threshold, also known as adaptive threshold, seeks to solve the constraints of global threshold by focusing on images at sub-level prior to the second interpolation of pixel threshold values. Local thresholding is particularly effective for images with scattered and small objects, and images with inconsistent backgrounds/uneven light intensities (Chong et al., 2023). These descriptions support the findings of Peng et al. (2024), who found that adaptive thresholding outperformed the Otsu, histogram, iterative and fixed thresholding techniques.

Considering that the recent trend in flocculation modelling and image analysis is to achieve automatic segmentation process and seamless application of flocculation data using ML algorithms, it is considered essential to investigate the feasibility of modifying thresholding for flocculation image analysis. While this may be difficult due to the high variability of water and wastewater characteristics, as well as coagulant type and dosage, investigating the responsiveness of different thresholding techniques to various types of water, coagulant type and dosage in relation to flocs properties are considered worthy of consideration. Rogowska (2009) proposed that identifying the sensitivity of segmentation techniques to changes in flocs area and perimeter is a reliable approach for determining the appropriate threshold. Fig. 3c shows the output of various threshold techniques applied to a floc image.

3.4. Morphological operations

The proper handling of flocculation digital images is critical to obtaining a valid result. Although, floc properties such as size, shape and fractal dimension are essential features used in many flocculation models. Extracting accurate morphological floc features from 2D images could help understand the nature of the flocs generated and the impact of the flocculation conditions (G_f, T and coagulant effect). For instance, floc porosity is critical for understanding flocs' resistance to shear stress and modelling aggregate settling velocity (Moruzzi et al., 2020a). Erosion, dilation, opening and closing are examples of general morphological operators, and the watershed image segmentation algorithm could be used in reconstruction. Erosion and dilation perform opposite functions; erosion shrinks the boundaries of the target objects (flocs), removing the pixel at the edges, while dilation works by filing the edges of the object and creating small holes within the flocs. Opening is a process that combines erosion followed by dilation to smooth edges, remove small undesirable noise, and break unwanted connections between two or more objects while preserving the overall object structure. Also, closing is the inverse of opening (dilation followed by erosion). It is an effective operator for smoothing and repairing broken objects while preserving their overall shape, as well as filling minor gaps (Lerallut et al., 2007).

Baum et al. (2023) used an erosion morphological operator with a 5-pixel wide disk size on floc images acquired from an automated microscope before applying Otsu thresholding. Cardona et al. (2018) modified particle image obtained from a particle size and shape experiment using PVM by closing with 8-pixel disk radius. It is important to highlight that morphological operators are not often employed to treat flocculation images, especially when specific floc textural features are of interest. This is because image analysis error is typically higher for smaller aggregates (Chakraborti et al., 2003; Moruzzi et al., 2017, 2018) and these operators may induce their disappearance, affecting the accuracy of the result. Although studies have not established the extent to which these operators cause errors in floc textural features, which reflects a significant gap that needs to be explored for proper guidance on their usage for floc image analysis. Nonetheless, they could be used, particularly when floc size and shape are the variables of interest.

3.5. Feature extraction

The acquisition of detailed floc properties is critical in digital image analysis. Vision intelligence algorithms have made it possible to extract a wide range of floc attributes directly. Ideally, the connected component is established when the thresholding or other segmentation method is used (Baum et al., 2023; Rogowska, 2009). Specific function extracts aggregate properties such as equivalent diameter, area, perimeter, major and minor axis lengths, porosity, elongation, orientation, eccentricity, entropy, energy, correlation, rectangular and ellipsoidal aspect ratios, among others. These features can be process for direct modelling or relevant analysis.

Traditional ML and emerging DL techniques for image-based flocculation modelling are presented in Sections 4 and 5. The discussion offers best practices in semantic segmentation using ML and DL models from diverse domains, with potential application in flocculation

modelling, as this is an emerging application in water treatment modelling.

4. Traditional machine learning semantic segmentation

Semantic segmentation techniques using ML are applicable to the majority of classification algorithms, although with certain exceptions. In general, ML algorithms for semantic segmentation are divided into two types: traditional ML and DL algorithms. Random Forest (RF) and its variants (Extreme Gradient Boosting, for example), Support Vector Machine (SVM), and K-Nearest Neighbour (KNN) are examples of traditional ML algorithms. These algorithms are robust and generate generalized classification and regression models. These have influenced their use in modelling various aspects of water and wastewater treatment (Arab et al., 2022; Bressane et al., 2023; Chong et al., 2024; Wang et al., 2022a).

The distinct benefit of ML models for segmentation is the ability to compare a wide range of image attributes to the annotated image in order to efficiently predict the ideal pixel value of the image, improving the object and background classification accuracy. In the case of flocculation images, several semantic segmentation operators (e.g. gaussian filter, Otsu, numerous manual threshold values, edge detection, etc.) could be combined to provide a dataset for training ML models, which could result in a more generalized model. Furthermore, the variable of importance could be ranked, providing more precise information on the impact of individual operators. Since the model learns from diverse datasets and the labelled images, this approach may help reduce the time required for conducting sensitivity or error analysis on the images.

Nonetheless, these techniques are seldomly employed for segmenting floc images. This might be attributed primarily to the reported drift to increased computing power of DL since the release of the convolutional neural network model (Mo et al., 2022). ML algorithms require labelled data (referred as masks) for training, which is an enormous task that requires expert's knowledge and time consuming. This is a general disadvantage of ML and DL implementation in image segmentation. As a result, ML models are rarely used in recent flocculation or floc image modelling research, although several studies have used them in modelling water treatment processes and water quality (Dongsheng et al., 2017; Wang, 2016; Wang et al., 2022a). Mo et al. (2022) supported this assumption, arguing that recent high computational power of DL algorithms are now the mainstream solutions for semantic segmentation tasks. Nonetheless, ML models may achieve high accuracy with a moderate dataset and lesser computational power than DL algorithms, which are often trained with large datasets. Further readings on the operational technicalities of major traditional ML algorithm for semantic segmentation are presented in section S2. The next section presents best practices in semantic segmentation using DL algorithms, as well as their potential for developing a generalized floc image segmentation model.

5. Deep learning semantic segmentation algorithms

DL-based semantic segmentation algorithms have proven relevant in digital image processing. The recent studies on flocculation imaging and water treatment modelling, though very limited, have demonstrated that advanced intelligence algorithms are critical for optimising treatment processes. Few studies have so far used DL algorithms, such as MaskRCNN (Daus et al., 2024) and DeepLab v3+ (Fukasawa et al., 2024), to model flocculation processes/floc images in water treatment and sludge dewatering domains. Additionally, Yokoyama et al. (2024) concluded that DL-based semantic segmentation model could facilitate real-time automation in their study that deployed a Vision Transformer (ConvNeXt) for predicting degree of flocculation (DF) by automatically controlling the flocculant dosage. These are significant and landmark efforts towards image-based DL model application to flocculation and system automation. Nonetheless, several challenges like the choice of DL algorithm, data preprocessing/ground truthing, algorithm computational complexities, and so on, still limit the use of DL for flocculation image segmentation and modelling.

Since the introduction of Jonathan Long's fully convolutional network (FCN) algorithm (Long et al., 2015; Shelhamer et al., 2017), the development of convolutional neural network algorithms (CNN) for semantic segmentation tasks has been revolutionized. The fully connected and end-to-end trainable network can learn from uninformed size images and provide highly precise maps of similar sizes. As a result, numerous DL semantic segmentation algorithms have been developed. These included the popular DeepLab and its versions (Chen et al., 2017; 2018a, 2018b), U-Net (Diakogiannis et al., 2020; Ronneberger et al., 2015), SegNet (Badrinarayanan et al., 2017), and Pyramid Scene Parsing Network – PSPNet (Zhao et al., 2017), among others. Currently, there is an exhaustive collection of hybrid models derived from DL semantic segmentation algorithms using feature engineering and transfer learning approaches. Semantic segmentation methods are integrated with classical image classification models, such as the VGG16, GoogLeNet, and ResNet, to name a few. Table 4 presents a summary of major DL semantic segmentation algorithms. It is important to clarify that this is an inexhaustive list of pretrained models, and only the major releases are captured.

Of all these algorithms, U-Net is one of the most widely used, due to ease of manipulating architectural designs and exceptional accuracy, as proven by ResUNet (Diakogiannis et al., 2020), UNet++ (Zhou et al., 2020), FG-Unet (Stoian et al., 2019) and others. The U-Net, a fully connected encoder–decoder network with skip links, is named for its U shape (Fig. 4a). The encoding layer consists of 3×3 convolutional filters

Table 4
Summary of State-of-the-art DL-based semantic segmentation algorithms. Adapted from Mo et al. (2022)and Singh and Rani (2020).

