Journal of Hydroinformatics

Journal of Hydroinformatics Vol 25 No 6, 2660 doi: 10.2166/hydro.2023.215

Development of a lightweight convolutional neural network-based visual model for sediment concentration prediction by incorporating the IoT concept

Cheng-Chia Huang[a,*], Che-Cheng Chang[b], Chiao-Ming Chang[b] and Ming-Han Tsai[b]

[a] Department of Water Resources Engineering and Conservation, Feng Chia University, Taichung City, Taiwan
[b] Department of Information Engineering and Computer Science, Feng Chia University, Taichung City, Taiwan
*Corresponding author. E-mail: cchiahuang@fcu.edu.tw

ABSTRACT

Sediment concentration (SC) monitoring has always been a pressing issue in water resource management, as many existing instruments still face challenges in accurately measuring due to environmental factors and instrument limitations. A robust technology is worth presenting to apply in the field site. This study firstly uses mean-absolute-error (MAE), root-mean-square error (RMSE), correlation coefficient (CC), and Nash–Sutcliffe efficiency coefficient (NSE) to describe the performance of the proposed convolutional neural network (CNN). Moreover, adapting the ensemble learning concept to compare the multiple machine learning (ML) approaches, the CNN presents the highest predicted accuracy, 91%, better than SVM (79%), VGG19 (63%) and ResNet50 (35%). As a result, the proposed CNN framework can appropriately apply the monitoring needs. The primary purpose is to develop a simple, accurate, and stable SC monitoring technology. Instead of some complex architectures, a simple and small neural network is adopted to implement real-time application (RTA). Via our design, such a traditional but critical issue can be improved to a new state. For example, by incorporating the concept of the Internet of Things (IoT) with our design, the distributed computing system for large-scale environmental monitoring can be realized quickly and easily.

Key words: Convolutional neural network (CNN), distributed computing system, Internet of Things (IoT), real-time application (RTA), sediment concentration (SC)

HIGHLIGHTS

- Safety: Contactless technology.
- Simple: Improving the computation efficiency for real-time prediction.
- Accuracy: Presenting satisfied accuracy by comparing it with the current technology.
- Stability: Overcoming outlier trouble in the prediction process to show sufficient stability.
- Development: Incorporating the IoT to develop the large-scale environmental monitoring platform quickly and easily.

1. INTRODUCTION

Flood damage is regarded as the most frequent natural disaster. Besides, the vast sedimentation is often carried by the enormous flooding due to broken watersheds and causes debris flow, reservoir siltation, river erosion and deposition, and water quality deterioration. This compound disaster frequently occurs in the current period because climate change contributes to dramatic flooding and sedimentation issues. Many studies have been conducted on sediment transportation during flood events. Still, several disturbing factors in field sites cause incomplete data collection, such as adverse weather conditions and the limitations of the measuring instrument. Despite developing early warning systems for decades, the accuracy has yet to be effectively improved due to insufficient sediment data. Above all, a fast, accurate, and reliable measurement technology needs to be proposed to reduce the uncertainty of the collected data and further supply sufficient data to improve the sedimentation prediction ability.

Traditional technology in sediment concentration (SC) measurement in the field presents challenges because the data representation often occurs during flooding. It is challenging to collect, and the above paragraph shows the reason. The measuring approaches for SC are divided into direct sampling and indirect measuring. Suction sampling is a commonly used direct method that is inexpensive and simple (Nielsen 1984; Bosman *et al.* 1987). The accuracy of the suction sampling

method is high; however, it is time-consuming because the water samples need to be analyzed in the laboratory, thus being hard to apply in the field site for large-scale, real-time monitoring during flood periods. Indirect measuring methods have been developed in the past decades to address these limitations (Wren *et al.* 2000; Barua *et al.* 2001). The acoustic method for measuring SC is a contactless technique that uses the decay rate of sound frequency to estimate the concentration of suspended particles. However, the accuracy is limited by the need for prior calibration and a strong relationship between the reflection signal and concentration. In other words, it requires extensive laboratory work, and the established relationship may be difficult to transfer to other locations (Thorne *et al.* 1991; Lohrmann 2001; Landers *et al.* 2016; Sirabahenda *et al.* 2019; Lin *et al.* 2020). Laser diffraction is a technique for measuring SC that uses light. The laser emits a beam into the sample, scattered, reflected, and absorbed by suspended particles. The method determines the concentration by analyzing the volume of the particle. However, laser diffraction needs to be calibrated and limited to a single-point measurement (Agrawal & Pottsmith 1994; Guillén *et al.* 2000; Merten *et al.* 2014; Su *et al.* 2016; Coleman *et al.* 2020). Time domain reflectometry (TDR) is another technique for measuring SC. It uses the relationship of different dielectrics and sediment samples to estimate the concentration. TDR has been applied to sediment observations in vertical profiles, but the limitation of TDR is a significant error in low-concentration environments (Chung *et al.* 2013; Mishra *et al.* 2018; Miyata *et al.* 2020; Chung & Wang 2022). Except for the above measuring technique, the numerical model is considered convenient to solve the SC issue, and the improvement of the SC estimation was gradually obtained. However, the accurate concentration value is challenging to simulate in the numerical model due to parameter settings in the different field sites. Additionally, the mentioned approaches are difficult for real-time monitoring (Fang & Wang 2000; van Maren *et al.* 2015; Huang *et al.* 2019; Orseau *et al.* 2021). Note that a proper measuring technology needs to be developed, and the interdisciplinary artificial intelligence (AI) is an appropriate way to optimize the current water resources management method.

In hydraulic engineering, AI has been applied to estimate the time-space distribution, surface runoff, groundwater level, and water quality (Khan *et al.* 2019; Hasda *et al.* 2020; Ighalo *et al.* 2021; Iriany *et al.* 2022). Moreover, several studies mentioned the monthly, weekly, or daily SC value estimation by using different ML algorithms (AlDahoul *et al.* 2021, 2022; Ehteram *et al.* 2021; Fan *et al.* 2023; Latif *et al.* 2023). The time scale is the other important issue, and several papers applied to predict the short lead-time SC value and obtained reasonable results (Chang *et al.* 2020; Huang *et al.* 2021; Jhong *et al.* 2021). Computer vision is a field of AI that enables a computer to derive meaningful information from digital images, videos, and other visual inputs that may not be found with the naked eye. For example, convolutional neural network (CNN) convolves the data into a smaller area to extract features (Hwang 2017; Hwang & Chen 2017), e.g., colors, edges, textures, and so on. Via the combinations of these features, we can detect, recognize, or predict some further events. For instance, Chang *et al.* (2021) propose an architecture based on CNN to detect vehicles and predict the future position of each vehicle for collision avoidance. Next, we can train a machine to perform these ML functions in much less time than the procedure of retinas, optic nerves, and visual cortices. For example, CNN is also one of the fastest and most accurate algorithms for image applications, which can inspect a large number of images per minute (Hwang 2017; Hwang & Chen 2017; Chang *et al.* 2022). Obviously, it surpasses human capabilities. On the basis of the above discussions, we can realize that CNN suits the estimation of SC by using different images. More specifically, the variation of the SC needs to observe the details of colors and textures of water and inspect the sample fast, accurately, and stably, and both are the capabilities of CNN. Above all, the above references indicate that accurate SC estimation is a crucial issue, and the current technology is still under the improvement process. The ultimate goal should establish a robust SC measuring approach to apply to the real-time warning in the field site.

