UWS Academic Portal

A flocculation tensor to monitor water quality using a deep learning model

Zhu, Guocheng; Lin, Jialin; Fang, Haiquan; Yuan, Fang; Li, Xiaoshang; Yuan, Cheng; Hursthouse, Andrew S.

Link to publication on the UWS Academic Portal

ORIGINAL PAPER

A flocculation tensor to monitor water quality using a deep learning model

Guocheng Zhu[1] · Jialin Lin[1] · Haiquan Fang[1] · Fang Yuan[2] · Xiaoshang Li[3] · Cheng Yuan[2,3] · Andrew S. Hursthouse[4]

Abstract

The increasing quantities of polluted waters are calling for advanced purification methods. Flocculation is an essential component of the water purification process, yet flocculation is commonly not optimal due to our poor understanding of the flocculation process. In particular, there is little knowledge on the mechanisms ruling the migration of pollutants during treatment. Here we have created the first tensor diagram, a mathematical framework for the flocculation process, analyzed its properties with a deep learning model, and developed a classification scheme for its relationship with pollutants. The tensor was constructed by combining pixel matrices from a variety of floc images, each with a particular flocculation period. Changing the factors used to make flocs images, such as coagulant dose and pH, resulted in tensors, which were used to generate matrices, that is the tensor diagram. Our deep learning algorithm employed a tensor diagram to identify pollution levels. Results show tensor map attributes with over 98% of sample images correctly classified. This approach offers potential to reduce the time delay of feedback from the flocculation process with deep learning categorization based on its clustering capabilities. The advantage of the tensor data from the flocculation process improves the efficiency and speed of response for commercial water treatment.

Keywords Flocculation · Tensor · Tensor diagram · Deep learning · Model

Introduction

There are numerous methods used in water treatment processes, such as pretreatment, conventional treatment, advanced treatment (Crini et al. 2019; Nadia Morin-Crini et al. 2022; Zhu et al. 2022). Additionally, the management of water treatment will be enhanced by any improvement in consistent function of treatment processes. As the demand for increasingly refined operational management of drinking water plants grows, the potential of increasing the amount of data used in the operation system is being investigated to increase energy saving and reduce the consumption of chemical agents, increasing the quality and efficiency of the production process (Eggimann et al. 2017; Li et al. 2021a; Worm et al. 2010). Many water treatment plants still require the development of a dosing system to meet the effluent quality requirements, and excessive manual dosing is frequently used (Zhang 2005). This situation is not conducive to the development of modern production, water supply, and drinking water quality requirements, resulting in a range of issues such as high chemical consumption, poor economic benefits, unstable water quality, and high labor intensity for workers (dos Santos et al. 2017; Imen et al. 2016; Zaque et al. 2018). In the water treatment industry, research on automatic control of coagulation dosing is both necessary and urgent.

The key to automating production and lowering reagent costs is automatic control of chemical dosing, which includes basic feed-forward feedback and hybrid control (Chen and Hou 2006; Liu et al. 2017). The feedback signal is crucial, and when combined with a feed-forward system, it

✉ Guocheng Zhu
zhuguoc@hnust.edu.cn

✉ Andrew S. Hursthouse
andrew.hursthouse@uws.ac.uk

1 College of Civil Engineering, Hunan University of Science and Technology, Xiangtan 411201, Hunan, China

2 General Water of China Co., Ltd., Beijing 100022, China

3 Xiangtan Middle Ring Water Business Limited Corporation, Xiangtan 411201, Hunan, China

4 School of Computing, Engineering and Physical Sciences, University of the West of Scotland, Paisley PA1 2BE, UK

allows for precise control of chemical dose. Feedback signals are frequently classified as a current signal (Liu et al. 2004), an optical signal (Sun and Zhong 2002), or an image signal (Xie et al. 2015). Interference from the external environment has a significant negative influence on signal feedback. Furthermore, the feedback control system is problematic due to the time lag between the point of dosing with chemicals and obtaining water treatment results such as residual turbidity (Li et al. 2021b). Stream current detector (SCD) (Dentel et al. 1989), transmitted light fluctuation (TLF) (Of et al. 2001), and image signal such as fractal dimension (Chen and Zhang 2010) are the most commonly used online feedback methods. For the control of coagulant dose, the stream current detector is positively related to zeta potential, but it is easily affected by environmental factors such as pH and electrolyte concentration (Kim et al. 2017). The TLF can reflect the size and the amount of flocs through a transmitted light fluctuation value (Gregory 1997), but it is difficult to maintain the stability and sensitivity of the value (Wang 2005). Image signal should be a promising technology that has received a lot of attention, and it can provide feedback on water quality through structure parameters of floc images, thereby determining the chemical dose (Lin and Ika 2019; Yu et al. 2015). However, there are inherent challenges, such as floc collection. Both floc fragmentation and image fuzziness will complicate the analysis of floc structure, lowering the speed of signal feedback. Furthermore, sample collection and calculation are both time-consuming, adding to operator demands. Studies on technologies that can reduce issues related to image acquisition, design, and manual calculation of features have received a lot of attention to improve the quality of the feedback signal.

Shortening signal feedback time by an effective model that demands relevant information as input, acquired before flocs settling, is an important way for solving the flocculation time-delay problem. There appears to be a lack of interest in manual design information that necessitates lengthy preprocessing. To make good predictions, we used flocs images as input data because a deep learning model, such as a convolutional neural network (CNN), can extract image features without the need for any pre-treatment (Yamamura et al. 2020), which has been successful in the development of other learning networks, particularly in the field of computer vision (Traore et al. 2018; Yuan et al. 2020). Few studies have looked into the use of floc image features extracted by a CNN model to shorten the flocculation time-delay. Because a tensor is the data format for deep learning models, developing flocculation tensor data that enhances the performance of deep learning and is a significant potential development in water treatment.