Year	Algorithm	Accuracy (%)	Domain
2015	FCN	62.2	Street scenes segmentation
	U-Net	77.5	Biomedical Image
	DeconvNet	72.5	Street scenes segmentation
	SegNet	57.0 - 79.1	Street scenes segmentation
	Deeplab V1	71.6 (IoU)	Street scenes segmentation
2016	Deeplab v3	81.3 and 85.7 (mIoU)	Scene segmentation
	Dilated Convolutions	75.3 (mIoU)	Street scenes segmentation
2017	PSPNet	85.4	Street scenes segmentation
	RefineNet	83.4	Person segmentation
	Deeplab v2	79.7 (mIoU)	Scene segmentation
	Large Kernel Matters	82.2	Street scenes segmentation
2018	ERFNet	68.0 - 86.5	Autonomous vehicle
	Deeplab v3+	82.1 and 89	Scene segmentation
	DUC	80.1 (mIoU)	Street scenes segmentation
	ICNET	69.5 (mIoU)	Street scenes segmentation
	BiSeNet	68.4 (mIoU)	Street scenes segmentation
	AdaptSegNet	65.1 (mIoU)	Street scenes segmentation
	EncNet	85.9 (mIoU)	Street scenes segmentation
2021	Mseg*	34.19 - 80.7	Multi-domain segmentation
2023	Segment Anything Model (SAM)	> 80 (mIoU)	Multi-domain segmentation

The asterix (*) implies that datasets are combination of scene images, particles, (model validation) among others. Detailed list can be found in the original article on Mseg by Lambert et al. (2021).

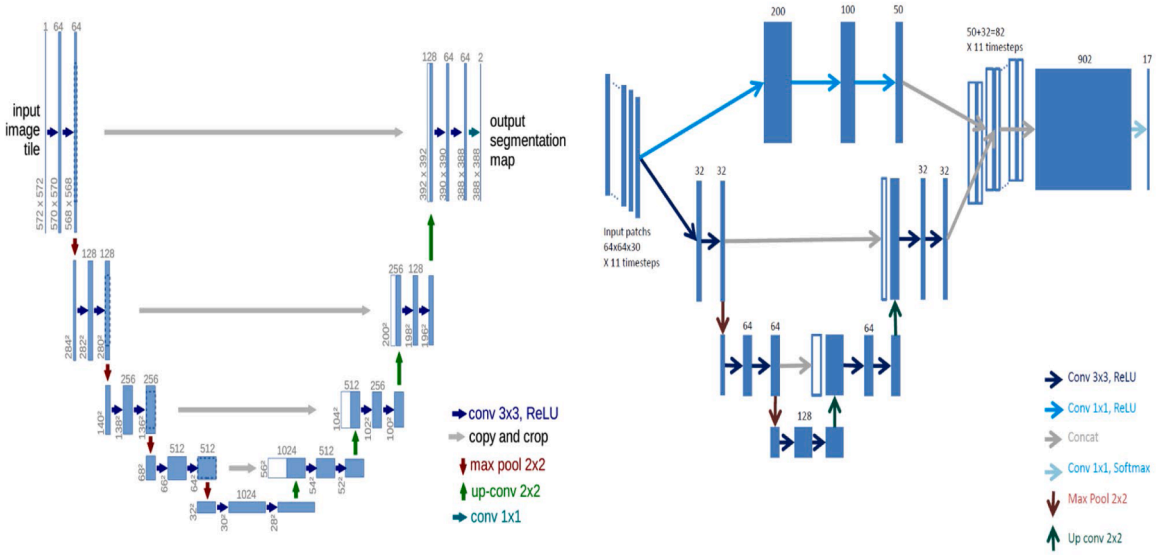


Fig. 4. Schematic description of (a) U-Net (Ronneberger et al., 2015) and (b) FG-UNet architecture (Stoian et al., 2019).

with a dropout layer, connected to successive layer using max pooling (Zhao and Zhang, 2024). Each encoding layer is concatenated to the corresponding decoding layer. The algorithm was originally designed to segment medical images (Ronneberger et al., 2015), but it has also successfully segmented digital images from other fields, including particle images (Shi et al. 2022), making it a potential fit for image-based flocculation modelling. Fig. 4a shows the original U-Net design with an input image resolution of 572 × 572. Fig. 4b depicts FG-UNET, an adapted U-Net with a resolution of 64 × 64 (Stoian et al. 2019). In general, the difficulty of ground truthing images limits the application of DL in semantic segmentation (Liu, 2024). Ground truthing (manual image labelling – masks) is time-consuming and requires expert knowledge, making it a major bottleneck in semantic segmentation with DL. To address this challenge, recent research has investigated novel methodologies using weakly supervised semantic segmentation techniques.

5.1. *Evaluation metrics for semantic segmentation*

To quantitatively assess the performance of semantic segmentation models, several key metrics are commonly employed. These metrics provide objective measures of how well a model can accurately delineate objects within an image. The Mean Intersection over Union (mIoU), also known as the Jaccard index, is a widely used metric that quantifies the overlap between the predicted segmentation mask and the ground truth mask (van Beers, 2018). It is calculated by dividing the intersection of the predicted and ground truth areas by their union, and this calculation is then averaged over all classes present in the image (or, in binary segmentation, over foreground and background). A higher mIoU score indicates a greater degree of spatial overlap and, consequently, better segmentation accuracy. The F1-score provides a balanced measure of a model's precision and recall. Precision (also known as Positive Predictive Value) measures the proportion of correctly identified positive pixels (e.g., floc pixels) out of all pixels that the model predicted as positive. Recall (also known as Sensitivity or True Positive Rate) measures the proportion of correctly identified positive pixels out of all actual positive pixels in the ground truth (Rizwan I Haque and Neubert, 2020). The F1-score is the harmonic mean of precision and recall, calculated as: F1-score = 2 * (Precision * Recall) / (Precision + Recall) (Rizwan I Haque and Neubert, 2020). This metric is particularly valuable in scenarios where class imbalance might skew other metrics, as it provides a robust assessment of segmentation performance by considering both false positives and false negatives. For a detailed reading on the evaluation metrics applied to semantic segmentation, kindly refer to the review conducted by Rizwan I Haque and Neubert (2020).

5.2. *Weakly supervised semantic segmentation*

Appropriate image annotation or data labelling (masks) are required for training DL model. Its accuracy and quality determine the model's segmentation and generalization capabilities (Liu, 2024; Mo et al., 2022). One of the earlier challenges of semantic segmentation with DL is the time required to gather enough masks for training a highly generalized model (Liu, 2024). Here the work of Big Data and the creation of some large database, which has facilitated the overall development of DL semantic segmentation algorithms is to be acknowledged. Nonetheless, domain-specific models (new or existing), such as flocculation, require sufficient labelled data for training. Similarly, data augmentation strategies, while transforming the data shortage problem into a huge milestone, can only be effective if the masks are accurately annotated.

The demand for semantic segmentation has influenced problem-specific attention and modifications. Regarding data annotation, detailed image labelling (pixel annotations) takes time and is laborious, aside from the need for domain knowledge (Lambert et al., 2021). For instance, Lambert et al. (2021) spent more than a year and four months annotating the 80,000 images used for training the MSeg algorithm. This problem introduces *"weakly-supervised semantic segmentation"* as a novel concept, with the goal of training a model with sparsely labelled images (annotations) with good generalization accuracy. With this method, image masks are easier to generate in a short period of time than the full pixel annotations using the usual approach (full hand annotation). This emerging concept includes the usage of annotation bounding boxes (Dai et al., 2015), scribble annotation (Wang et al., 2019a), and point annotation (Qian et al., 2019). Table 5 summarizes the advantages and disadvantages of the current weakly-supervised semantic segmentation concepts. A detailed review on weakly supervised semantic segmentation can be found in Mo et al. (2022).

For flocculation images, this concept gives an innovative technique to handle large floc image datasets, which could encourage the adoption of DL semantic segmentation models in flocculation modelling, further advancing water treatment modelling. Daus et al. (2024) suggested that manual annotation exaggerates mean floc size of smaller particles, introduces bias, and decreases the accuracy of DL models. The authors presented a novel copy-paste mask creation and data augmentation approach.

Table 5
Summarized performance of weakly-supervised semantic segmentation algorithms adapted from Mo et al. (2022).

Annotation type	mToU % range	Advantages	Disadvantages
Image-level labelling / pseudo-pixel	40.6 - 66.7 33.4	• Improved image annotation speed and ease of use. • Objects of interest can easily be identified.	• The approach consumes time when handling very large datasets, compared to other methods. • Boundary layers of images may not be properly identified. • Objection location information is often missing and needs to be integrated for improved performance. • Pseudo pixels metric could mislead segmentation if not effectively handled or modelled as a loss function.
Bounding box	48 - 64.6	• Provide precise location of objects. • Highly efficient for images with blurry object edges	• It depends on the image annotation quality.
Scribbling	63.1 - 76	• Provide precise location of objects. • Easier than the bounding box method. • Highly efficient for images without thorough pixel annotation. • Highly efficient for images with blurry object edges.	• Requires domain expert knowledge. • Possess high selectivity and variation in scribbling pattern between the same or similar images.
Point	46.1 - 70.5 30 19.6	• The simplest form of annotation. • Excellent performance when handling large dataset. • Compatible with diverse scene images. • Highly efficient for images with blurry object edges.	• Seldomly applied in semantic segmentation. • Lesser information for the algorithm to learn. • Poor performance when handling complex datasets. • May not be compatible for particle-based images or floc images, when considering the primary particle.

The authors developed a multi-stage approach to generating floc masks. The denoised image was first segmented using the watershed technique, after which contours are identified. Selective segmentation was further carried out by cutting selected flocs (objects of interest) with bounding boxes, which were then blended into a new background to generate the masks. Data augmentation was performed by flipping, scaling, and rotating images and objects during the copy-paste operation, as well as truncation and occlusion. Although their approach is more trustworthy than manual annotation, further research into reliable alternatives is recommended.