This study aims to design a computer vision method based on the technique of CNN to provide the estimation of SC with accuracy and stability. Instead of some complex architectures, a simple and small neural network is adopted to improve the computation efficiency. The proposed CNN-based prediction model is more suitable for real-time prediction, which is the first contribution. Secondly, since our network only has a small number of trainable parameters, it is easier to be ported to embedded systems in practice. By comparing several famous ML algorithms, an accurate, efficient, and stable real-time approach can be proposed to predict the turbidity water concentration for water resources management. Most importantly, such a traditional but critical issue can be improved to a new state via our design. It is suitable to incorporate IoT and develop a distributed computing system for large-scale environmental monitoring in the future. For example, in Figure 1(a), a camera sensor is mounted on an embedded system to perceive the environment. Also, there is a GPS (global positioning system) component to position the system and a battery to provide power. All devices are placed in a waterproof case. Next, in Figure 1(b), we can spread a number of systems to an area that we intend to monitor. Via the ability of wireless communication, we can remotely receive the perceptions from all systems to continuously monitor the environment. Obviously, this design helps us

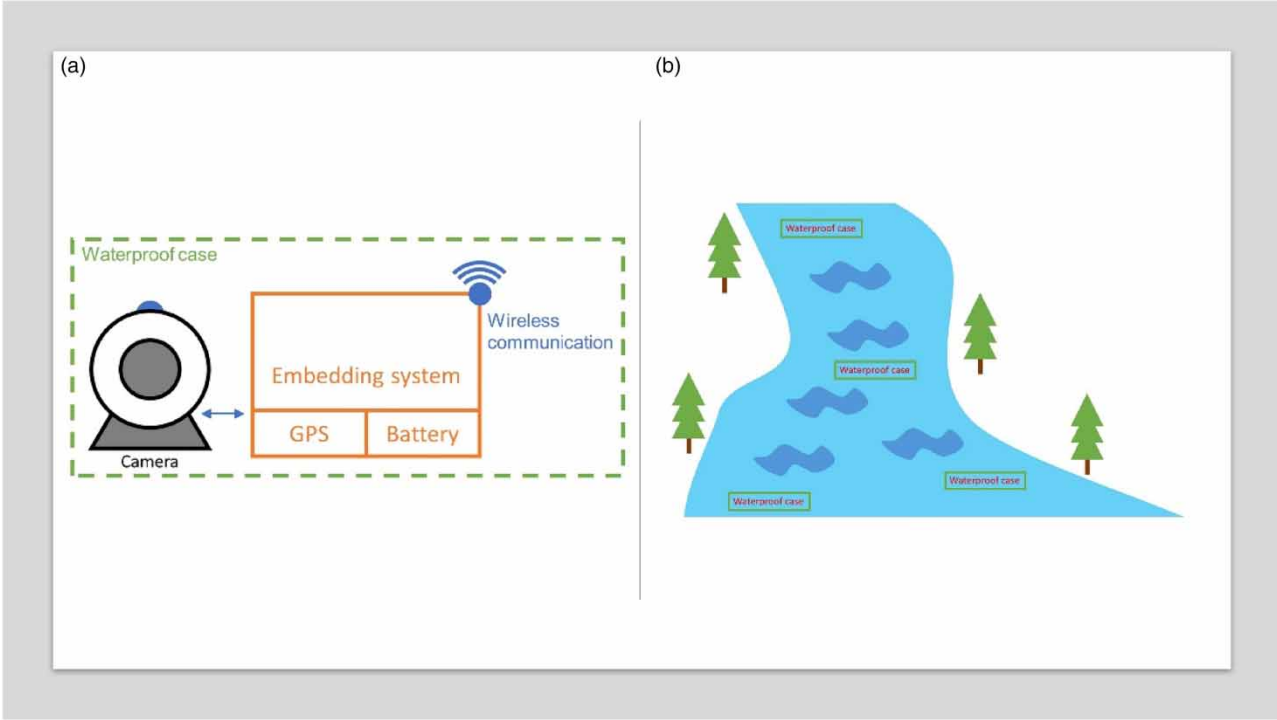

Figure 1 | An example for monitoring the environment: (a) a feasible architecture and (b) spreading cases to the environment.

reduce labor costs since we can remotely receive the perceptions from all systems to continuously monitor the environment. In addition, thanks to the architecture of the embedded system and our lightweight neural network, our system has lower power consumption to make it able to function for longer.

2. METHODOLOGY

Traditional technology in SC measurement in the field presents challenges because data representation often occurs during heavy rainfall or typhoon events. It is difficult to collect and hard to be applied for large-scale and real-time monitoring during flood periods. Therefore, we designed a novel method based on CNN to properly deal with the SC issue estimation.

2.1. The neural network

A neural network is an ML model which is inspired by the architectures of neurons. Owing to the properties of NN, it can be collocated with other techniques to cope with large and complicated issues, e.g., classifying images, speech recognition, optimization, and so on those references (Pedregosa *et al.* 2011; Chollet 2015; Chang *et al.* 2022; Géron 2022). Namely, it is versatile, scalable, and powerful.

In the field of NN, one of the commonest architectures is Perceptron. For each neuron, the inputs and sole output are numbers. Moreover, each input is associated with a weight and an extra bias. Particularly, the weighted sum is applied to an activation function to produce the output. Next, multi-layer perceptron (MLP) is an improved version of the classical perceptron, which is made up of an input layer, hidden layer(s), and an output layer, where every layer except the unique output layer is fully connected to the next layer. Notice that the hidden layer is constructed from the notion of classical perceptron (Chang *et al.* 2022; Géron 2022). The computation of classical perceptron is shown in the following equations:

$$y_i = x_{n-1}w_{n-1} + \ldots + x_0 w_0 + b_i \tag{1}$$

$$z_i = act_i(y_i) \tag{2}$$

where y_i is the weighted sum; x is the input; w is the weight of x; b_i is the bias; z_i is the output produced from the activation function $act_i()$.

2.2. The convolutional layer

The idea of convolution is inspired by the organization of the animal visual cortex and works well (Géron 2022). Briefly, individual neurons respond to the stimuli in a receptive field. Since the receptive fields of different neurons partially overlap, they finally tile the entire visual field. This can be approximated by the operation of convolution (Hwang 2017; Hwang & Chen 2017). On the other hand, the convolutional layer is generally paired with the pooling layer to reduce the dimensions by combining the outputs of neuron clusters in the previous layer into one single neuron in the next layer. Namely, the design can reduce the quantity of weights and complexity of calculations compared with traditional neural layers in the network. Hence, we can add some convolutional layer(s) as well as the pooling layer(s) in the first half of the neural network to introduce its advantages to the traditional neural network.

Technically speaking, a convolutional neural network is an architecture composed of at least some convolutional layer(s) as well as pooling layer(s) along with a neural network, and the two kinds of layers often pair in practice. Thus, we can use several simple equations to describe the computational cost via the change of neurons in successive layers. Assume that an image with $n*n$ pixels is considered here. If w filters with $m*m$ pixels and the pooling region with $k*k$ are used in the convolutional layer and pooling layer, we need $w*[(n-m+1)/k]*[(n-m+1)/k]$ neurons in the successive layer to receive such output.

2.3. The comparison among three well-known classifiers and ours

In this section, we introduce three well-known classifiers that are adopted to compare with ours for the purpose of showing the superiority of our method, i.e., SVM (support vector machine), VGG19 (Visual Geometry Group 19), and ResNet50 (Residual Network 50).

Firstly, SVM is a classical ML model capable of performing classification for small- or medium-sized datasets (Géron 2022). It constructs a hyper-plane or set of hyper-planes in a high dimensional space, which can be utilized for classification. Generally, when the problem is not separable, the support vectors are the samples within the margin boundaries. Here is the brief mathematical formulation: given training samples x_i and a vector y, the goal is to find w and b such that the prediction given by $\boldsymbol{sign}(\boldsymbol{w}^T\emptyset(\boldsymbol{x_i}) + \boldsymbol{b})$ is correct for most samples. Furthermore, if the number of classes is greater than 2, the strategy of OVR (One-Vs-Rest) is utilized to help us reach the goal (https://scikit-learn.org/stable/modules/svm.html#svc).