As a result, we created a new tensor of flocculation flocs made of images captured throughout the flocculation procedure. Each tensor had the shape (t, n, m), which is related to a certain pollutant level, where t is the number of images and n and m are the image matrix's row and column numbers. Turbidity was used as a term to describe the pollutant in this study. The tensor diagram, a matrix constructed from the tensor, was then generated. A convolutional neural network model was built utilizing the tensor diagram, and the effectiveness of feedback signals was examined. The learning mechanisms were evaluated by measuring relative cosine similarity (RCS) and T-distributed stochastic neighbor embedding (t-SNE).

Materials and methods

Water samples

For the treatment of a real water sample, a coagulation experiment was performed on a program-controlled jar test apparatus (ZR4-6, ZhongRun Water Industry Technology Development Co., Ltd., Shenzhen, China). The water samples were taken from the Xiangjiang River in Xiangtan City, which is the primary source of drinking water for the city's drinking water plants. As a coagulant, an analytical grade polymeric aluminum chloride purchased from Tianjing Kaitong Chemical Co. Ltd. in China was used in the flocculation experiment. The coagulant concentration in the water sample was calculated using the amount of Al_2O_3 component measured with an acid–base titration according to Chinese Standard GB 15892-2009. The turbidity water was tested using a portable turbidity meter (2100Q, HACH Company, USA). The ultraviolet absorbance at 254 nm (UV254) was measured on a TU-1901 UV–visible spectrophotometer (Purkinje General Instrument Co., Ltd., Beijing, China).

Flocculation experiment

For the treatment of a real water sample, a coagulation experiment was performed on a program-controlled jar test apparatus (ZR4-6, ZhongRun Water Industry Technology Development co. Ltd, Shenzhen, China). In detail, 1 L of each water sample was transferred into a plexiglass cylinder beaker, and the initial pH of the sample was adjusted to the set value using 0.5 mol/L hydrochloric acid and 0.5 mol/L sodium hydroxide. The sample was rapidly mixed at a predetermined agitation speed (rpm) for a fixed time, followed by a slow mixing phase at a predetermined agitation speed for set flocculation time, and finally a 30 min settling time. Water quality was measured by extracting water from the beaker 2 cm below the water surface. The G values were obtained from the jar test apparatus and adjusted by varying the stirring speed or reaction time. A digital image capture instrument (GSY-753, Shenzhen Woshijie Electronic Technology Co., Ltd, China) was fixed on the outer wall of the

beaker. Video of flocculation process was captured during the above slowly stirring phase, and images were extracted from the videos within a fixed time interval ranging from 1 to 6 s. Figure S1 shows an experimental setup that includes a video recording of flocs during the flocculation process. During the experiments, sample numbers collected were 5259 (coagulant dosing) and 4264 (pH variation).

Tensor diagram

Experimental conditions for tensor data collection are shown in Table S1 (supplementary information). A tensor for flocculation was created in this study. In a two-dimensional coordinate system, a three-dimensional tensor graph was constructed. The flocculation period is shown in Fig. 1a by the abscissa (x). The pollutant class level can be found on the y-axis. Remaining turbidity is the term we use here. Coordinate values shaped the pictures matrix's form structure (n, m). The rows and columns are represented by the numbers n and m, respectively. There are tensors in the tensor diagram (y value), n tensors in the shape structure (t, n, m), and the number of images in each row is represented by t.

Deep leaning model

We used the tensor diagram as input data and the class label as output values to establish the convolutional neural network model which was used to predict the class label of the final effluent corresponding to the input tensor diagram consisting of image data and the corresponding effluent turbidity was determined. Experimental conditions for tensor data collection in different models is shown in Table S1. Finally, the effective convolutional neural network model was used to predict the class label of the final effluent corresponding to the input tensor diagram consisting of image data, and the corresponding effluent turbidity was determined. The model was self-constructed using interpreter python and pytorch package, and it included three convolutional layers, three pooling layers, and two fully connected layers. Figure S2 shows the entire flowchart for the construction of the convolutional neural network framework. Prior to simulation, the images were pre-treated and reduced to a 48×48 pixel size in order to make the image sizes the same. The total number of images was divided into three datasets: 80% as training samples, 10% as verification samples, and 10% as test samples. In detail, the sample image data were fed into the first convolution layer, which calculated convolution using the kernel parameter. The output of the convolution operation is fed into the pooling layer, which employs the maximum pooling method to extract the most features. Convolutional kernel and polling kernel sizes were set to five and two, respectively. Strides were limited to two. The data were treated twice by convolution and polling operations with the same parameter configurations.
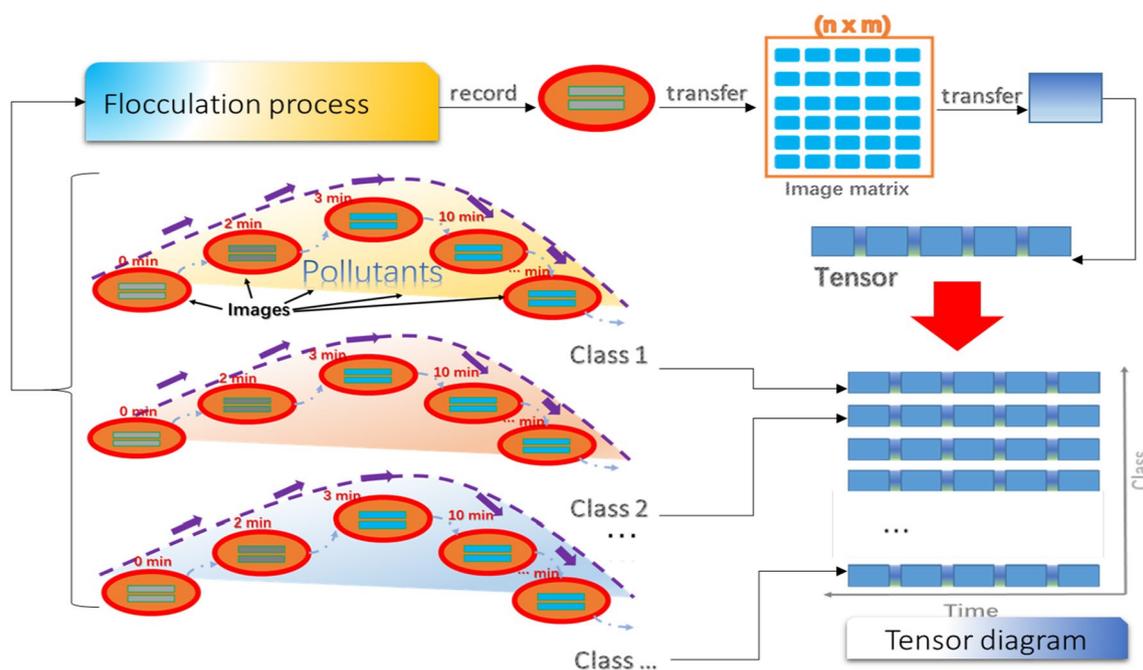