5.3. Real-time semantic segmentation

Real-time semantic segmentation is a novel and interesting approach to DL applications. This technique has been widely implemented in autonomous driving, video surveillance, sensor-based AI applications (Ha et al., 2017; Wang et al., 2019a), and more recently in water treatment (Yokoyama et al., 2024). Real-time segmentation was inspired by the challenge of DL segmentation models' high computational powers. This hindered the adaptability of many DL models in providing real-time inference, which might encourage promoting sustainable and real-world solutions. As a result, transfer learning is employed to create a light computational algorithm, such as the U-Net architecture, which is efficient and robust. This concept is generally in three forms, as briefly explored below.

The U-Net encoder is the architecture's backbone and accounted for the majority of its computational power (Mo et al., 2022). SwiftNet (Oršic et al., 2019) and Deep Feature Aggregation Network – DFANet (Li et al., 2019) are examples of architectures with lightweight backbone that use less computational power. SwiftNetRN-18, proposed by Oršic et al. (2019) explored ImageNet, ResNet-18 (He et al., 2016) and MobileNet v2 (Sandler et al., 2018) as the backbone, with an up sampling layer via a laterally connected decoder. DFANet slightly modified the Xception model and used it as the backbone to achieve a very accurate real-time semantic segmentation (Li et al., 2019). The architecture used an innovative approach by reusing high-level features that connect structural details and pixel information. While this approach is unique, Mo et al. (2022) argued that its main weakness is the disparity between the backbone architecture used, which is unfamiliar with semantic segmentation tasks (i.e., the majority of image recognition classification algorithms). Consequently, the backbone may fail to accurately cover the multi-scale information inherent in the images unless global average pooling is utilized, which is not a guaranteed solution.

Specialized backbone architectures for real-time semantic segmentation of high and low spectral images have been investigated, including the Lightweight Encoder–Decoder Network (LEDNet) designed by Wang et al. (2019a) and the low-rank-to-high-rank context reconstruction framework designed by Chen et al. (2020), among others. LEDNet used an encoder network made up of split-shuffle blocks with downsampling and shuffling, which is connected to a decoder network with an attention pyramid network (APN) for semantic segmentation. These concepts are highly efficient, lightweight, and have demonstrated remarkable accuracy and precision. However, they often require a huge amount of data with precise masks for training.

YOLO ("You only look once") is a lightweight algorithm capable of real-time semantic segmentation, image classification, oriented object detection, and other tasks (Chen and Zhu, 2024; Liu et al., 2022; Redmon et al., 2016). YOLO has been extensively explored for real-time image analysis and modelling tasks. The YOLO architecture is in three components: the backbone, the neck, and the head. The YOLO series has been improved after its initial publication in 2016 by Redmon et al. (2016). Wang et al. (2024) recently unveiled YOLOv9, an improvement over the previous YOLOv8. The YOLOv9 implementation was built on YOLOv7 (Wang et al., 2022a) and dynamic YOLOv7 architecture (Lin et al., 2023). The YOLOv7 architecture previously introduced the trainable bag-of-freebies and E-ELAN (Extended Efficient Layer Aggregation Networks), in addition to the previous CSPNet backbone (Wang et al., 2019a) architecture in YOLOv5 (Chen and Zhu, 2024; Wang et al., 2022a). The authors of YOLOv9 introduced reversible programable gradient information (PGI) to mitigate the information bottleneck encountered in its predecessors (information loss during the input data feedforward process). The CSPNet architecture was then replaced with GLEAN, which improved both the architecture and model inference performance.

Despite these tremendous and robust algorithms being available via open-source platforms, we were unable to locate any studies on flocculation or water treatment that investigated any of these real-time segmentation algorithms. Although Chen and Zhu (2024) developed a real-time waste recognition model for surface water cleanup with an unmanned boat based on the YOLOv5s algorithm. It is believed that these algorithms will be adopted by future flocculation and water treatment modelling studies with interest in real-time modelling.

6. Future perspectives: pathway towards automation

The significant progress on flocculation modelling, digital floc image acquisition and recent image-based DL techniques indicates great potential and possible improvements in water treatment and process modelling. Limited studies have been carried out on advanced DL and ML algorithms implementation, primarily because of factors such as computational complexities and ground truthing, among others. The limited known findings demonstrate the potential for effective use of existing algorithms to accelerate the attainment of full automation in smart water and wastewater treatment. Of all the current challenges, generating timely and accurate masks is a major limitation. This could also be attributed to the complexities and variations in floc images caused by the camera lens, floc properties of interest, water quality, flocculant/coagulant type and other factors. As established by Daus et al. (2024), manual annotation is unreliable for flocs image ground truthing; thus, the best and one-size-fits-all strategy is to use a sensitivity analysis (error analysis) via the traditional segmentation method. This will assist authors in estimating the error in the mask generation; hence, masks with the least error can be used to train the DL/ML model, resulting in excellent segmentation accuracy. Such an approach will significantly reduce the time required to generate masks.

Additionally, sparse data labelling approaches such as the copy-paste method (Daus et al., 2024) and the weakly supervised approaches would provide a novel way to explore DL robustness in segmenting floc images. An exploration of the compatibility of weakly supervised model could further guide on the probable detail and the extent necessary for mask generation. Although Fukasawa et al. (2024) trained a DeepLab v3+ model with only 10 images, we believe it ought to have benefitted from data augmentation, increasing the model's generalizability potential. Therefore, we strongly believe that the combination of data augmentation and weakly supervised learning may facilitate mask generation, complementing the sensitivity analysis, which also requires physical examination of sparsely selected images.

Furthermore, floc images of diverse water type, coagulants, camera properties/imaging device, varying gradient of velocity and flocculation conditions can easily be annotated using sensitivity/error analysis and sparse sampling technique, to develop an accurate large floc-dataset, which can facilitate the development of the first highly generalized floc image segmenting DL model. This kind of model can be deployed for high inference and automatic modelling of flocculation kinetic and aggregate settling velocity – the most essential aspect of water treatment process.

Given the recent automatic flocculant dose obtained with the MaskRCNN model by Yokoyama et al. (2024), real-time flocculation kinetics and flocs evolution modelling are considered realistic, leveraging on the present pool of real-time segmentation techniques. In addition, the rapidly evolving field of Artificial Intelligence and its application through generative AI and Agentic AI (Zhang and Ng, 2024), coupled with the available IoT devices (Das et al., 2025), could facilitate the deployment of these robust models with sensing device in water treatment. Such milestone will facilitate full system automation in water treatment as initiated by Yokoyama et al. (2024) and other authors. A pictorial overview of the extensive discussions and future directions on possible DL algorithms adaptable for flocculation kinetics and real-time water treatment modelling is presented in Fig. 5.

Aside from flocculation modelling, the DL-based imaging modelling could provide real-time particle/pollutant tracking in water works (for example, microplastic monitoring and removal), real-time tracking of aggregate settling behaviour, and pollutant detection in reactors. For example, the high inference and existing pool of studies on the YOLO model series show that real-time particle aggregation and breakage tracking are feasible. Therefore, research aimed at achieving this goal would transform the dynamics of water treatment modelling and floc monitoring, facilitating smart water treatment. Similar results could be achieved by properly training real-time segmentation algorithms (e.g. LEDNet, DFANet, etc.). Following the pool of evidence, it is considered safe to foresee full system automation in water and wastewater treatment, which may be achieved by leveraging IoT and DL algorithms, thereby facilitating the attainment of industry 4.0 and 5.0.

Every prospective research direction is bound to encounter significant limitations at full-scale implementations. In the context of flocculation monitoring and modelling using image-based DL techniques, we believe that some pertinent issues surrounding image capturing, like the noise in images, camera lens and field of view, water turbidity and flocs size of interests, etc., have been addressed by previous studies. For

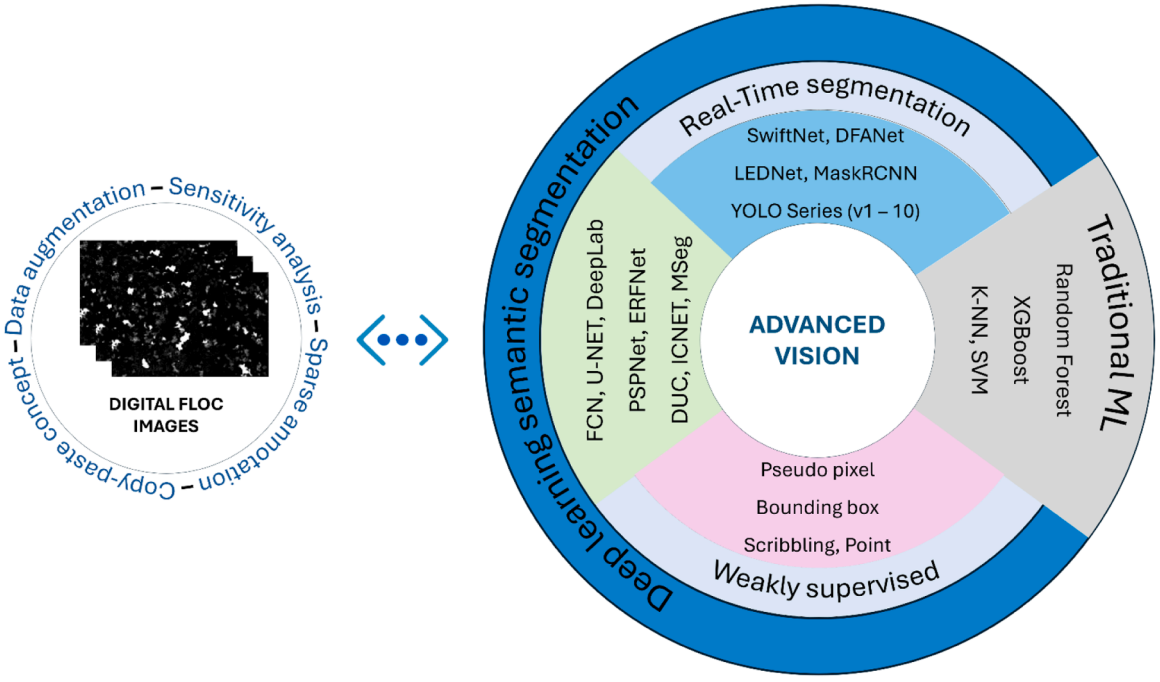


Fig. 5. Overview of advanced vision ML and DL techniques applicable for implementing automated image-based water treatment modelling.