The next one is VGG19 from the VGG research laboratory at Oxford University, which also has a simple architecture (Simonyan & Zisserman 2015; Géron 2022). Briefly, there are a few convolutional layers with a pooling layer repeatedly until reaching a total of 16 or 19 convolutional layers that depend on different variants. The last one is ResNet50, which is based on the residual network architecture (He *et al.* 2015; Géron 2022). Noteworthily, like other variants with 101 and 152 layers, the ResNet architectures use an extremely deep design. Note that since the inputs and outputs of our work are different from the two originals, we have to modify them to fit our requirements. Next, for these two originals, using only our sediment samples is not enough to train them and causes very terrible performance. Hence, we incorporate the concept of transfer learning based on the concept of pre-training weight with ImageNet. We freeze several upper convolutional layers to keep the characters of the originals and only retrain the successive layers to fit our requirements such as 16 classes.

As a result, we choose a classical method, a shallower convolution neural network, and a deeper convolution neural network to compare with ours and give comprehensive discussions.

2.4. Statistical indicators

The clear water is considered as 0 ppm, and it is set as the initial value. During normal periods, a minimal amount of sediment is suspended in the river water. An SC below 100 ppm is considered relatively clear water and has been the lower limit in this study. During flood events, the SC increases significantly, and water intake treatment should be halted if the SC exceeds 6,000 ppm. This study has raised the upper experiment limit to 7,800 ppm to enhance practicality. Hence, the SC prediction of different samples from 100 to 7,800 ppm (mg/l) is desired in this research. The first experiment used clear water (0 ppm) to capture background images, followed by sequentially capturing images with increasing SC. Additionally, sediment was added each time and thoroughly mixed using a stirring rod before capturing the images. The photography environment was maintained at a constant temperature (28 °C) and consistent lighting conditions (Figure 2) to ensure uniform image quality throughout the captures.

Four frequently used statistical indicators, MAE, RMSE, CC, and NSE, are adopted as the objective indicators for the accuracy prediction between observation data and CNN. The above statistical indicators are often used to evaluate the prediction

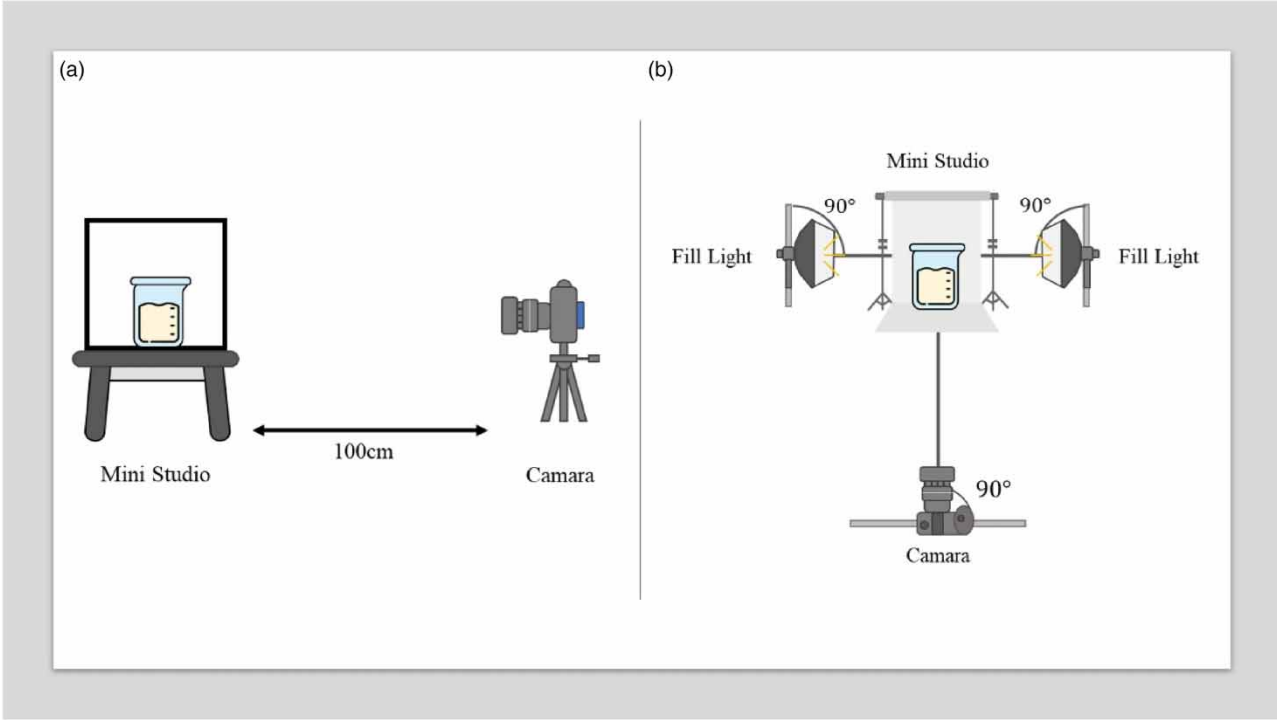

Figure 2 | Camera and scenario set-up: (a) horizontal view and (b) top view.

ability of the AI model. MAE and RMSE are supposed to be positive values, and a lower value means better performance of SC estimation. The range of CC is between −1 and 1, representing a negative correlation to a positive correlation. In other words, the perfect prediction value of CC equals 1. NSE is determined by subtracting the error variance ratio in the model from one divided by the variance in the observation. The lower value indicates poor prediction performance, and the perfect prediction of the NSE value is 1. The mentioned statistical indicators are shown in the following equations

$$\mathrm{MAE} = \frac{1}{N}\sum |(P_i - M_i)^2| \tag{3}$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(P_i - M_i)^2}{N}} \tag{4}$$

$$\mathrm{CC} = \frac{\sum_{i=1}^{n}(P_i - \bar{P})(M_i - \bar{M})}{\sqrt{\sum_{i=1}^{n}(P_i - \bar{P})^2 \sum_{i=1}^{n}(M_i - \bar{M})^2}} \tag{5}$$

$$\mathrm{NSE} = 1 - \frac{\sum_{i=1}^{N}(P_i - M_i)^2}{\sum_{i=1}^{N}(M_i - \bar{M})^2} \tag{6}$$

where N is the sample number; P and M are the prediction and measurement values, respectively; $\bar{S}_p$ and $\bar{M}$ are the means of prediction and measurement value, respectively; C_{h} and C_{m} means 'hit' and 'miss' which are determined as stated above.

3. EXPERIMENTAL SET-UP

3.1. Camera and scenario set-up

In this work, FUJIFILM X-S10 is used to take all experimental photos, which is an APS-C (Advanced Photo System Classic) camera. The settings are shown below:

- Resolution: 6,240*4,160;
- Aperture: f/8;
- Shutter speed: 1/25 s;
- ISO: 200;
- Focal length: 35 mm;
- Color temperature: 4,000;
- Mode: manual mode.

On the other hand, the experimental scenario is illustrated in Figure 2. The camera is set horizontally to the sample in the film studio, where the distance is 100 cm from the former to the latter, which is shown in Figure 2(a). Then two fill lights are set horizontally on both sides of the film studio to reduce the contrast to match the dynamic range of the photo. Namely, it is for the purpose of recording the same number of details seen by humans (Figure 2(b)).

3.2. Experimental set-up

Except for those expensive surveying instruments, most portable devices have a tolerance of $+/-$ 250 ppm. Hence, we set goals to design a classifier for the interval of 500 ppm. We prepare 60 samples for the training and testing procedures, ranging from 100 to 7,800 ppm for 100–150 ppm random increments, where the resolution of each image is 6,240*4,160. Next, we do not use a whole image as a training or testing sample since it will induce a very complex network, which is unsuitable for the embedded system. According to the characteristics of CNN, we use the concept of a sliding window to segment images to obtain more samples and keep the size of the network. Accordingly, the number of original samples is enough for our experiment. The sample pictures are shown in Figure 3. In addition, the adopted calculated formula of SC is shown below.

$$S_s = \frac{W_s}{V_w} \tag{7}$$

where S_s refers to sedimentation (mg/l); W_s refers to sediment weight; V_w refers to water volume.