Fig. 1 A two-dimensional coordinate plot of a tensor diagram of flocculation flocs. The flocculation time is shown on the x-axis. The pollutant class level is shown on the y-axis. In practice, the value on the x-axis was used to determine how much sample to take, which was based on the ratio of the total flocculation time to the sampling interval time. This study showed that a tensor diagram is made up of a group of tensors and that it can be used to predict the class level of a pollutant when a particular factor is assessed

After completing the convolutional feature extraction, the final convolution layer is flattened and fed into the fully connected layer, where the dropout layer was introduced to inactivate some neurons with a fixed probability equal to 0.5, and then the rest of the neurons come into the fully connected layer, which was repeated twice, and finally the model provided turbidity class labels. The dropout function was used to improve generalization capability, thereby alleviating the convolutional neural network model's overfitting problem. The activation function and classifier in the model were set to rectified linear unit (ReLU) and Softmax, respectively. When compared to the Sigmoid and Tanh functions, the rectified linear unit eliminates the gradient vanishing problem (Pan et al. 2021). The learning rate was set to 0.01 by default. The total number of epochs was set to 90. The turbidity class label was the output value as target variable of prediction. After the model has been properly trained, the sample data, including training data and test data, can be used to predict the turbidity class signal and evaluate the model performance by measuring the accuracy rate. The detailed parameter settings were shown in Text S1. After the model has been properly trained, the sample data, including training data and test data, can be used to predict the turbidity class signal and evaluate the model performance by measuring the accuracy rate.

Clustering and similarity measurement

The accuracy rate in % is used to evaluate the convolutional neural network model performance by calculating the percentage ratio of correctly predicted samples to total samples. The cross-entropy is used as a loss function to evaluate the difference between predicted values (turbidity class label) and measured values (turbidity class label). Convolutional features are a high-dimensional dataset that is difficult to investigate visually. To determine whether they were separable, they must be shown in low-dimensional space using a non-linear mapping technique that reduces their dimensionality to low dimensions. The t-SNE is a technique that solves crowding problems by combining a student t-distribution with a heavy-tailed probability distribution (Van Der Maaten 2014). Further details are given in Text S2. The degree of cosine similarity between the convolutional features is then derived from Kaur and Aggarwal (2013).

$$\cos(\theta) = \frac{a \cdot b}{\|a\| \times \|b\|}$$

where a and b are pixel matrices.

Because we compared more than three samples and determined the degree of similarity, we proposed a relative cosine similarity (RCS) value. In Eq. 1, each sample feature matrix (a) would be compared to a reference value. The average feature map matrix value of all samples was used as the reference value (b).

Results and discussion

Effect of tensor's deep learning by varying coagulant dosage and pH

The effect of coagulant dosage on deep leaning model accuracy was investigated by adjusting coagulant dose levels from 0.5 to 32 mg/L, referred to as the Mod-Dos model, respectively. Raw turbidity (NTU) 10.8 NTU, temperature 20.7 °C, UV254 0.15 cm^{-1}, pH 7.62, Gt 6600, time 15 min were other raw water parameters during image collection. The total number of images in this study is 5259, divided into seven classes. Figure 2 shows how the loss function and percentage accuracy rate change as the number of epochs increases.

The model's results are given in Fig. 4a, b with the effect of coagulant dosage (denoted as Mod-Dos). The loss values for train and validation declined as the epoch increased, to 0.12 at epoch 42 from 1.63 and 0.08 at epoch 36 from 1.13, respectively. The loss value could not be reduced by increasing the epoch. With the growth of the epoch, the accuracy rates for training and validation climbed to 95.78% at epoch 32 from 30.23 and 96.93% at epoch 26 from 44.33%, respectively. They did not increase as the epoch was increased. The ultimate training and testing accuracy rates were both 92%, with an average training error of 0.25. Because certain classes had values that were near to one another and further tweaking to the sample structure, 98% accuracy rate could be achieved. Detailed information is shown in Text S3 and Fig. S3.