example, Moruzzi et al. (2024) used the NiDIA technique to acquire high resolution images from multi chambers pilot plant and performed detailed kinetic modelling. Regarding the implementation at a full-scale treatment plant, we believed that situations such as staff training, regulatory compliance, and particularly its integration with legacy systems are essential bottlenecks and restraints that will need to be addressed. Nonetheless, successful implementation of ML and DL models for optimising large-scale treatment has been reported over the last decade, as seen in Egypt Nasr et al. (2012), Sweden (Wang et al., 2021, 2022b), among others.

In terms of installations, the major barrier can be traced to the acquisition of appropriate hardware, such as image capturing devices. Nonetheless, the computational complexities, inference rate, and sensor compatibility in data transfer, and relevant infrastructure are critical challenges that could limit image-based DL applications in water treatment. Regardless, these limitations can be overcome with the emerging pool of AI virtual resources through Agentic AI, further proving that DL can be implemented in full-scale treatment system. Besides, studies like Fukasawa et al. (2024) and Yokoyama et al. (2024) have proven that DL models ca be seamlessly deployed at full-scale treatment plant.

Lastly, ethical considerations in image database management and data sharing policies are the other major concerns that may arise from implementation. This may warrant the need for a robust data security system and cyber intrusion management system, under certain circumstances. However, floc image management can be achieved at the treatment plant level, where data acquisition is managed locally rather than through the internet. Besides, water treatment databanks also need to be created to facilitate global adoption, as seen in other fields like remote sensing. A unique attempt has been made by Liu et al. (2023), who created an open-access image database for chemical microparticles of diverse size range. Likewise, Bankole et al. (2025a) deployed an open-source flocculation kinetic tool that allows users to use the model without retaining their data by including an auto-destroy command in the model at the backend, to maintain users' data privacy.

7. Conclusion

This study critically reviews recent developments in flocculation modelling using ML and DL algorithms. We elucidate the application of computing platforms in enhancing floc image analysis and review best practices semantic segmentation using ML algorithms, particularly DL, with an emphasis on their applicability to improving flocculation and water treatment modelling. Our findings reveal significant gaps and challenges limiting DL adoption, such as computational complexities, floc image ground truthing and masks generation, and issues arising from variations in water quality and coagulant dosage, among other factors. Processing time, seamless data management, and real-time modelling has influenced a drift to low-level computing platforms for images analyses and modelling tasks, instead of the standalone image analysis software. Factors such as the camera depth of field, extent of illumination penetration, presence of artefacts (bubbles) on reactor walls, and image resolution affects the accuracy of image acquisition and downstream model accuracy. Despite the introduction of error analysis over the past decades, studies have not established the sensitivity of image processing (filters, reconstruction etc.) to floc texture measurements and image analysis errors.

In response to these gaps, we propose the adoption of sensitivity analysis as the most efficient and time-saving method for generating floc masks. Accurate annotated masks of floc images acquired under diverse treatment conditions (water quality, coagulants, gradient of velocity, etc.) can be achieved to form large database within short time. Also, the exploration of sparse sampling for weakly supervised segmentation in combination with data augmentation as a viable means of assessing DL robustness in segmenting floc images is proposed. The future of DL implementation in water treatment lies in the implementation of real-time semantic segmentation algorithms and object detection models, enabling real-time pollutant tracking, floc aggregation and breakage tracking, and particularly pollutant identification during treatment. Based on the proposed concept, autonomous modelling and the attainment of the industry 4.0 and 5.0 in water and wastewater treatment is envisaged.

CRediT authorship contribution statement

Abayomi O. Bankole: Writing – original draft, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Rodrigo Moruzzi:** Writing – review & editing, Validation, Supervision, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Rogério G. Negri:** Writing – review & editing, Supervision, Software. **John Bridgeman:** Writing – review & editing. **Soroosh Sharifi:** Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

Abayomi O. Bankole and Rodrigo Moruzzi are grateful to the Brazil Federal Agency for Support and Evaluation of Graduate Education (CAPES) in the scope of the program CAPES-Print, process number 88887.468887/2019-00 and 88887.310463/2018-00. Rodrigo Moruzzi is also grateful to CNPq, Brazil for grant numbers 44159/2023-0, 309788/2021-8 and 201641/2024-0 and FAPESP, Brazil 2023/080521. Rogerio Negri, appreciate CNPq Brazil for grant number 305220/2022-5.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.hazadv.2025.100870.

Data availability

No data was used for the research described in the article.

References

Abreu, D., Fernandes del Pozo, D., Deruyttere, E., Daly, A.J., Van der Meeren, P., Nopens, I., Gallo, P., 2025. A holistic determination of optimal coagulant dosage by exploring the dynamics of particle size distribution, floc structure, and electrochemistry. Chem. Eng. J. 508, 161055. https://doi.org/10.1016/j.cej.2025.161055.

Achite, M., Samadianfard, S., Elshaboury, N., Sharafi, M., 2022. Modeling and optimization of coagulant dosage in water treatment plants using hybridized random forest model with genetic algorithm optimization. Environ. Dev. Sustain. https://doi.org/10.1007/s10668-022-02523-z.

Al-Ani, S., Guo, H., Fyfe, S., Long, Z., Donnaz, S., Kim, Y., 2025. Effect of training sample size, image resolution and epochs on filamentous and floc-forming bacteria classification using machine learning. J. Environ. Manag. 379, 124803. https://doi.org/10.1016/j.jenvman.2025.124803.

Altowayti, W.A.H., Shahir, S., Othman, N., Eisa, T.A.E., Yafooz, W.M.S., Al-Dhaqm, A., Soon, C.Y., Yahya, I.B., Che Rahim, N.A.N., binti, Abaker, M., Ali, A., 2022. The role of conventional methods and artificial intelligence in the wastewater treatment: a comprehensive review. Processes 10 (9), 9. https://doi.org/10.3390/pr10091832. Article.

Alvi, M., Batstone, D., Mbamba, C.K., Keymer, P., French, T., Ward, A., Dwyer, J., Cardell-Oliver, R., 2023. Deep learning in wastewater treatment: a critical review. Water Res. 245, 120518. https://doi.org/10.1016/j.watres.2023.120518.

Amjad, H., Khan, Z., Tarabara, V.V., 2015. Fractal structure and permeability of membrane cake layers: effect of coagulation–flocculation and settling as pretreatment steps. Sep. Purif. Technol. 143, 40–51. https://doi.org/10.1016/j.seppur.2015.01.020.

Arab, M., Akbarian, H., Gheibi, M., Akrami, M., Fathollahi-Fard, A.M., Hajiaghaei-Keshteli, M., Tian, G., 2022. A soft-sensor for sustainable operation of coagulation