3.3. The design of CNN

The design of CNN is illustrated in Figure 4. The input is a part of the original color image, where the resolution is 224*224 pixels. Noteworthily, in accordance with our experiences, utilizing a part of the original color image as the input helps us reduce the computational complexity of the model and keep the critical features of the original color image (Géron 2022). Next, the output is set to 16 neurons to classify samples into 16 categories with 500 ppm intervals from 0 to 8,000 ppm. During the training procedure, several different kinds of layers are used, i.e., two pairs of convolutional layers and pooling layers for feature extraction and computation reduction, three dropout layers for overfitting prevention, and the unique flatten layer for reshaping the data stream.

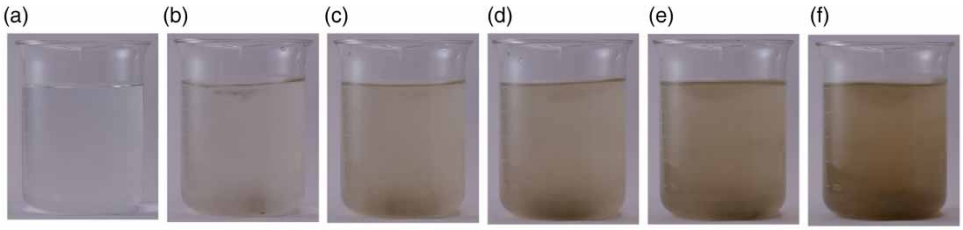

Figure 3 | Samples with various mixtures of sediment and water body: (a) 0 ppm; (b) 1,858.6 ppm; (c) 3,516.6 ppm; (d) 4,246 ppm; (e) 5,530.4 ppm; (f) 7,834.5 ppm.

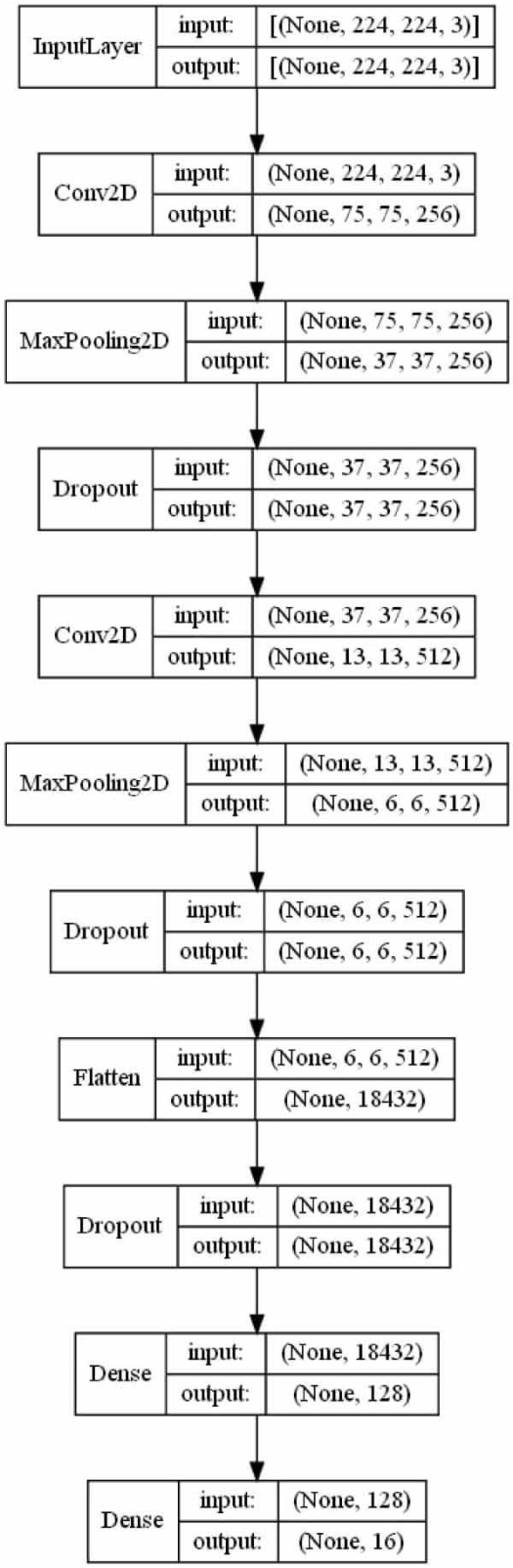

Figure 4 | The design of a convolutional neural network.

Next, the detailed settings of the learning procedure are given below:

- Batch size: 128;
- Dropout: 0.2;
- Learning rate: 0.001;
- Optimizer: Adam;
- Beta_1: 0.9;
- Beta_2: 0.999;

4. RESULTS AND DISCUSSION

4.1. Training and testing procedure

Firstly, we have 60 images captured by the camera. Via the technique of sliding window with a small range, 277,440 sub-images are obtained. Then, these sub-images are separated into 80% training data and 20% testing data for the purpose of overfitting prevention. Based on the CNN model introduced in the previous section, the hyperparameter, epochs, with four various settings are considered, i.e., 50, 100, 150, and 200. The line charts of the training procedure with different epochs are shown in Figure 5(a)–5(d). Here, we consider five successive data as a whole, and the reason is that doing so can present the experimental results more clearly. According to the training results, the accuracy at the end of all training procedures is 0.8634, 0.9061, 0.9111, and 0.9183, respectively. Next, via the testing data to realize the validation again, the testing accuracy of four various settings is 0.8493, 0.9125, 0.9054, and 0.8803, respectively, as shown in Figure 6(a)–6(d). Note that both training and testing procedures show sufficient accuracy; all values are higher than 0.84.

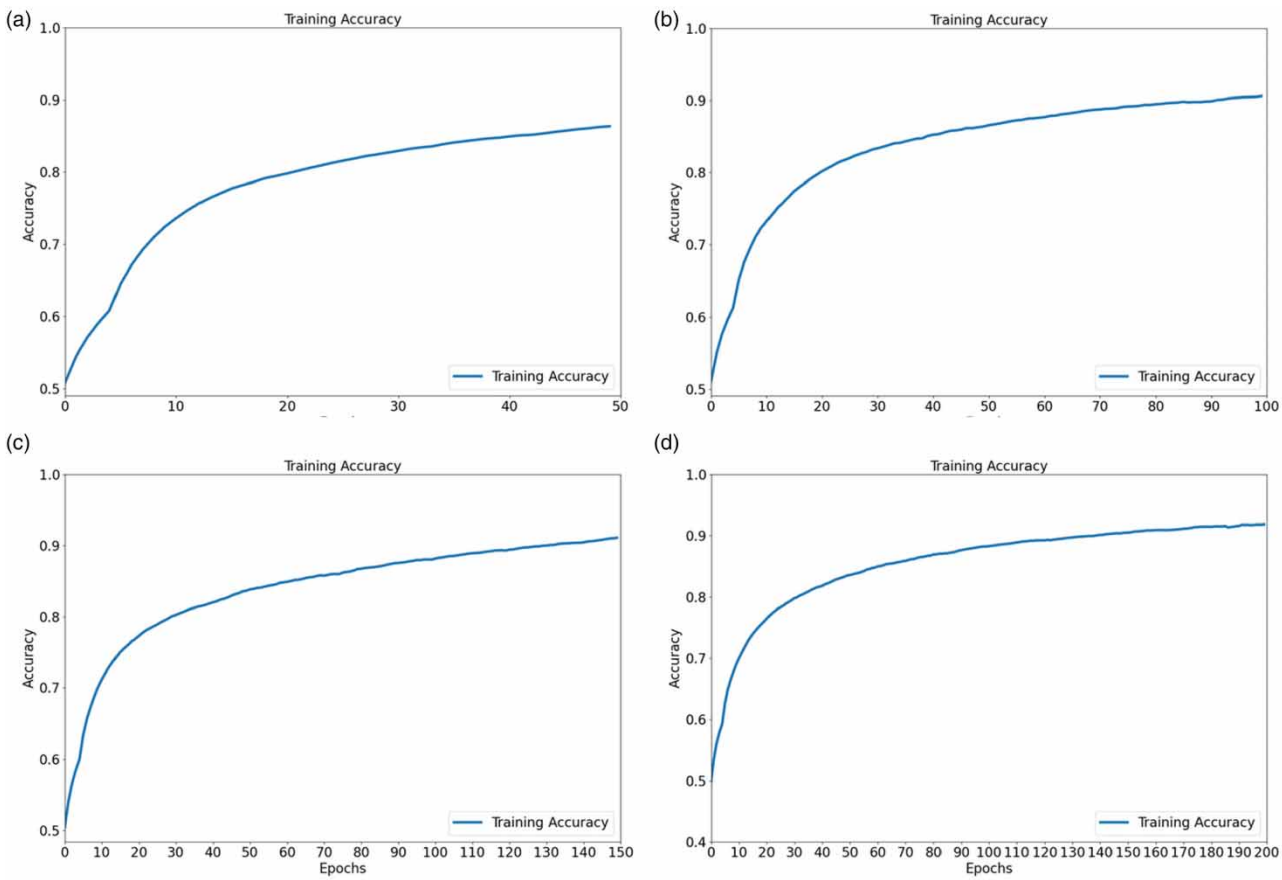

Figure 5 | The line chart of the training procedure with (a) 50; (b) 100; (c) 150; (d) 200 epochs.