Similar results were seen in other factor investigation such as pH (denoted as Mod-pH): By adjusting the pH level from 4 to 10, the influence of pH on the accuracy of the deep leaning model was examined. Other raw water parameters for images collection included raw turbidity (NTU) 15.2 NTU, temperature 11.8 °C, UV254 0.23 cm^{-1}, coagulant dosage 16 mg/L, Gt 6600, and time 15 min. The total number of images in this study is 4264, with 5 classes being used. The results of the experiment are shown in Fig. 4c, d as examples. The loss values for train and validation decreased as the epoch increased to a minimum of 0.24 at epoch 55 from 1.56 and 0.13 at epoch 55 from 1.52, respectively. The accuracy rates for training and validation increased to a maximum of 91.21% at epoch 42 from 31.12 and 98.11% at epoch 46 from 33.21%, respectively. The average training error was 0.46, and both the final training and testing accuracy rates were 98%. The above results showed that with the tensor the deep leaning model had a good effect on prediction of the water quality parameter.
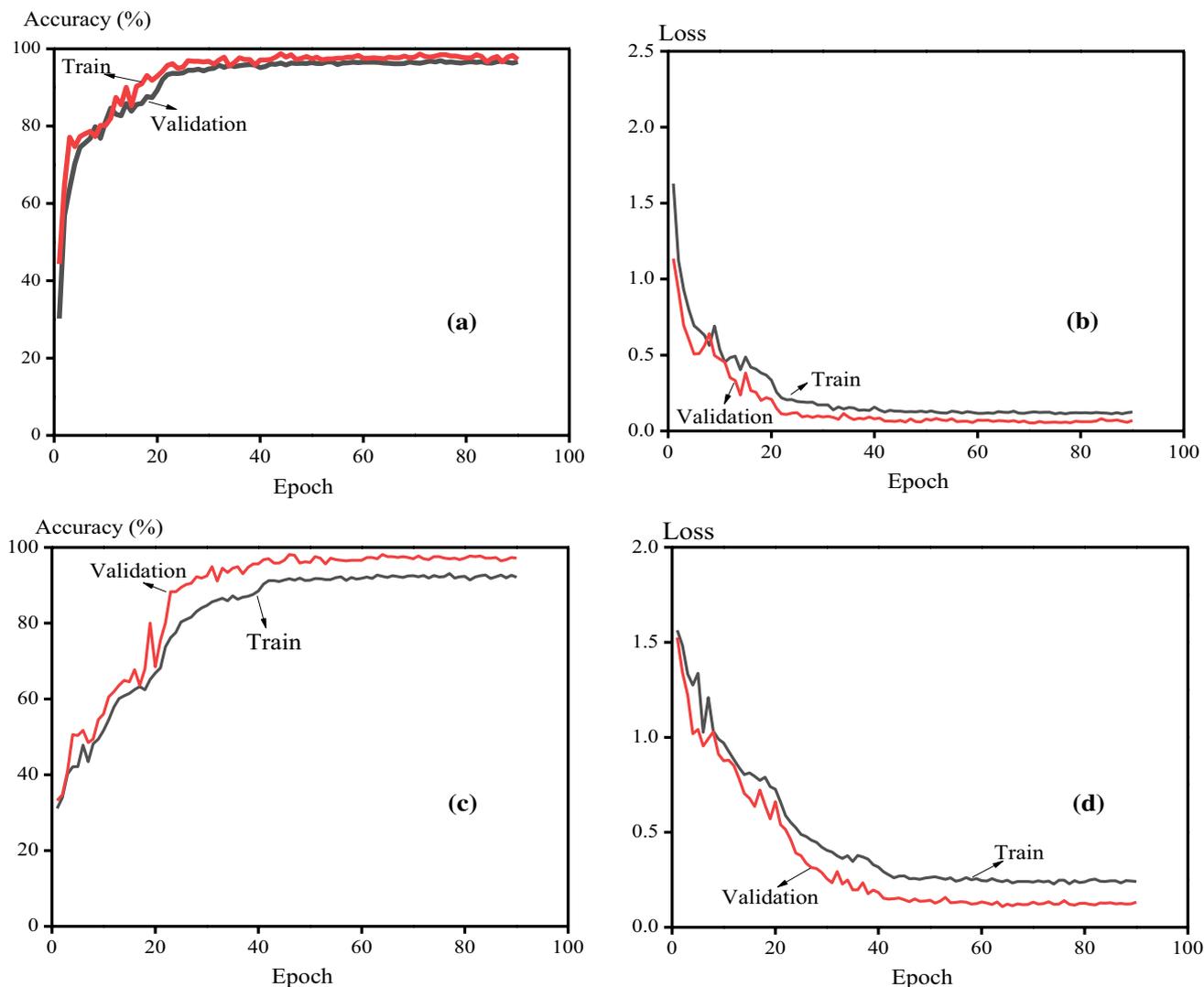
Fig. 2 Deep learning effect on prediction of turbidity signal with (**a**) Mod-Dos for training accuracy, (**b**) Mod-Dos for training loss, (**c**), Mod-pH for training accuracy, (**d**) Mod-pH for training loss. The results demonstrated that the deep learning model could achieve an accuracy of 98%, indicating that the tensor was highly sensitive to the deep learning model for predicting the turbidity signal validating the development of this type of model

According to our investigation in the drinking water plant, the raw water turbidity, coagulant dosage and flow rate have a great impact on the effluent turbidity, and other interference is often limited. In a fixed period of time, the water quality is very stable, and the creation of the water quality model based on the variation of a single factor is of great significance and important reference value for the deployment for real time treatment.