and flocculation units. Eng. Appl. Artif. Intell. 115, 105315. https://doi.org/10.1016/j.engappai.2022.105315.
Argaman, Y.A., 1971. Pilot-plant studies of flocculation. J. AWWA 63 (12), 775–777. https://doi.org/10.1002/j.1551-8833.1971.tb02612.x.
Argaman, Y., Kaufman, W.J., 1970. Turbulence and flocculation. J. Sanit. Eng. Div. 96 (2), 223–241. https://doi.org/10.1061/JSEDAI.0001073.
Badrinarayanan, V., Kendall, A., Cipolla, R., 2017. SegNet: a deep convolutional encoder-decoder architecture for image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 39 (12), 2481–2495. https://doi.org/10.1109/TPAMI.2016.2644615. IEEE Transactions on Pattern Analysis and Machine Intelligence.
Bankole, A.O., James, A.O., Odjegba, E.E., Bankole, A.R., Emmanuel, B.I., Fiore, F.A., Pu, J.H., Moruzzi, R., 2023. Factors affecting sanitation coverage in three income levels and potential toward achieving SDG 6.2. Water Policy 25 (2), 146–176. https://doi.org/10.2166/wp.2023.181.
Bankole, A.O., Moruzzi, R., Negri, R.G., Bressane, A., Reis, A.G., Sharifi, S., James, A.O., Bankole, A.R., 2024a. Machine learning framework for modeling flocculation kinetics using non-intrusive dynamic image analysis. Sci. Total Environ. 908, 168452. https://doi.org/10.1016/j.scitotenv.2023.168452.
Bankole, A.O., Moruzzi, R., Negri, R.G., Campos, L.C., 2025a. A novel open-source framework for automatic flocculation kinetics and retention time modelling using image analysis and swarm intelligence. J. Water Process Eng. 74, 107871. https://doi.org/10.1016/j.jwpe.2025.107871.
Bankole, A.O., Moruzzi, R., Negri, R.G., Campos, L.C., 2022a. FLOCXION - Flocculation Kinetics and Retention Time Optimisation using Swarm Intelligence and Image Analysis [Source Code]. https://doi.org/10.24433/CO.5496877.v1.
Bankole, A.O., Moruzzi, R., Negri, R.G., Oishi, C.M., Bankole, A.R., James, A.O., 2024b. MI-NiDIA: a scalable framework for modeling flocculation kinetics and floc evolution in water treatment. Softw. Impacts, 100662. https://doi.org/10.1016/j.simpa.2024.100662.
Bankole, A.O., Oluwasanya, G., Odjegba, E.E., 2022b. Evaluation of groundwater suitability in the cretaceous Abeokuta Formation, Nigeria: implications for water supply and public health. Groundw. Sustain. Dev., 100845 https://doi.org/10.1016/j.gsd.2022.100845.
Bao, W., Zhu, S., Guo, S., Wang, L., Huang, S., Fu, J., Ye, Z., 2018. Particle size distribution mathematical models and properties of suspended solids in a typical freshwater pond. Environ. Pollut. 241, 164–171. https://doi.org/10.1016/j.envpol.2018.05.063. Scopus.
Baum, A., Moiseyenko, R., Glanville, S., Martini Jørgensen, T., 2023. Image-based characterization of flocculation processes through PLS inspired representation learning in convolutional neural networks. J. Chemom. e3534. https://doi.org/10.1002/cem.3534. *n/a*(n/a).
Borges, M.C.P., Abreu, S.B., Lima, C.H.R., Cardoso, T., Yonamine, S.M., Araujo, W.D.V., Silva, P.R.S., Machado, V.B., Moraes, V., Silva, T.J.B., Reis, V.A., Santos, J.V.R., Reis, M.L., Canamary, É.A., Vieira, G.C., Meireles, S., 2022. The Brazilian National System for Water and Sanitation Data (SNIS): providing information on a municipal level on water and sanitation services. J. Urban Manag. https://doi.org/10.1016/j.jum.2022.08.002.
Bradski, G., 2000. The OpenCV library. Dr. Dobb's J. Softw. Tools.
Bratby, J., 2016. Coagulation and flocculation in water and wastewater treatment. Water Intell. Online 15 (0). https://doi.org/10.2166/9781780407500, 9781780407500–9781780407500.
Bratby, J., Miller, M., Marais, G., 1977. Design of flocculation systems from batch test Data. Water SA 3 (4), 173–182.
Bratby, J.R., 1981. Interpreting laboratory results for the design of rapid mixing and flocculation systems. J. AWWA 73 (6), 318–325. https://doi.org/10.1002/j.1551-8833.1981.tb04721.x.
Bressane, A., Goulart, A.P.G., Melo, C.P., Gomes, I.G., Loureiro, A.I.S., Negri, R.G., Moruzzi, R., dos Reis, A.G., Formiga, J.K.S., da Silva, G.H.R., Thomé, R.F., 2023. A non-hybrid data-driven fuzzy inference system for coagulant dosage in drinking water treatment plant: machine-learning for accurate real-time prediction. Water 15 (6), 6. https://doi.org/10.3390/w15061126. Article.
Bridgeman, J., Dapelo, D., 2025. Flocculator design–the case for vorticity. J. Water Process Eng. 74, 107649. https://doi.org/10.1016/j.jwpe.2025.107649.
Buades, A., Coll, B., Morel, J.M., 2011. Non-Local means denoising. Image Process.On Line 1, 208–212. https://doi.org/10.5201/ipol.2011.bcm_nlm.
Cardona, J., Ferreira, C., McGinty, J., Hamilton, A., Agimelen, O.S., Cleary, A., Atkinson, R., Michie, C., Marshall, S., Chen, Y.C., Sefcik, J., Andonovic, I., Tachtatzis, C., 2018. Image analysis framework with focus evaluation for *in situ* characterisation of particle size and shape attributes. Chem. Eng. Sci. 191, 208–231. https://doi.org/10.1016/j.ces.2018.06.067.
Chaki, J., Dey, N., 2018. A Beginner's Guide to Image Preprocessing Techniques. CRC Press. https://doi.org/10.1201/9780429441134.
Chakraborti, R., Atkinson, J., Benschoten, J., 2000. Characterization of alum floc by image analysis. Environ. Sci. Technol. 34. https://doi.org/10.1021/es990818o.
Chakraborti, R.K., Gardner, K.H., Atkinson, J.F., Van Benschoten, J.E., 2003. Changes in fractal dimension during aggregation. Water Res. 37 (4), 873–883. https://doi.org/10.1016/S0043-1354(02)00379-2.
Chatterjee, S., Ghosh, S., Dawn, S., Hore, S., Dey, N., 2016. Forest type classification: a hybrid NN-GA model based approach. In: Satapathy, S.C., Mandal, J.K., Udgata, S.K., Bhateja, V. (Eds.), Information Systems Design and Intelligent Applications. Springer, India, pp. 227–236. https://doi.org/10.1007/978-81-322-2757-1_23.
Chen, L., Zhu, J., 2024. Water surface garbage detection based on lightweight YOLOv5. Sci. Rep. 14 (1), 6133. https://doi.org/10.1038/s41598-024-55051-3.
Chen, L.C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L., 2018a. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. IEEE Trans. Pattern Anal. Mach. Intell. 40 (4), 834–848. https://doi.org/10.1109/TPAMI.2017.2699184. IEEE Transactions on Pattern Analysis and Machine Intelligence.
Chen, L.C., Papandreou, G., Schroff, F., & Adam, H. (2017). Rethinking atrous convolution for semantic image segmentation (arXiv:1706.05587). arXiv. 10.48550/arXiv.1706.05587.
Chen, L.C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018b). Encoder-decoder with atrous separable convolution for semantic image segmentation (arXiv: 1802.02611). arXiv. 10.48550/arXiv.1802.02611.
Chen, W., Zhu, X., Sun, R., He, J., Li, R., Shen, X., Yu, B., 2020. Tensor low-rank reconstruction for semantic segmentation. In: Vedaldi, A., Bischof, H., Brox, T., Frahm, J.M. (Eds.), Computer Vision–ECCV 2020. Springer International Publishing, pp. 52–69. https://doi.org/10.1007/978-3-030-58520-4_4.
Chong, J.W.R., Khoo, K.S., Chew, K.W., Ting, H.Y., Iwamoto, K., Ruan, R., Ma, Z., Show, P.L., 2024. Artificial intelligence-driven microalgae autotrophic batch cultivation: a comparative study of machine and deep learning-based image classification models. Algal Res. 79, 103400. https://doi.org/10.1016/j.algal.2024.103400.
Chong, J.W.R., Khoo, K.S., Chew, K.W., Ting, H.Y., Show, P.L., 2023. Trends in digital image processing of isolated microalgae by incorporating classification algorithm. Biotechnol. Adv. 63, 108095. https://doi.org/10.1016/j.biotechadv.2023.108095.
Chowdhury, S., Karanfil, T., 2024. Applications of artificial intelligence (AI) in drinking water treatment processes: possibilities. Chemosphere 356. https://doi.org/10.1016/j.chemosphere.2024.141958. Scopus.
Cleasby, J.L., 1984. Is velocity gradient a valid turbulent flocculation parameter? J. Environ. Eng. 110 (5), 875–897. https://doi.org/10.1061/(ASCE)0733-9372(1984)110:5(875). Scopus.
Crittenden, J.C., Trussell, R.R., Hand, D.W., Howe, K.J., Tchobanoglous, G., 2012. MWH's Water Treatment: Principles and Design. John Wiley & Sons, Inc. https://doi.org/10.1002/9781118131473.
Dai, J., He, K., Sun, J., 2015. BoxSup: exploiting bounding boxes to supervise convolutional networks for semantic segmentation. In: Proceedings of the IEEE International Conference on Computer Vision (ICCV), pp. 1635–1643. https://doi.org/10.1109/ICCV.2015.191.
Das, S., Khondakar, K.R., Mazumdar, H., Kaushik, A., Mishra, Y.K., 2025. AI and IoT: supported sixth generation sensing for water quality assessment to empower sustainable ecosystems. ACS ES&T Water. https://doi.org/10.1021/acsestwater.4c00360.
Daus, S., Buchwald, T., Peuker, U.A., 2024. In-line image analysis of particulate processes with deep learning: optimizing training data generation via copy-paste augmentation. Powder Technol. 443, 119884. https://doi.org/10.1016/j.powtec.2024.119884.
Diakogiannis, F.I., Waldner, F., Caccetta, P., Wu, C., 2020. ResUNet-a: a deep learning framework for semantic segmentation of remotely sensed data. ISPRS J. Photogramm. Remote Sens. 162, 94–114. https://doi.org/10.1016/j.isprsjprs.2020.01.013.
Dongsheng, W., Yongjie, L., Lei, Z., 2017. Raw water quality assessment oriented to the drinking water treatment based on SVM model. In: Proceedings of the 29th Chinese Control And Decision Conference (CCDC), pp. 6236–6241. https://doi.org/10.1109/CCDC.2017.7978293.
Farhi, N., Kohen, E., Mamane, H., Shavitt, Y., 2021. Prediction of wastewater treatment quality using LSTM neural network. Environ. Technol. Innov. 23, 101632. https://doi.org/10.1016/j.eti.2021.101632.
Fukasawa, A., Yamato, T., Watanabe, S., 2024. Automatic control of polymer dosage using floc images in sludge dewatering plant. Environ. Process. 11 (2), 29. https://doi.org/10.1007/s40710-024-00707-5.
Giering, S.L.C., Hosking, B., Briggs, N., Iversen, M.H., 2020. The interpretation of particle size, shape, and carbon flux of marine particle images is strongly affected by the choice of particle detection algorithm. Front. Mar. Sci. 7. https://doi.org/10.3389/fmars.2020.00564.
Govoreanu, R., Saveyn, H., Van der Meeren, P., Nopens, I., Vanrolleghem, P.A., 2009. A methodological approach for direct quantification of the activated sludge floc size distribution by using different techniques. Water Sci. Technol. 60 (7), 1857–1867. https://doi.org/10.2166/wst.2009.535.
Gregory, J., 2009. Monitoring particle aggregation processes. Adv. Colloid Interface Sci. 147–148, 109–123. https://doi.org/10.1016/j.cis.2008.09.003.
Ha, Q., Watanabe, K., Karasawa, T., Ushiku, Y., Harada, T., 2017. MFNet: towards real-time semantic segmentation for autonomous vehicles with multi-spectral scenes. In: Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 5108–5115. https://doi.org/10.1109/IROS.2017.8206396.
Haarhoff, J., Joubert, H., 1997. Determination of aggregation and breakup constants during flocculation. Water Sci. Technol. 36 (4), 33–40. https://doi.org/10.1016/S0273-1223(97)00416-2.
Haarhoff, J., Van Beek, J.C., Van Zyl, H.J., 1996. Practical application of the Argaman-Kaufman flocculation model. In: Proceedings of the Beinnial Conference of the Water Institutes of Southern Africa, 8. https://wisa.org.za/wp-content/uploads/2018/12/WISA1996-034Ppdf.
Hamada, M., Zaqoot, H.A., Jreiban, A.A., 2018. Application of artificial neural networks for the prediction of Gaza wastewater treatment plant performance-Gaza strip. J. Appl. Res. Water Wastewater 5 (1), 399–406. https://doi.org/10.22126/arww.2018.874.
He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778. https://doi.org/10.1109/CVPR.2016.90.
He, W., Nan, J., Li, H., Li, S., 2012. Characteristic analysis on temporal evolution of floc size and structure in low-shear flow. Water Res. 46 (2), 509–520. https://doi.org/10.1016/j.watres.2011.11.040.