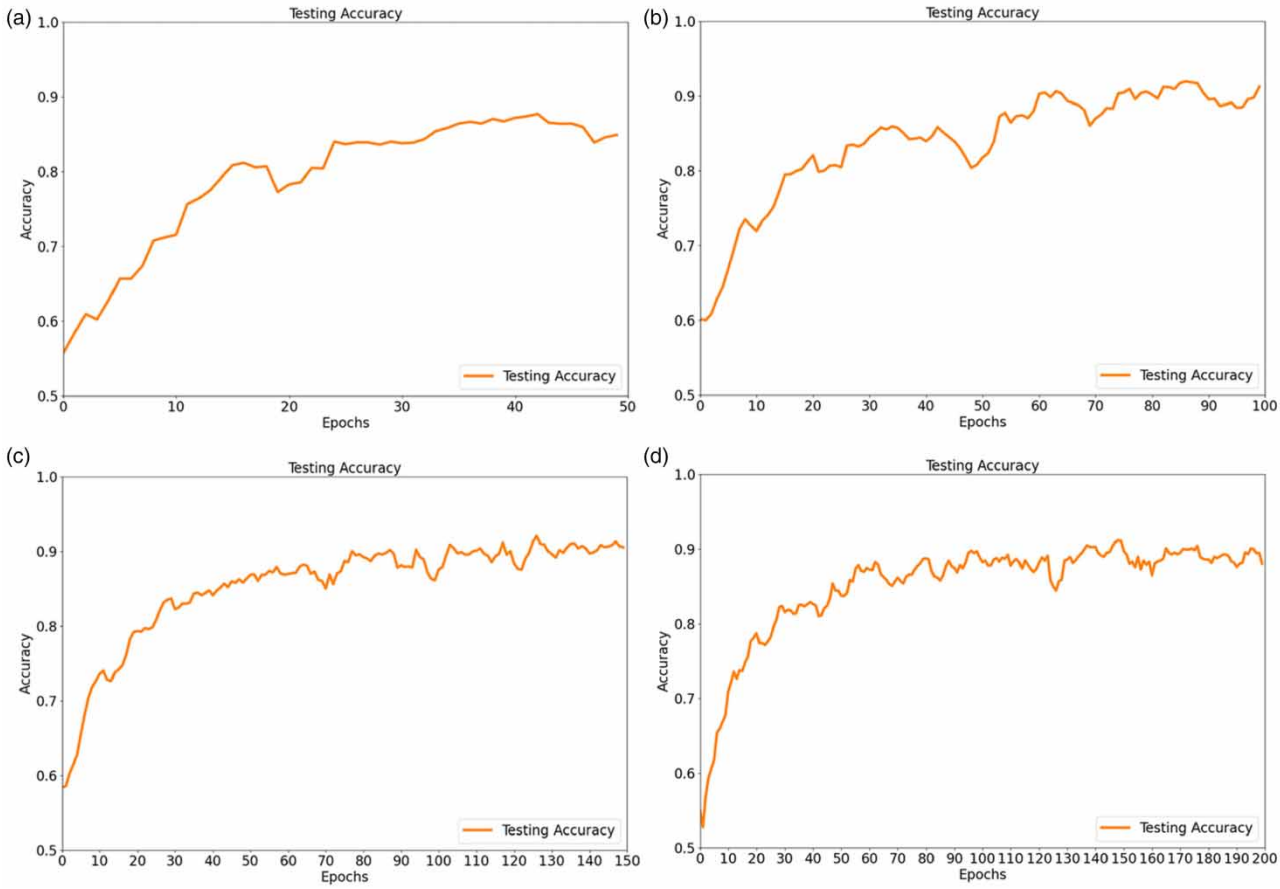

Figure 6 | The line chart of the testing procedure with (a) 50; (b) 100; (c) 150; (d) 200 epochs.

Figure 7 shows the training and testing procedure line chart with 50, 100, 150, and 200 epochs. The epoch value presents a continued rising trend of prediction accuracy in the training procedure. In addition, the performance indicated that the epoch value is significantly improved from 50 to 100, and the accuracy improvement becomes insignificant when the epochs are larger than 100. The 100 epochs should be the better selection because the latter three are almost identical. The same trend in the testing procedure shows the improvement of the prediction accuracy from 50 to 100 epochs is more significant than others. The results meet the property that increasing the number of epochs may cause overfitting of the CNN model. In accordance with the above description, a significant result can be proposed that the training and testing procedure with 100 epochs is suitable among all procedures from the viewpoint of accuracy.

On the other hand, this study draws the heatmaps based on (a) training and (b) testing data and corresponding predictions under 100 epochs in order to observe whether the position of a sample affects the accuracy or not (Figure 8). According to the results, we can observe that the properties of the two heatmaps are very similar. Note that training and testing data are between 0.7833 and 1.000, indicating a highly related relationship. More specifically, the experiments using the bottom of training/testing samples possess higher performance averagely and vice versa. The intuition is that the settling of suspended sands in water is formed based on the density difference between the suspended sands and water. Hence, sands are caused by gravity to settle downward until reaching the sediment layer within the water. Namely, the samples in the bottom layer are mixed better and more stable.

Last but not least, our CNN has only 3,548,816 trainable parameters, making it easier to port to embedded systems in practice due to its lower computational complexity, lower latency, lower memory usage, and less power consumption. More specifically, via our experimental computer, which is with CPU: Intel Core i7-8750, RAM: 16GB DDR4, and GPU: Nvidia GeForce RTX 2060, we gather statistics on the length of completing various training and predicting procedures, i.e., 50 epochs needing 4 h

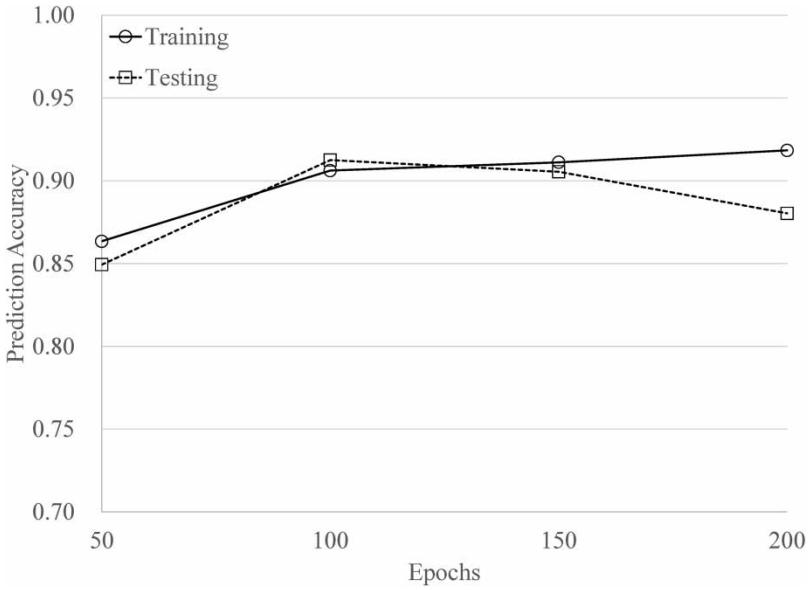

Figure 7 | The line chart of the training and testing procedure with 50, 100, 150, and 200 epochs.