Similarity of the flocculation tensor

Because those convolutional features of tensor are so similar (Fig. 3), it is possible that all of the features generated in a flocculation process that belong to the same class can be successfully applied to a model. It is more likely to employ early stage flocculation photos to reduce the time it takes to predict effluent turbidity signal, hence minimizing the time lag of a model forecast. As a result, it will be intriguing to observe if the tensor images' convolutional features are similar. As a result, under the effect of coagulant dosage (denoted as Mod-Dos) and the pH (denoted as Mod-pH), the RCS values of the convolutional features of those tensor's images that were assessed during flocculation time were investigated (Fig. 3).

As demonstrated in Fig. 3a–j, high similarity degree values were found in the same class sample, indicating that the form of variation in the degree of similarity of most convolutional features of flocculation images appeared to
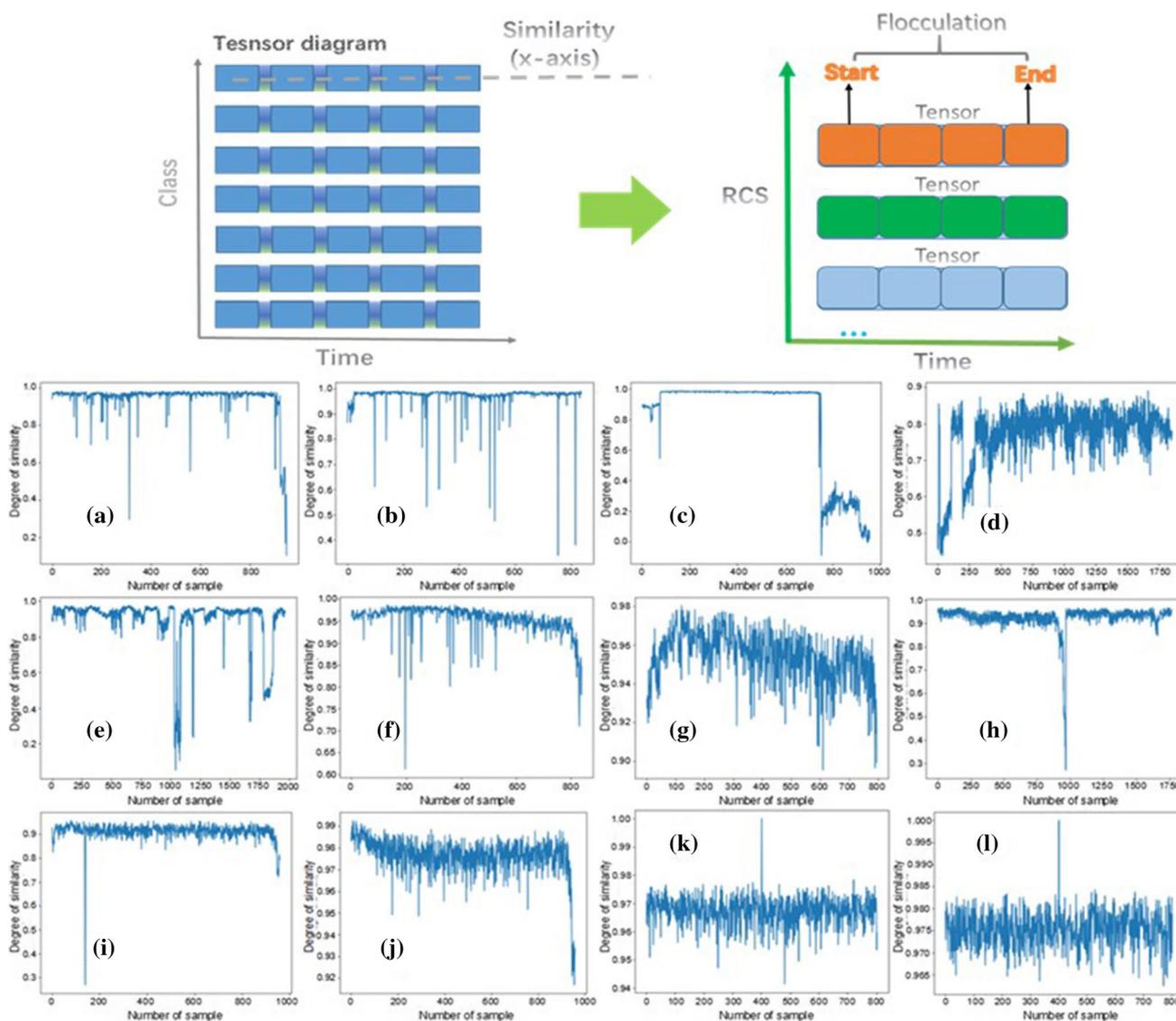
Fig. 3 Similarity convolutional features in those samples at different classes in Mod-Dos and Mod-pH: **a–e** Class 1–5 of Mod-Dos; **f–j** Class 1–5 of Mod-pH; **k** Class 1 of Mod-Dos with dataset at sampling interval 5 s. **l** Class 1 of Mod-pH with dataset at sampling interval 5 s. The x-axis stands for number of samples. The y-axis presented the degree of similarity as RCS value. The results showed that most samples (images) that constituted the tensor had similar features in the same class, suggesting we could randomly select any sample from these samples to predict feedback signal. Earlier stage samples were prioritized because their time-delay influence was the least. Environmental influence may cause sample deviations. We can define conditions to deduct those samples that did not meet the conditions in a program system, which is not a problem for this technique