Hounslow, M.J., Ryall, R.L., Marshall, V.R., 1988. A discretized population balance for nucleation, growth, and aggregation. AIChE J. 34 (11), 1821–1832. https://doi.org/10.1002/aic.690341108.

Hu, P., Liang, L., Li, B., Xia, W., 2021. Delving into the heterocoagulation between coal and quartz at different shear rates by the focused beam reflectance measurement (FBRM) and particle vision and measurement (PVM) techniques. Fuel 286, 119445. https://doi.org/10.1016/j.fuel.2020.119445.

James, A.O., Bankole, A.O., Moruzzi, R., Silva, G.H.R., 2025. Alkaline coagulation for separation of outdoor anaerobically cultured microalgae using natural-based coagulant. Algal Res. 85, 103844. https://doi.org/10.1016/j.algal.2024.103844.

James, A.O., Bankole, A.O., Pompei, C.M.E., Dantas, G.A.S.A., Ruas, G., Silva, G.H.R., 2023. Exploration of microalgae-activated sludge growth performance in lab-scale photobioreactors under outdoor environmental conditions for wastewater biotreatment. Phycology 3 (4), 4. https://doi.org/10.3390/phycology3040033. Article.

Jarvis, P., Jefferson, B., Gregory, J., Parsons, S.A., 2005. A review of floc strength and breakage. Water Res. 39 (14), 3121. https://www.academia.edu/8717549/A_review_of_floc_strength_and_breakage.

Jin, B., Wilén, B.M., Lant, P., 2003. A comprehensive insight into floc characteristics and their impact on compressibility and settleability of activated sludge. Chem. Eng. J. 95 (1), 221–234. https://doi.org/10.1016/S1385-8947(03)00108-6.

Kim, C.M., Parnichkun, M., 2016. MLP, ANFIS, and GRNN based real-time coagulant dosage determination and accuracy comparison using full-scale data of a water treatment plant. J. Water Supply Res. Technol. Aqua 66 (1), 49–61. https://doi.org/10.2166/aqua.2016.022.

Lambert, J., Liu, Z., Sener, O., Hays, J., & Koltun, V. (2021). MSeg: a composite dataset for multi-domain semantic segmentation (arXiv:2112.13762). arXiv. 10.48550/arXiv.2112.13762.

Lawler, D.F., 1997. Particle size distributions in treatment processes: theory and practice. Water Sci. Technol. 36 (4), 15–23. https://doi.org/10.1016/S0273-1223(97)00414-9.

Lerallut, R., Decencière, É., Meyer, F., 2007. Image filtering using morphological amoebas. Image Vis. Comput. 25 (4), 395–404. https://doi.org/10.1016/j.imavis.2006.04.018.

Li, H., Xiong, P., Fan, H., Sun, J., 2019. DFANet: deep feature aggregation for real-time semantic segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 9514–9523. https://doi.org/10.1109/CVPR.2019.00975.

Li, L., Rong, S., Wang, R., Yu, S., 2021. Recent advances in artificial intelligence and machine learning for nonlinear relationship analysis and process control in drinking water treatment: a review. Chem. Eng. J. 405, 126673. https://doi.org/10.1016/j.cej.2020.126673.

Li, M., Hu, K., Wang, J., 2020. Using digital image processing to characterize flocculation of papermaking wastewater. In: Arai, K., Kapoor, S., Bhatia, R. (Eds.), Advances in Information and Communication. Springer International Publishing, pp. 770–774. https://doi.org/10.1007/978-3-030-39442-4_56.

Li, S., Liu, Y., Wang, Z., Dou, C., Zhao, W., 2024a. Constructing a visual detection model for floc settling velocity using machine learning. J. Environ. Manag. 370, 122805. https://doi.org/10.1016/j.jenvman.2024.122805.

Li, S., Liu, Y., Wang, Z., Dou, C., Zhao, W., Shu, H., 2024b. Constructing a visual detection method for coagulation effect based on image feature machine learning. J. Water Process Eng. 68, 106354. https://doi.org/10.1016/j.jwpe.2024.106354.

Liang, J., Liu, L., 2023. Prediction of optimal coagulant dosage based on FCM-ISSA-ANFIS hybrid model. Pol. J. Environ. Stud. 32 (6), 5171–5183. https://doi.org/10.15244/pjoes/168940.

Liang, L., Peng, Y., Tan, J., Xie, G., 2015. A review of the modern characterization techniques for flocs in mineral processing. Miner. Eng. 84, 130–144. https://doi.org/10.1016/j.mineng.2015.10.011.

Lin, Z., Wang, Y., Zhang, J., & Chu, X. (2023). DynamicDet: a unified dynamic architecture for object detection (arXiv:2304.05552). arXiv. 10.48550/arXiv.2304.05552.

Liu, J., 2024. A review on current progress of semantic segmentation. Appl. Comput. Eng. 33, 108–117. https://doi.org/10.54254/2755-2721/33/20230242.

Liu, J., Kuang, W., Liu, J., Gao, Z., Rohani, S., Gong, J., 2022. *In-situ* multi-phase flow imaging for particle dynamic tracking and characterization: advances and applications. Chem. Eng. J. 438, 135554. https://doi.org/10.1016/j.cej.2022.135554.

Liu, J., Zhang, Q., Chen, M., Gao, Z., Rohani, S., Gong, J., 2023. A verified open-access AI-based chemical microparticle image database for *in-situ* particle visualization and quantification in multi-phase flow. Chem. Eng. J. 451, 138940. https://doi.org/10.1016/j.cej.2022.138940.

Long, J., Shelhamer, E., Darrell, T., 2015. Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3431–3440. https://doi.org/10.1109/CVPR.2015.7298965.

Lopez-Exposito, P., Negro, C., Blanco, A., 2019. Direct estimation of microalgal flocs fractal dimension through laser reflectance and machine learning. Algal Res. 37, 240–247. https://doi.org/10.1016/j.algal.2018.12.007.

Lu, C.F., Spielman, L.A., 1985. Kinetics of floc breakage and aggregation in agitated liquid suspensions. J. Colloid Interface Sci. 103 (1), 95–105. https://doi.org/10.1016/0021-9797(85)90080-3.

Ma, Q., Liu, Y., He, Z., Wang, H., Wang, R., Kong, Y., Li, Z., 2023. Using image texture to monitor the growth and settling of flocs. AQUA Water Infrastruct. Ecosyst. Soc. 72 (10), 1825–1836. https://doi.org/10.2166/aqua.2023.014.

Maheshan, C.M., Prasanna Kumar, H., 2019. Performance of image pre-processing filters for noise removal in transformer oil images at different temperatures. SN Appl. Sci. 2 (1), 67. https://doi.org/10.1007/s42452-019-1800-x.

Maiti, A., Chakravarty, D., Biswas, K., Halder, A., 2017. Development of a mass model in estimating weight-wise particle size distribution using digital image processing. Int. J. Min. Sci. Technol. 27 (3), 435–443. https://doi.org/10.1016/j.ijmst.2017.03.015.