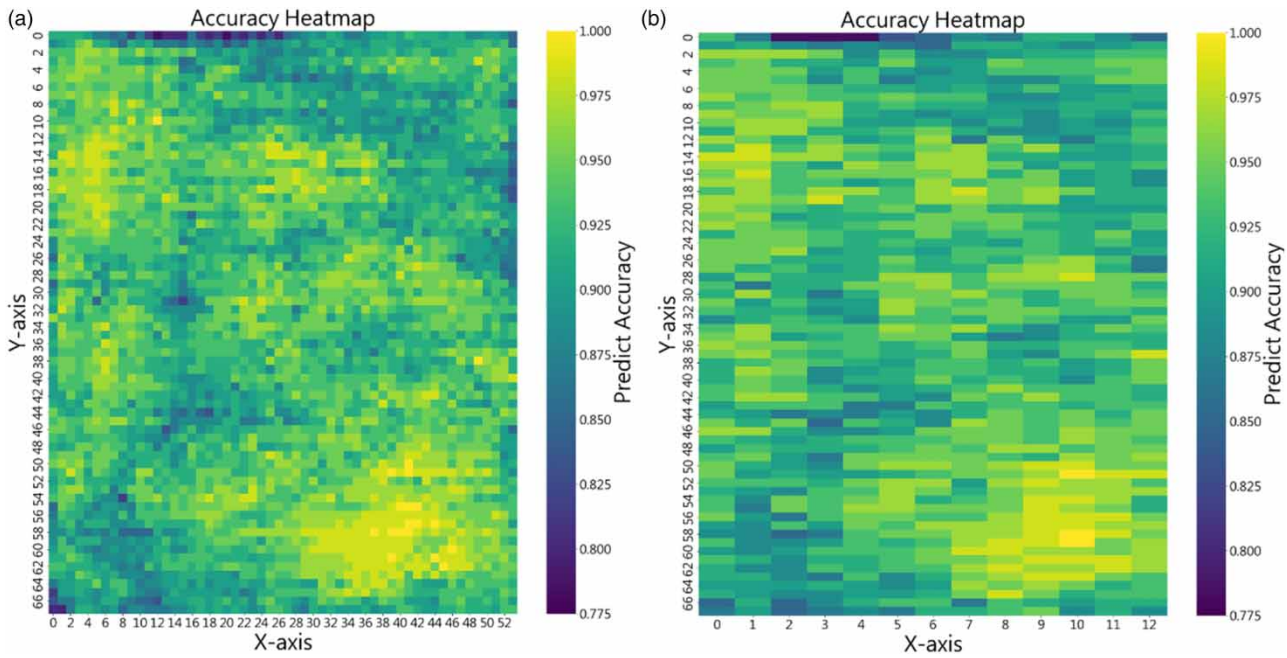

Figure 8 | The heatmaps based on (a) training (b) testing data and corresponding predictions under 100 epochs.

and 55 min, 100 epochs 9 h and 10 min, 150 epochs 13 h and 43 min, 200 epochs 20 h and 23 min, and each prediction needing only 0.07 s. A significant result shows that even the length of the training procedure with the setting of 200 epochs is affordable and acceptable, and the length of each prediction is excellent. The data, as mentioned earlier, are shown in Figure 9.

4.2. Comparison of prediction accuracy between observation and CNN

Figure 10 compares the SC of measured data and prediction value to evaluate the practicality of the proposed model. A significant result presented that the comparison points are close to the fitting line, which means the prediction model indicates a

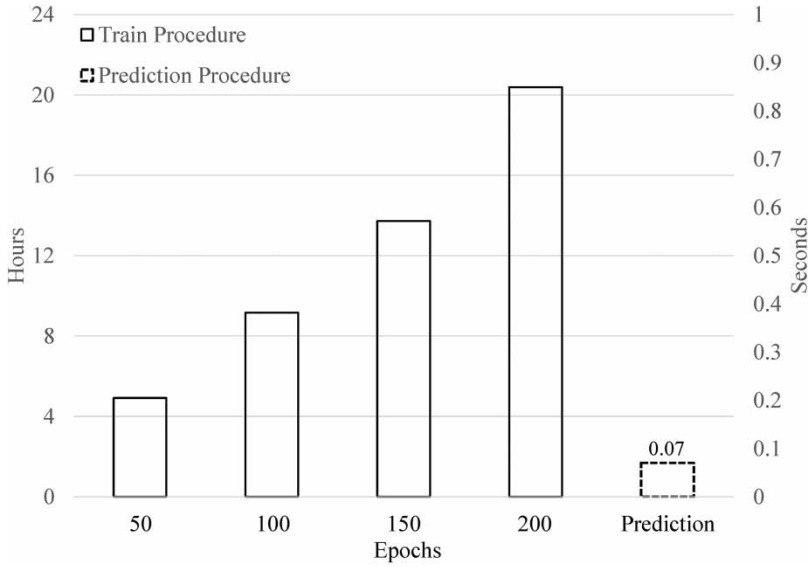

Figure 9 | Computer time length of training and predicting procedures.

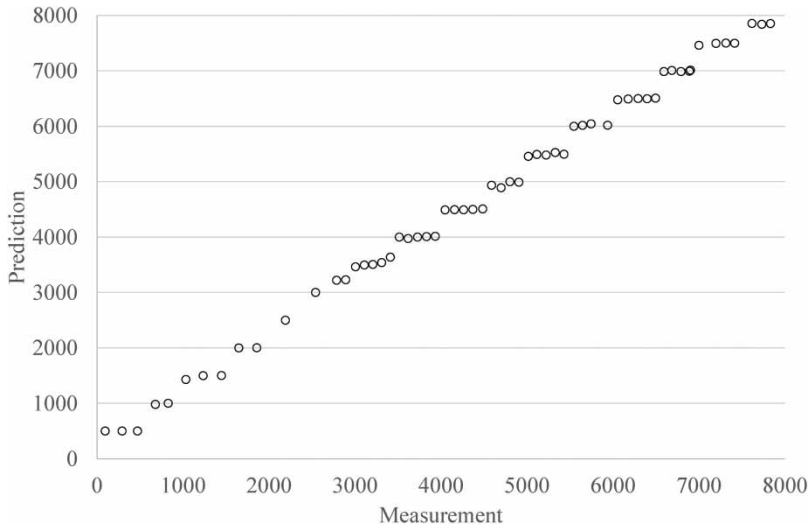

Figure 10 | Comparison of measured data and prediction model.

sufficient prediction ability. Then, based on the measured and predicted values, four statistical indexes, MAE, RMSE, CC, and NSE, were compared to further analyze the prediction performance.

In this study, grouping was conducted with intervals of 1,000 ppm. The main reason for this choice was to ensure that each subgroup contained at least five or more data points for analysis. The comparison spacing is divided into eight parts, from 0 to 1,000, 1,000 to 2,000, 2,000 to 3,000, 3,000 to 4,000, 4,000 to 5,000, 5,000 to 6,000, 6,000 to 7,000, and 7,000 to 8,000 ppm. The values of MAE are 269.17, 241.19, 384.38, 296.37, 219.37, 281.91, 217.56, and 199.44, respectively. The maximum and minimum values are 384.38 and 199.44 at 2,000–3,000 and 7,000–8,000 intervals. The same trend is presented in RMSE; the maximum and minimum values are 389.56 and 242.59. Additionally, CC and NSE are adopted to evaluate the correlation and error. The range of CC is between 0.88 and 0.99; the range of NSE is from 0.72 to 0.81. Both CC and NSE show satisfactory results. All statistical data are shown in Table 1.