be a horizontal line. A small number of samples showed similarity values that were not parallel to the horizontal line. In Fig. 3a, c, k, for example, the main variation emerged in the early and late stages of flocculation, with relatively stabilized values in the median stage. The divergence would make it more difficult to recognize such images with fewer pixels, impacting on turbidity feedback. However, because the majority of the samples had a high degree of similarity, we were able to build an effective model. As demonstrated in Fig. 3k, l, the similarity of practically all samples in the same class was around 0.97 throughout the flocculation process. Although the accuracy of the Mod-Dos and Mod-pH with the dataset for sampling time interval of 5 s was lower than that of the dataset for 1 s, the degree of similarity of the convolutional features in the models was not so low. This work demonstrated that the similarity between convolutional features is an inherent quality of the flocculation image and that fluctuation in similarity is unaffected by changes in factors.
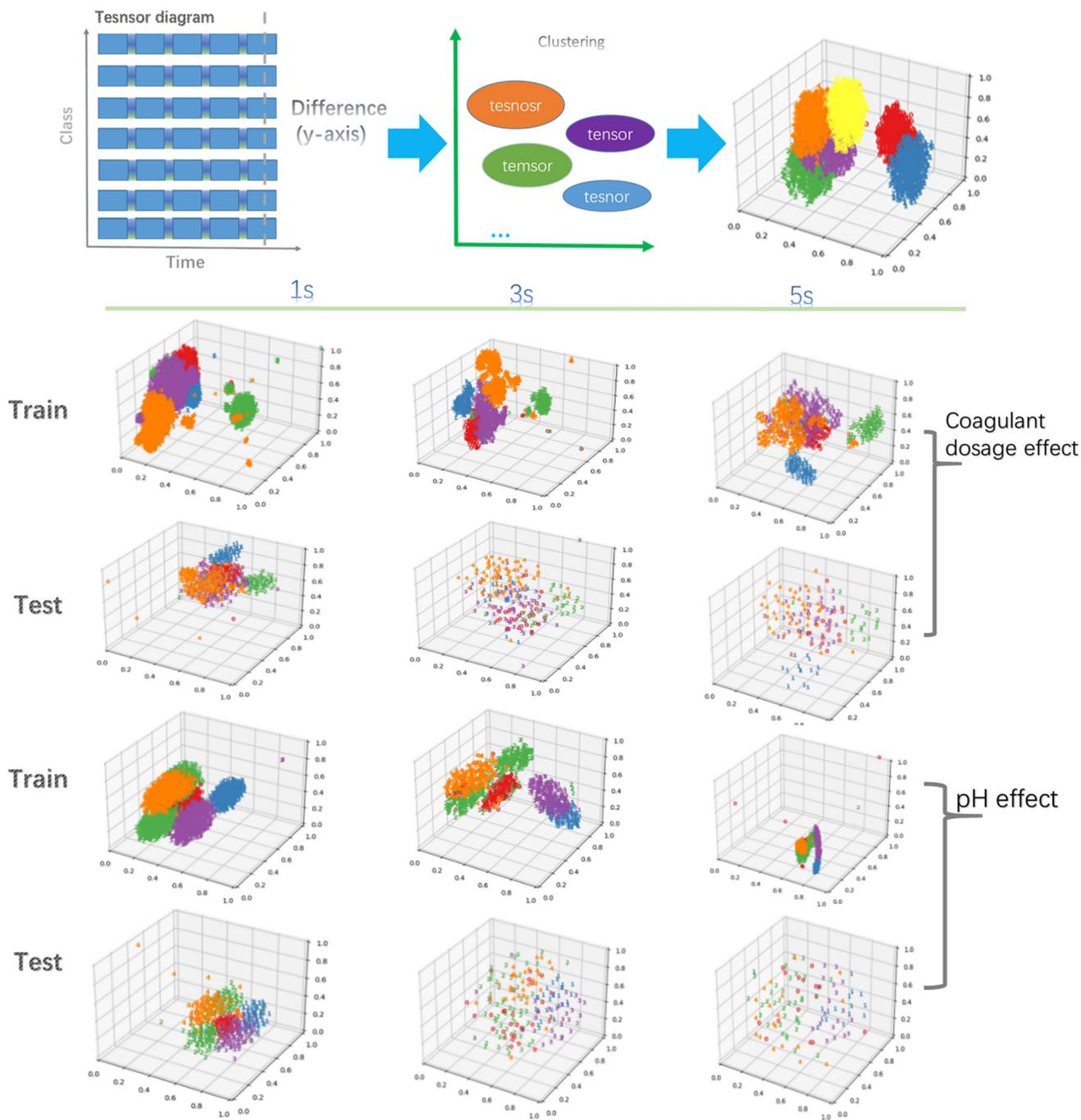
Fig. 4 The plots of tensor clustering of flocs. The plot shows that the tensor's convolutional features showed sample clustering. The results demonstrated that the tensor's model effect on sample clustering was effective, an important property. The sample size (controlled by time such as 1 s, 3 s, and 5 s) and factor were unlikely to influence the effect or nature. The nature was the fundamental basis of the tensor to construct an effective deep learning model after the feedback signal had a meaningful relationship between pollutants

Because the majority of the samples in the entire flocculation process were identical, it was logical to group them together into a single class corresponding to a fixed turbidity signal. It also shows variances in convolutional characteristics among distinct classes, allowing them to be distinguished from one another. To analyze the characteristics of these differences, we assessed the RCS values of training samples, validation samples, and test samples. The data in Tables S2–5 show the results of varying the degree of similarity and the probability density distribution of convolutional features of the tensors. The RCS values appear to be made up of each line segment that represents each

class. It means that the RCS values of samples from different classes differed and could be distinguished by the similarity values. The probability density distributions of the three datasets were also similar in shape, indicating an acceptable dataset distribution. When using a model with a greater accuracy or a model with a lower accuracy with varied sampling time interval, the results did not change. We can see that the similarity of the convolutional feature of the flocculation image is an inherent property.