Marques, R.D.O., Filho, S.S.F., 2017. Flocculation kinetics of low-turbidity raw water and the irreversible floc breakup process. Environ. Technol. 38 (7), 901–910. https://doi.org/10.1080/09593330.2016.1236149.

Marques, R.D.O., Filho, S.S.F., 2022. Further investigation of the irreversible floc breakup in flocculation kinetics modelling. Water Supply 22 (4), 3814–3823. https://doi.org/10.2166/ws.2022.023.

Mo, Y., Wu, Y., Yang, X., Liu, F., Liao, Y., 2022. Review the state-of-the-art technologies of semantic segmentation based on deep learning. Neurocomputing 493, 626–646. https://doi.org/10.1016/j.neucom.2022.01.005.

Moruzzi, R.B., Bridgeman, J., Silva, P.A.G., 2020a. A combined experimental and numerical approach to the assessment of floc settling velocity using fractal geometry. Water Sci. Technol. 81 (5), 915–924. https://doi.org/10.2166/wst.2020.171.

Moruzzi, R.B., Campos, L.C., Sharifi, S., da Silva, P.G., Gregory, J., 2020b. Nonintrusive investigation of large Al-kaolin fractal aggregates with slow settling velocities. Water Res. 185, 116287. https://doi.org/10.1016/j.watres.2020.116287.

Moruzzi, R.B., da Silva, P.G., Sharifi, S., Campos, L.C., Gregory, J., 2019. Strength assessment of Al-Humic and Al-Kaolin aggregates by intrusive and non-intrusive methods. Sep. Purif. Technol. 217, 265–273. https://doi.org/10.1016/j.seppur.2019.02.033.

Moruzzi, R.B., de Oliveira, A.L., da Conceição, F.T., Gregory, J., Campos, L.C., 2017. Fractal dimension of large aggregates under different flocculation conditions. Sci. Total Environ. 609, 807–814. https://doi.org/10.1016/j.scitotenv.2017.07.194.

Moruzzi, R.B., de Oliveira, A.L., Sharifi, S., Bankole, A.O., Campos, L.C., 2024. Enhancing flocculation kinetics assessment: integrating aggregate size distribution into experimental and modelling frameworks. J. Water Process Eng. 63, 105433. https://doi.org/10.1016/j.jwpe.2024.105433.

Moruzzi, R.B., de Oliveira, S.C., 2013. Mathematical modeling and analysis of the flocculation process in chambers in series. Bioprocess. Biosyst. Eng. 36 (3), 357–363. https://doi.org/10.1007/s00449-012-0791-4.

Moruzzi, R.B., Oliveira, A.L.D., Almeida, T.D., 2018. Fractal aggregates evolution during flocculation. Braz. J. Chem. Eng. 35, 1203–1210. https://doi.org/10.1590/0104-6632.20180354s20170231.

Moruzzi, R.B., Reali, M.A.P., 2010. Characterization of micro-bubble size distribution and flow configuration in DAF contact zone by a non-intrusive image analysis system and tracer tests. Water Sci. Technol. 61 (1), 253–262. https://doi.org/10.2166/wst.2010.784.

Muhaimin, M., Chaerunisaa, A.Y., Bodmeier, R., 2021. Real-time particle size analysis using focused beam reflectance measurement as a process analytical technology tool for continuous microencapsulation process. Sci. Rep. 11 (1), 19390. https://doi.org/10.1038/s41598-021-98984-9.

Nasr, M.S., Moustafa, M.A.E., Seif, H.A.E., El Kobrosy, G., 2012. Application of Artificial Neural Network (ANN) for the prediction of EL-AGAMY wastewater treatment plant performance-EGYPT. Alex. Eng. J. 51 (1), 37–43. https://doi.org/10.1016/j.aej.2012.07.005.

Nazemzadeh, N., Nielsen, R.F., Gernaey, K.V., Mansouri, S.S., Andersson, M.P., 2021. Implementation of first-principles surface interactions in a hybrid machine learning assisted modelling of flocculation. Comput. Aided Chem. Eng. 50, 845–850. https://doi.org/10.1016/B978-0-323-88506-5.50132-7. Scopus.

Nelli, F., Ge, Z., Blackall, L., Taheriashtiani, N., Henry, R., Brumley, D.R., Grace, M., Jex, A., Burch, M., Lin, T.F., Bertelkamp, C., Willis, A., Gao, L., Schmidt, J., Crosbie, N.D., Zamyadi, A., 2024. Machine learning driven image segmentation and shape clustering of algal microscopic images obtained from various water types. Comput. Struct. Biotechnol. Rep. 1, 100014. https://doi.org/10.1016/j.csbr.2024.100014.

Odjegba, E.E., Bankole, A.O., Layi-Adigun, B.O., Dada, V.O., 2021. Water, sanitation, and hygiene in healthcare centres: appraisal in a pandemic. J. Water Sanit. Hyg. Dev. 11 (6), 926–936. https://doi.org/10.2166/washdev.2021.075.

Odjegba, E.E., Bankole, A.O., Sadiq, A., Layi-Adigun, B.O., Adebimpe, A.M., Kosemani, M.O., Ojo, E.B., Adewuyi, M.A., 2024. Water, sanitation and hygiene (WASH) index for primary healthcare facilities: towards achieving WASH security. Heliyon 10 (15). https://doi.org/10.1016/j.heliyon.2024.e35548.

Oliveira, A., da, S., Lopes, V.D.S., Filho, U.C., Moruzzi, R.B., de Oliveira, A.L., 2018. Neural network for fractal dimension evolution. Water Sci. Technol. 78 (4), 795–802. https://doi.org/10.2166/wst.2018.349.

Oršic, M., Krešo, I., Bevandic, P., Šegvic, S., 2019. In defense of pre-trained ImageNet architectures for real-time semantic segmentation of road-driving images. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 12599–12608. https://doi.org/10.1109/CVPR.2019.01289.

Pablo, G.G.J., Maria, J.L.A., Felipe, M.U.A., 2022. Low-cost methodology for the characterization of floc size in low turbidity and low alkalinity waters using image analysis. Water Pract. Technol. 17 (4), 887–900. https://doi.org/10.2166/wpt.2022.026.

Park, P.K., Lee, C.H., Lee, S., 2006. Permeability of collapsed cakes formed by deposition of fractal aggregates upon membrane filtration. Environ. Sci. Technol. 40 (8), 2699–2705. https://doi.org/10.1021/es0515304.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Müller, A., Nothman, J., Louppe, G., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É.