This study further investigates the meaning of statistical index to the prediction performance. The highest difference of MAE is at the interval 2,000–3,000 ppm, a value of 384.83. It is a sufficient accuracy in the field site application because

Table 1 | The value of the different statistical index

Index ppm	MAE	RMSE	CC	NSE
0–1,000	269.17	310.44	0.86	0.72
1,000–2,000	241.19	272.87	0.90	0.81
2,000–3,000	384.38	389.56	0.98	0.70
3,000–4,000	296.37	319.80	0.92	0.81
4,000–5,000	219.37	253.90	0.89	0.78
5,000–6,000	281.91	312.92	0.88	0.76
6,000–7,000	217.56	254.78	0.88	0.76
7,000–8,000	199.44	242.59	0.90	0.72

the accuracy of the current instrument is 500–1,000 ppm. The maximum error of RMSE is in the interval 2,000–3,000 ppm, a value of 389.56. In addition, the minimum error value is 242.59, located at the interval 7,000–8,000 ppm. A Significant result is presented that the accuracy and the stability are acceptable because the value of MAE and the difference in different intervals of RMSE is small enough. CC is applied to assess the degree of correlation between the measurement and prediction. In addition, the ranges in value from -1 to 1. A CC closer to 1 or -1 indicates a stronger positive or negative linear correlation between the two variables, while a CC close to 0 indicates no linear correlation between the two variables. In this study, the range of CC is from 0.86 to 0.95 in different intervals, which means the measured and predicted values present a high correlation. NSE is commonly applied in the hydraulic engineering field to analyze the fitting of a prediction model to measured data. It is commonly applied in fields such as hydrology and water resources management. NSE ranges from negative infinity to 1, with a value closer to 1 indicating a better fitting. Conversely, a negative value signifies poor model performance. The range of NSE shows a good fitting in the proposed prediction model because all values are higher than 0.7, and the best performance is located at the interval 7,000–8,000 ppm with a value of 0.90.

Overall, three significant advantages are shown in the proposed model. The satisfied accuracy compared to the current instrument; the sufficient stability to apply in the field site; and last but not least, the high correlation of measurement and prediction is presented in all intervals, which means this model can be applied in 0–8,000 ppm. A practical application procedure could be formulated in the future.

4.3. Comparison of prediction performance for four prediction models

In this subsection, via the same dataset, three well-known classifiers are adopted to compare with ours, i.e., SVM, VGG19, and ResNet50. The detailed results are given below. Firstly, Table 2 shows the number of parameters of each network. Since the number of parameters affects the time of the predicted procedure, we can observe that the methods need a similar ratio of time to finish their predicted procedure. This conforms with the characteristics of neural networks. Note that the specifications of the experimental computer are CPU: Intel Core i7-97,003.00 GHz, RAM: 16 GB DDR43, 200 MHz, and GPU: Nvidia GeForce RTX 2080 8 GB. Next, according to the results in Table 2, we can see that our method demonstrates its superiority in predicted accuracy as well.

Table 2 | Comparison of parameters and the time of the predicted procedure

Model	The number of parameters	The time of the predicted procedure (s)	Predicted accuracy (%)
Our CNN with 100 epochs	3,548,816	0.07	91
SVM	Not applicable	0.73	79
VGG19	20,032,592	0.09	63
ResNet50	23,620,496	0.09	35

5. CONCLUSION

This study firstly considers the SC issue as a classification topic for 16 classes with the lightweight CNN. Our design and experiments illustrate that the CNN-based method for SC prediction during flood periods is feasible and works well. The proposed CNN model can potentially develop a practical tool in the field site due to the following advantages.

Firstly, the proposed CNN-based model should be compared with the well-known classification algorithms instead of regression methods. This is why we chose SVM, VGG19, and ResNet50. This prediction model presents satisfied accuracy by comparing it with the current technology. The proposed CNN model performs better than the other well-known models, SVM, VGG19, and ResNet50. In addition, sufficient accuracy is presented by comparing it to the current technology (lower than 500 ppm). Second, this model overcomes outlier trouble often occurring in the prediction process and shows sufficient stability. Third, it can be a simple network architecture and training procedure. Since our network only has a small number of trainable parameters, it is easier to be ported to embedded systems in practice. Most importantly, via our design, such a traditional but critical issue can be improved to a new state. This study proposes a practical concept by incorporating the Internet of Things (IoT) with our design, a contactless technology, the distributed computing system for large-scale environmental monitoring can be realized quickly and easily.

This study is worth continuing, such as challenging the higher concentration prediction, more accurate performance, and different particle sizes/shapes of the sediment. Furthermore, developing a real-time response system for the prediction of the turbidity water to improve the water resources management potential. The above vision will be presented in the future study.

ACKNOWLEDGEMENT

The financial support for this study was provided by a grant from the National Science and Technology Council (project number: 111-2625-M-035-007-MY3).

DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

CONFLICT OF INTEREST

The authors declare there is no conflict.

REFERENCES

Agrawal, Y. C. & Pottsmith, H. C. 1994 Laser diffraction particle sizing in STRESS. *Continental Shelf Research* **14** (10), Article 10. https://doi.org/10.1016/0278-4343(94)90030-2.

AlDahoul, N., Essam, Y., Kumar, P., Ahmed, A. N., Sherif, M., Sefelnasr, A. & Elshafie, A. 2021 Suspended sediment load prediction using long short-term memory neural network. *Scientific Reports* **11** (1), 7826.

AlDahoul, N., Ahmed, A. N., Allawi, M. F., Sherif, M., Sefelnasr, A., Chau, K. & El-Shafie, A. 2022 A comparison of machine learning models for suspended sediment load classification. *Engineering Applications of Computational Fluid Mechanics* **16** (1), 1211–1232.

Barua, D. K., Shen, C., Lemmin, U., Wren, D. G., Barkdoll, B. D., Kuhnle, R. A. & Derrow, R. W. 2001 Field techniques for suspended-sediment measurement. *Journal of Hydraulic Engineering* **127** (9), Article 9. https://doi.org/10.1061/(ASCE)0733-9429(2001)127:9(784).

Bosman, J. J., van der Velden, E. T. J. M. & Hulsbergen, C. H. 1987 Sediment concentration measurement by transverse suction. *Coastal Engineering* **11** (4), Article 4. https://doi.org/10.1016/0378-3839(87)90033-0.

Chang, M.-J., Lin, G.-F., Chen, P.-A., Lee, F.-Z. & Lai, J.-S. 2020 Development of a real-time forecasting model for turbidity current arrival time to improve reservoir desilting operation. *Hydrological Sciences Journal* **65** (6), 1022–1035.

Chang, C.-C., Ooi, Y.-M. & Sieh, B.-H. 2021 IoV-based collision avoidance architecture using machine learning prediction. *IEEE Access* **9**, 115497–115505.

Chang, C.-C., Ooi, Y.-M., Tsui, S.-T., Chiang, T.-H. & Tsai, M.-H. 2022 Utilizing ensemble learning to improve the distance information for UWB positioning. *Applied Sciences* **12** (19), 9614.

Chollet, F. 2015 Keras documentation. *Keras. Io*, **33**.

Chung, C.-C., Lin, C.-P., Wu, I.-L., Chen, P.-H. & Tsay, T.-K. 2013 New TDR waveguides and data reduction method for monitoring of stream and drainage stage. *Journal of Hydrology* **505**, 346–351.

Chung, C.-C. & Wang, Y.-K. 2022 Practical assessment of real-time suspended sediment load monitoring using time domain reflectometry. *Water Resources Research* **58** (9), e2022WR032289.

Coleman, D. J., Ganju, N. K. & Kirwan, M. L. 2020 Sediment delivery to a tidal marsh platform is minimized by source decoupling and flux convergence. *Journal of Geophysical Research: Earth Surface* **125** (8), Article 8. https://doi.org/10.1029/2020JF005558.