Clustering of the flocculation tensor

Datasets with a potential clustering function are credited as the primary cause for excellent deep learning. The convolutional features of tensor have been shown to have good deep learning performance, with 98% prediction performance, the maximum, on feedback of turbidity signal. The tensor' clustering effect was explored using T-distributed stochastic neighbor embedding (t-SNE) visualizations in a low-dimensional state with a 30% confusion degree. Figure 4 shows the results of t-SNE visualizations of tensors' convolutional features.

As shown in Fig. 4, the convolutional features efficiently distinguished the training and testing samples, which were labeled in color numbers in three-dimensional space. We are encouraged to enhance the model prediction performance through increasing the number of samples resulting in better predictions. We examined the effect of total number of samples on the accuracy in an experiment with coagulant dose of 16 mg/L, raw water turbidity of 8.26 NTU, and sampling time intervals of 1–6 s. With the 5-s dataset, the model's prediction accuracy rate was significantly lower than that with 1-s dataset. The detailed discussion on the influence of number of samples on model prediction is shown in Text S4 and Fig. S4. However, from the model with the 1-s dataset to the model with 5 s, there was always a clustering effect, and the clustering effect of the samples was more visible with a large number of samples, such as for the 1-s dataset. These results showed that the clustering separation among different class samples is a characteristic of the tensor's convolution feature and can be displayed in low-dimensional space. Low-dimensional clustering features suggest that high-dimensional data may have a good learning effect in sample class recognition, as indicated by manifold learning, and if high-dimensional data have a good machine learning potential and it is also applied to dimensional state (Turaga et al. 2020; Guérin et al. 2021; Liu et al. 2019; Wong et al. 2018). This identifies the clustering characteristic as a crucial method for convolutional neural network to learn well. It has been shown that extracting data features using the clustering method is a frequent characteristic to achieve the goal of successful deep learning (Dong et al. 2017; Hsu and Lin 2017; Jalal et al. 2017). The convolution feature of the tensor has a strong clustering property, which makes model construction much easier.

Conclusion

The flocculation tensor that we developed for water quality signal feedback or prediction was incorporated into a deep learning network throughout this study. The turbidity signal feedback was studied using the tensor, and the accuracy rate approached 98%. The tensor has two significant features that we discovered. We were able to select earlier flocculation photos because the tensor component had a similarity that allowed us to reduce the signal feedback time-delay. Because the tensor was structured in such a way that it could be clustered, we were able to discover the basic reason for excellent deep learning. In the end, our most significant contribution was the development and structure of the flocculation tensor, along with a demonstration of the flocculation tensor's effectiveness in training a deep learning model for signal feedback, which resulted in the reduction of the influence of time-delay paving the way for further development and implementation in practice.

Declarations

References

Chen CL, Hou PL (2006) Fuzzy model identification and control system design for coagulation chemical dosing of potable water. Water Supply 6(3):97–104. https://doi.org/10.2166/ws.2006.782

Chen W, Zhang J (2010) Operation and management of urban water system. China Construction Industry Press, Beijing

Crini G, Lichtfouse E (2019) Advantages and disadvantages of techniques used for wastewater treatment. Environ Chem Lett 17:145–155. https://doi.org/10.1007/s10311-018-0785-9

Dentel SK, Thomas AV, Kingery KM (1989) Evaluation of the streaming current detector—I: use in jar tests. Water Res 23(4):413–421. https://doi.org/10.1016/0043-1354(89)90132-2

Dong X, Qian L, Huang L (2017) Short-term load forecasting in smart grid: a combined CNN and K-means clustering approach, pp. 119–125. https://doi.org/10.1109/BIGCOMP.2017.7881726

Dos Santos FCR, Librantz AFH, Dias CG, Rodrigues SG (2017) Intelligent system for improving dosage control. Acta Scientiarum Technol 39(1):33–38. https://doi.org/10.4025/actascitechnol.v39i1.29353

Eggimann S, Mutzner L, Wani O, Schneider MY, Spuhler D, Moy de Vitry M, Beutler P, Maurer M (2017) The potential of knowing more: a review of data-driven urban water management. Environ Sci Technol 51(5):2538–2553. https://doi.org/10.1021/acs.est.6b04267

Gregory J (1997) The density of particle aggregates. Water Sci Technol 36(4):1–13. https://doi.org/10.1016/j.aca.2007.07.011

Guérin J, Thiery S, Nyiri E, Gibaru O, Boots B (2021) Combining pretrained CNN feature extractors to enhance clustering of complex natural images. Neurocomputing 423:551–571. https://doi.org/10.1016/j.neucom.2020.10.068

Hsu CC, Lin CW (2017) Cnn-based joint clustering and representation learning with feature drift compensation for large-scale image data. IEEE Trans Multimed 20(2):421–429. https://doi.org/10.1109/TMM.2017.2745702

Imen S, Chang N-B, Yang YJ, Golchubian A (2016) Developing a model-based drinking water decision support system featuring remote sensing and fast learning techniques. IEEE Syst J 12(2):1358–1368. https://doi.org/10.1109/JSYST.2016.2538082

Jalal DK, Ganesan R, Merline A (2017) Fuzzy-C-means clustering based segmentation and CNN-classification for accurate segmentation of lung nodules. Asian Pac J Cancer Prev Apjcp 18(7):1869–1874. https://doi.org/10.22034/APJCP.2017.18.7.1869

Kaur S, Aggarwal D (2013) Image content based retrieval system using cosine similarity for skin disease images. Adv Comput Sci Int J 2(4):89–95