(2018). Scikit-learn: machine learning in python (arXiv:1201.0490). arXiv. 10.48550/arXiv.1201.0490.
Peng, S., Guo, Y., Wang, J., Wang, Y., Zhang, W., Zhou, X., Jiang, L., Lai, B., 2024. The coagulation-precipitation turbidity prediction model for precision drug delivery system based on deep learning and machine vision. J. Environ. Chem. Eng. 12 (2), 112211. https://doi.org/10.1016/j.jece.2024.112211.
Pisa, I., Morell, A., Vicario, J.L., Vilanova, R., 2020. LSTM-based IMC approach applied in Wastewater Treatment Plants: performance and stability analysis. IFAC-PapersOnLine 53 (2), 16569–16574. https://doi.org/10.1016/j.ifacol.2020.12.782.
Qian, R., Wei, Y., Shi, H., Li, J., Liu, J., Huang, T., 2019. Weakly supervised scene parsing with point-based distance metric learning. In: Proceedings of the AAAI Conference on Artificial Intelligence, 33, p. 01. https://doi.org/10.1609/aaai.v33i01.33018843. Article.
Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: unified, real-time object detection. 779–788. https://www.cv-foundation.org/openaccess/content_cvpr_2016/html/Redmon_You_Only_Look_CVPR_2016_paper.html.
Rizwan I Haque, I., Neubert, J., 2020. Deep learning approaches to biomedical image segmentation. Inform. Med. Unlocked 18, 100297. https://doi.org/10.1016/j.imu.2020.100297.
Rogowska, J., 2009. Chapter 5–Overview and fundamentals of medical image segmentation. In: Bankman, I.N. (Ed.), Handbook of Medical Image Processing and Analysis (Second Edition). Academic Press, pp. 73–90. https://doi.org/10.1016/B978-012373904-9.50013-1.
Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: convolutional networks for biomedical image segmentation (arXiv:1505.04597). arXiv. 10.48550/arXiv.1505.04597.
Safeer, S., Pandey, R.P., Rehman, B., Safdar, T., Ahmad, I., Hasan, S.W., Ullah, A., 2022. A review of artificial intelligence in water purification and wastewater treatment: recent advancements. J. Water Process Eng. 49, 102974. https://doi.org/10.1016/j.jwpe.2022.102974.
Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.C., 2018. MobileNetV2: inverted residuals and linear bottlenecks. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 4510–4520. https://doi.org/10.1109/CVPR.2018.00474.
Sankaranarayanan, S., Likozar, B., Navia, R., 2019. Real-time particle size analysis using the focused beam reflectance measurement probe for *in situ* fabrication of polyacrylamide–Filler composite materials. Sci. Rep. 9 (1), 10126. https://doi.org/10.1038/s41598-019-46451-x.
Santos Nunes, G., Sammarro Silva, K.J., Souza Freitas, B.L., Belini, V.L., Sabogal-Paz, L. P., 2022. *In-situ* microscopy investigation of floc development during coagulation-flocculation with chemical and natural coagulants. Sep. Sci. Technol. 57 (14), 2312–2322. https://doi.org/10.1080/01496395.2022.2056055.
Saqr, A.M., Kartal, V., Karakoyun, E., Abd-Elmaboud, M.E., 2025. Improving the accuracy of groundwater level forecasting by coupling ensemble machine learning model and coronavirus herd immunity optimizer. Water Resour. Manag. https://doi.org/10.1007/s11269-025-04210-w.
Seghir, S., Hasseine, A., Rasteiro, M.G., 2022. Describing the flocculation of PCC particles using population balance modelling approaches. Chem. Eng. Res. Des. 186, 638–646. https://doi.org/10.1016/j.cherd.2022.08.038.
Shelhamer, E., Long, J., Darrell, T., 2017. Fully convolutional networks for semantic segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 39 (4), 640–651. https://doi.org/10.1109/TPAMI.2016.2572683. IEEE Transactions on Pattern Analysis and Machine Intelligence.
Shi, P., Duan, M., Yang, L., Feng, W., Ding, L., Jiang, L., 2022. An improved U-net image segmentation method and its application for metallic grain size statistics. Materials 15 (13), 13. https://doi.org/10.3390/ma15134417. Article.
Singh, M., Ranade, V., Shardt, O., Matsoukas, T., 2022. Challenges and opportunities concerning numerical solutions for population balances: a critical review. J. Phys. A Math. Theor. 55 (38), 383002. https://doi.org/10.1088/1751-8121/ac8a42.
Singh, R., & Rani, R. (2020). Semantic segmentation using deep convolutional neural network: a review (SSRN Scholarly Paper 3565919). 10.2139/ssrn.3565919.
Smoluchowski, M., 1917. Mathematical theory of the kinetics of the coagulation of colloidal solutions. Z. Phys. Chem. 92 (155), 129–135.
Stoian, A., Poulain, V., Inglada, J., Poughon, V., Derksen, D., 2019. Land cover maps production with high resolution satellite image time series and convolutional neural networks: adaptations and limits for operational systems. Remote Sens. 11 (17), 17. https://doi.org/10.3390/rs11171986. Article.
Sun, S., Weber-Shirk, M., Lion, L.W., 2016. Characterization of flocs and floc size distributions using image analysis. Environ. Eng. Sci. 33 (1), 25–34. https://doi.org/10.1089/ees.2015.0311.
Tawfik, M.S., Adishesha, A.S., Hsi, Y., Purswani, P., Johns, R.T., Shokouhi, P., Huang, X., Karpyn, Z.T., 2022. Comparative study of traditional and deep-learning denoising approaches for image-based petrophysical characterization of porous media. Front. Water 3. https://doi.org/10.3389/frwa.2021.800369.
Thomas, D.N., Judd, S.J., Fawcett, N., 1999. Flocculation modelling: a review. Water Res. 33, 1579–1592. https://doi.org/10.1016/S0043-1354(98)00392-3.
Vahedi, A., Gorczyca, B., 2012. Predicting the settling velocity of flocs formed in water treatment using multiple fractal dimensions. Water Res. 46 (13), 4188–4194. https://doi.org/10.1016/j.watres.2012.04.031.
van Beers, F. (2018). Using intersection over union loss to improve binary image segmentation Bachelor, University of Groningen. https://fse.studenttheses.ub.rug.nl/18139/.
Wang, B., Qi, G., Tang, S., Zhang, T., Wei, Y., Li, L., & Zhang, Y. (2019a). Boundary perception guidance: a scribble-supervised semantic segmentation approach. 3663–3669. https://www.ijcai.org/proceedings/2019/508.
Wang, C.Y., Bochkovskiy, A., & Liao, H.Y.M. (2022a). YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors (arXiv:2207.02696). arXiv. 10.48550/arXiv.2207.02696.
Wang, C.Y., Yeh, I.H., & Liao, H.Y.M. (2024). YOLOv9: learning what you want to learn using programmable gradient information (arXiv:2402.13616). arXiv. http://arxiv.org/abs/2402.13616.
Wang, D. (2016). Research on raw water quality assessment oriented to drinking water treatment based on the SVM model. 10.2166/WS.2015.186.
Wang, D., Thunéll, S., Lindberg, U., Jiang, L., Trygg, J., Tysklind, M., 2022b. Towards better process management in wastewater treatment plants: process analytics based on SHAP values for tree-based machine learning methods. J. Environ. Manag. 301, 113941. https://doi.org/10.1016/j.jenvman.2021.113941.
Wang, D., Thunéll, S., Lindberg, U., Jiang, L., Trygg, J., Tysklind, M., Souihi, N., 2021. A machine learning framework to improve effluent quality control in wastewater treatment plants. Sci. Total Environ. 784, 147138. https://doi.org/10.1016/j.scitotenv.2021.147138.
Wang, W., 2006. Image analysis of particles by modified Ferret method–best-fit rectangle. Powder Technol. 165 (1), 1–10. https://doi.org/10.1016/j.powtec.2006.03.017.
Watanabe, Y., 2017. Flocculation and me. Water Res. 114, 88–103. https://doi.org/10.1016/j.watres.2016.12.035.
Wei, H., Gao, B., Ren, J., Li, A., Yang, H., 2018. Coagulation/flocculation in dewatering of sludge: a review. Water Res. 143, 608–631. https://doi.org/10.1016/j.watres.2018.07.029.
Xu, T.F., He, C.L., Wang, G.T., Li, B.W., Tao, Y., 2024. Real-time monitoring of activated sludge flocs via enhanced mask region-based Convolutional Neural networks. Environ. Res. 262 (Pt 1), 119792. https://doi.org/10.1016/j.envres.2024.119792.
Yokoyama, H., Yamashita, T., Kojima, Y., Nakamura, K., 2024. Deep learning-based flocculation sensor for automatic control of flocculant dose in sludge dewatering processes during wastewater treatment. Water Res., 121890 https://doi.org/10.1016/j.watres.2024.121890.
Yu, W.Z., Gregory, J., Campos, L., 2010. Breakage and regrowth of Al-humic flocs–effect of additional coagulant dosage. Environ. Sci. Technol. 44 (16), 6371–6376. https://doi.org/10.1021/es1007627.
Yu, X., Chen, S., Zhang, X., Wu, H., Guo, Y., Guan, J., 2023. Research progress of the artificial intelligence application in wastewater treatment during 2012–2022: a bibliometric analysis. Water Sci. Technol. 88 (7), 1750–1766. https://doi.org/10.2166/wst.2023.296.
Yuan, F., Cai, Y.Q., Chen, W.L., & Cheng, E. (2016). Waterworks flocculants system based on machine vision and fuzzy control. 5–9. Scopus. 10.1109/RVSP.2015.10.
Yukselen, M.A., Gregory, J., 2002. Breakage and re-formation of Alum Flocs. Environ. Eng. Sci. 19 (4), 229–236. https://doi.org/10.1089/109287502760271544.
Zhang, H., Ng, C., 2024. Applications of artificial intelligence, machine learning, and data analytics in water environments. ACS ES&T Water 4 (3), 761–763. https://doi.org/10.1021/acsestwater.4c00140.
Zhang, K., Achari, G., Li, H., Zargar, A., Sadiq, R., 2013. Machine learning approaches to predict coagulant dosage in water treatment plants. Int. J. Syst. Assur. Eng. Manag. 4 (2), 205–214. https://doi.org/10.1007/s13198-013-0166-5. Scopus.
Zhang, X., Zhu, J., Li, Z., Li, J., Ren, P., 2021. Effect of hydrodynamic breakage on floc evolution and turbidity reduction in flocculation and sedimentation processes. Water Supply 22 (2), 1409–1420. https://doi.org/10.2166/ws.2021.345.
Zhang, Z., Yang, J., Ding, L., Zhao, Y., 2012. An improved estimation of coal particle mass using image analysis. Powder Technol. 229, 178–184. https://doi.org/10.1016/j.powtec.2012.06.027.
Zhao, H., Shi, J., Qi, X., Wang, X., & Jia, J. (2017). Pyramid scene parsing network. 2881–2890. https://openaccess.thecvf.com/content_cvpr_2017/html/Zhao_Pyramid_Scene_Parsing_CVPR_2017_paper.html.
Zhao, J., Hu, Y., Feng, Q., Gao, C., 2024. Research on image segmentation processing of coal particle flocs combined with clarity detection. Int. J. Coal Prep. Util. 0 (0), 1–16. https://doi.org/10.1080/19392699.2024.2402439.
Zhao, L., Zhang, Z., 2024. A improved pooling method for convolutional neural networks. Sci. Rep. 14 (1), 1589. https://doi.org/10.1038/s41598-024-51258-6.
Zhou, Z., Siddiquee, M.M.R., Tajbakhsh, N., Liang, J., 2020. UNet++: redesigning skip connections to exploit multiscale features in image segmentation. IEEE Trans. Med. Imaging 39 (6), 1856–1867. https://doi.org/10.1109/TMI.2019.2959609. IEEE Transactions on Medical Imaging.
Zhu, Z., 2022. Predicting the flocculation kinetics of fine particles in a turbulent flow using a Budyko-type model. Environ. Sci. Pollut. Res. https://doi.org/10.1007/s11356-022-21518-x.