Ehteram, M., Ahmed, A. N., Latif, S. D., Huang, Y. F., Alizamir, M., Kisi, O., Mert, C. & El-Shafie, A. 2021 Design of a hybrid ANN multi-objective whale algorithm for suspended sediment load prediction. *Environmental Science and Pollution Research* **28**, 1596–1611.

Fan, J., Liu, X. & Li, W. 2023 Daily suspended sediment concentration forecast in the upper reach of Yellow River using a comprehensive integrated deep learning model. *Journal of Hydrology* **623**, 129732.

Fang, H.-W. & Wang, G.-Q. 2000 Three-dimensional mathematical model of suspended-sediment transport. *Journal of Hydraulic Engineering* **126** (8), Article 8. https://doi.org/10.1061/(ASCE)0733-9429(2000)126:8(578).

Géron, A. 2022 *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow*. O'Reilly Media, Inc, Sebastopol, CA.

Guillén, J., Palanques, A., Puig, P., Durrieu de Madron, X. & Nyffeler, F. 2000 *Field Calibration of Optical Sensors for Measuring Suspended Sediment Concentration in the Western Mediterranean*. https://doi.org/10.3989/scimar.2000.64n4427

Hasda, R., Rahaman, M. F., Jahan, C. S., Molla, K. I. & Mazumder, Q. H. 2020 Climatic data analysis for groundwater level simulation in drought prone Barind Tract, Bangladesh: Modelling approach using artificial neural network. *Groundwater for Sustainable Development* **10**, 100361.

He, K., Zhang, X., Ren, S. & Sun, J. 2015 Deep residual learning for image recognition. arXiv:1512.03385.

Huang, C.-C., Lai, Y. G., Lai, J.-S. & Tan, Y.-C. 2019 Field and numerical modeling study of turbidity current in Shimen Reservoir during typhoon events. *Journal of Hydraulic Engineering* **145** (5), 05019003.

Huang, C.-C., Chang, M.-J., Lin, G.-F., Wu, M.-C. & Wang, P.-H. 2021 Real-time forecasting of suspended sediment concentrations in reservoirs by the optimal integration of multiple machine learning techniques. *Journal of Hydrology: Regional Studies* **34**, 100804.

Hwang, K. 2017 *Cloud Computing for Machine Learning and Cognitive Applications*. Mit Press, Cambridge, MA.

Hwang, K. & Chen, M. 2017 *Big-data Analytics for Cloud, IoT and Cognitive Computing*. John Wiley & Sons, Hoboken, NJ.

Ighalo, J. O., Adeniyi, A. G. & Marques, G. 2021 Artificial intelligence for surface water quality monitoring and assessment: A systematic literature analysis. *Modeling Earth Systems and Environment* **7** (2), 669–681.

Iriany, A., Rosyida, D., Sulistyono, A. D., Wong, W.-K. & Sui, J. M. 2022 Cross covariance normalized weight of GSTAR-SUR model as input of neural network model on precipitation forecasting. *Journal of Management Information & Decision Sciences* **25** (1), 1–13.

Jhong, B.-C., Fang, H.-T. & Huang, C.-C. 2021 Assessment of effective monitoring sites in a reservoir watershed by support vector machine coupled with multi-objective genetic algorithm for sediment flux prediction during typhoons. *Water Resources Management* **35**, 2387–2408.

Khan, T. A., Alam, M., Ahmed, S. F., Shahid, Z. & Mazliham, M. S. 2019 A factual flash flood evaluation using SVM and K-NN. *2019 IEEE 6th International Conference on Engineering Technologies and Applied Sciences (ICETAS)*, Kuala Lumpur, 20-21 December 2019. IEEE, New York, pp. 1–6.

Landers, M. N., Straub, T. D., Wood, M. S. & Domanski, M. M. 2016 Sediment acoustic index method for computing continuous suspended-sediment concentrations. In *Sediment Acoustic Index Method for Computing Continuous Suspended-Sediment Concentrations (USGS Numbered Series 3-C5; Techniques and Methods, Vols. 3-C5, Issues 3-C5, P. 83)*. U.S. Geological Survey. https://doi.org/10.3133/tm3C5.

Latif, S. D., Chong, K. L., Ahmed, A. N., Huang, Y. F., Sherif, M. & El-Shafie, A. 2023 Sediment load prediction in Johor river: Deep learning versus machine learning models. *Applied Water Science* **13** (3), 79.

Lin, J., He, Q., Guo, L., van Prooijen, B. C. & Wang, Z. B. 2020 An integrated optic and acoustic (IOA) approach for measuring suspended sediment concentration in highly turbid environments. *Marine Geology* **421**, 106062. https://doi.org/10.1016/j.margeo.2019.106062.

Lohrmann, A. 2001 Monitoring sediment concentration with acoustic backscattering instruments. *Nortek Technical Note* **3**, 1–5.

Merten, G. H., Capel, P. D. & Minella, J. P. G. 2014 Effects of suspended sediment concentration and grain size on three optical turbidity sensors. *Journal of Soils and Sediments* **14** (7), Article 7. https://doi.org/10.1007/s11368-013-0813-0.

Mishra, P. N., Bore, T., Jiang, Y., Scheuermann, A. & Li, L. 2018 Dielectric spectroscopy measurements on kaolin suspensions for sediment concentration monitoring. *Measurement* **121**, 160–169. https://doi.org/10.1016/j.measurement.2018.02.034.

Miyata, S., Mizugaki, S., Naito, S. & Fujita, M. 2020 Application of time domain reflectometry to high suspended sediment concentration measurements: Laboratory validation and preliminary field observations in a steep mountain stream. *Journal of Hydrology* **585**, 124747. https://doi.org/10.1016/j.jhydrol.2020.124747.

Nielsen, P. 1984 Field measurements of time-averaged suspended sediment concentrations under waves. *Coastal Engineering* **8** (1), Article 1. https://doi.org/10.1016/0378-3839(84)90022-X.

Orseau, S., Huybrechts, N., Tassi, P., Pham Van Bang, D. & Klein, F. 2021 Two-dimensional modeling of fine sediment transport with mixed sediment and consolidation: Application to the Gironde Estuary, France. *International Journal of Sediment Research* **36** (6), Article 6. https://doi.org/10.1016/j.ijsrc.2019.12.005.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R. & Dubourg, V. 2011 Scikit-learn: Machine learning in python. *The Journal of Machine Learning Research* **12**, 2825–2830.

Simonyan, K. & Zisserman, A. 2015 Very deep convolutional networks for large-scale image recognition. arXiv:1409.1556

Sirabahenda, Z., St-Hilaire, A., Courtenay, S. C. & van den Heuvel, M. R. 2019 Comparison of acoustic to optical backscatter continuous measurements of suspended sediment concentrations and their characterization in an agriculturally impacted river. *Water* **11** (5), Article 5. https://doi.org/10.3390/w11050981.

Su, M., Yao, P., Wang, Z., Zhang, C., Chen, Y. & Stive, M. J. F. 2016 Conversion of electro-optical signals to sediment concentration in a silt–sand suspension environment. *Coastal Engineering* **114**, 284–294. https://doi.org/10.1016/j.coastaleng.2016.04.014.

Thorne, P. D., Vincent, C. E., Hardcastle, P. J., Rehman, S. & Pearson, N. 1991 Measuring suspended sediment concentrations using acoustic backscatter devices. *Marine Geology* **98** (1), Article 1. https://doi.org/10.1016/0025-3227(91)90031-X.

van Maren, D. S., van Kessel, T., Cronin, K. & Sittoni, L. 2015 The impact of channel deepening and dredging on estuarine sediment concentration. *Continental Shelf Research* **95**, 1–14. https://doi.org/10.1016/j.csr.2014.12.010.

Wren, D. G., Barkdoll, B. D., Kuhnle, R. A. & Derrow, R. W. 2000 Field techniques for suspended-sediment measurement. *Journal of Hydraulic Engineering* **126** (2), Article 2. https://doi.org/10.1061/(ASCE)0733-9429(2000)126:2(97).

First received 10 September 2023; accepted in revised form 13 October 2023. Available online 26 October 2023