Kim S-G, Son H-J, Lee J-K, Yeom H-S, Yoo P-J (2017) Evaluation of streaming current detector (SCD) and charge analyzing system (CAS) for automation of coagulant dosage determination. J Korean Soc Environ Eng 39(4):201–207. https://doi.org/10.4491/KSEE.2017.39.4.201

Li L, Rong S, Wang R, Yu S (2021a) 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 Y, Xu L, Shi T, Yu W (2021b) The influence of various additives on coagulation process at different dosing point: From a perspective of structure properties. J Environ Sci 101:168–176. https://doi.org/10.1016/j.jes.2020.08.016

Lin J-L, Ika AR (2019) Enhanced coagulation of low turbid water for drinking water treatment: dosing approach on floc formation and residuals minimization. Environ Eng Sci 36(6):732–738. https://doi.org/10.1089/ees.2018.0430

Liu W, Ratnaweera H (2017) Feed-forward-based software sensor for outlet turbidity of coagulation process considering plug flow condition. Int J Environ Sci Technol 14(8):1689–1696. https://doi.org/10.1007/s13762-017-1284-4

Liu HN, Xue YJ, Tan YQ, Qu JH (2004) On-line monitor of the streaming current one-factor and its automation system of adding coagulants. Proc Water Environ Fed 15:232–233. https://doi.org/10.2175/193864704784148321

Liu H, Lang B, Liu M, Yan H (2019) CNN and RNN based payload classification methods for attack detection. Knowledge-Based Syst 163:332–341. https://doi.org/10.1016/j.knosys.2018.08.036

Morin-Crini N, Lichtfouse E, Fourmentin M, Ribeiro ARL, Noutsopoulos C, Mapelli F, Fenyvesi É, Vieira MGA, Picos-Corrales LA, Moreno-Piraján JC, Giraldo L, Sohajda T, Huq MM, Soltan J, Torri G, Magureanu M, Bradu C, Crini G (2022) Removal of emerging contaminants from wastewater using advanced treatments: a review. Environ Chem Lett 20:1333–1375. https://doi.org/10.1007/s10311-021-01379-5

Of D, Gui-Bai LI, Yong-Ping T (2001) An on-line monitoring technique for transmitted light fluctuation in high-turbidity water and its application(continued). Ind Water Wastewater 32(2):1–3

Pan J, Qu L, Peng K (2021) Sensor and actuator fault diagnosis for robot joint based on deep CNN. Entropy 23(6):751. https://doi.org/10.3390/e23060751

Sun L, Zhong C (2002) Application of automatic control system of coagulant dosage in mine water treatment. Coal Mine Autom. https://doi.org/10.3969/j.issn.1671-251X.2002.01.010

Traore BB, Kamsu-Foguem B, Tangara F (2018) Deep convolution neural network for image recognition. Ecol Inform 48:257–268. https://doi.org/10.1016/j.ecoinf.2018.10.002

Turaga P, Anirudh R, Chellappa R (2020) Manifold learning. computer vision: a reference guide, 1–6

Van Der Maaten L (2014) Accelerating t-SNE using tree-based algorithms. J Mach Learn Res 15(1):3221–3245. https://doi.org/10.5555/2627435.2697068

Wang CG (2005) Control mode and optimization of the coagulate dosage technique of fluctuation of transmitted light. Environ Technol 6:14–17. https://doi.org/10.3969/j.issn.1004-7204.2005.06.004r

Wong LJ, Headley WC, Andrews S, Gerdes RM, Michaels AJ (2018) Clustering learned cnn features from raw i/q data for emitter identification, pp. 26–33, IEEE. https://doi.org/10.1109/MILCOM.2018.8599847

Worm G, Van der Helm A, Lapikas T, Van Schagen K, Rietveld L (2010) Integration of models, data management, interfaces and training support in a drinking water treatment plant simulator. Environ Modell Softw 25(5):677–683. https://doi.org/10.1016/j.envsoft.2009.05.011

Xie X, Wang J, Hu F (2015) An improved floc image segmentation algorithm based on particle swarm optimization and entropic. J Comput Inform Syst 11(6):2113–2120. https://doi.org/10.13762/jcis13762

Yamamura H, Putri EU, Kawakami T, Suzuki A, Ariesyady HD, Ishii T (2020) Dosage optimization of polyaluminum chloride by the application of convolutional neural network to the floc images captured in jar tests. Sep Purif Technol 237:116467. https://doi.org/10.1016/j.seppur.2019.116467

Yu W, Gregory J, Campos LC, Graham N (2015) Dependence of floc properties on coagulant type, dosing mode and nature of particles. Water Res 68:119–126. https://doi.org/10.1016/j.watres.2014.09.045

Yuan D, Li X, He Z, Liu Q, Lu S (2020) Visual object tracking with adaptive structural convolutional network. Knowledge-Based Syst 194:105554. https://doi.org/10.1016/j.knosys.2020.105554

Zaque RAM, da Silva WTP, Santos ADA (2018) Expert system for applying coagulant in water treatment: case study in Nobres (Brazil).

Water Pract Technol 13(4):832–840. https://doi.org/10.2166/wpt.2018.080

Zhang G (2005) Coagulant dosing control BP ANFIS feed-forward water treatment. University of Electronic Science and Technology of China, Chengdu

Zhu G, Chen J, Zhang S, Zhao Z, Luo H, Hursthouse AS, Wan P, Fan G (2022) High removal of nitrogen and phosphorus from black-odorous water using a novel aeration-adsorption system. Environ Chem Lett. https://doi.org/10.1007/s10311-022-01427-8