HYDROGEN RELEASE MODELLING FOR ANALYSIS USING DATA-DRIVEN AUTOENCODER WITH CONVOLUTIONAL NEURAL NETWORKS

Yunqing Xu [1], Shangzhi Liu [2], Zhongxu He [3], Jing Dang [4], Yi Liu [5,*]

[1] College of Mechanical and Electrical Engineering, China University of Petroleum (East China), Qingdao, 266580, China, z21040108@s.upc.edu.cn

[2] College of Mechanical and Electrical Engineering, China University of Petroleum (East China), Qingdao, 266580, China, s18040080@s.upc.edu.cn

[3] College of Mechanical and Electrical Engineering, China University of Petroleum (East China), Qingdao, 266580, China, s20040069@s.upc.edu.cn

[4] College of Chemistry and Chemical Engineering, China University of Petroleum (East China), Qingdao, 266580, China, z21030111@s.upc.edu.cn

[5] College of Chemistry and Chemical Engineering, China University of Petroleum (East China), Qingdao, 266580, China, liuyi@upc.edu.cn

ABSTRACT

High-accuracy gas dispersion models are necessary for predicting hydrogen movement, and for reducing the damage caused by hydrogen release accidents in chemical processes. In urban areas, where obstacles are large and abundant, computational fluid dynamics (CFD) would be the best choice for simulating and analyzing scenarios of the accidental release of hydrogen. However, owing to the large computation time required for CFD simulation, it is inappropriate in emergencies and real-time alarm systems. In this study, a non-linear surrogate model based on deep learning is proposed. Deep convolutional layer data-driven autoencoder and batch normalized deep neural network is used to analyze the effects of wind speed, wind direction and release degree on hydrogen concentration in real-time. The typical parameters of hydrogen diffusion accidents at hydrogen refuelling stations were acquired by CFD numerical simulation approach, and a database of hydrogen diffusion accident parameters is established. By establishing an appropriate neural network structure and associated activation function, a deep learning framework is created, and then a deep learning model is constructed. The accuracy and timeliness of the model are assessed by comparing the results of the CFD simulation with those of the deep learning model. To develop a dynamic reconfiguration prediction model for the hydrogen refuelling station diffusion scenario, the algorithm is continuously enhanced and the model is improved. After training is finished, the model's prediction time is measured in seconds, which is 10^5 times quicker than field CFD simulations. The deep learning model of hydrogen release in hydrogen refuelling stations is established to realize timely and accurate prediction and simulation of accident consequences and provide decision-making suggestions for emergency rescue and personnel evacuation, which is of great significance for the protection of human life, health and property safety.

Keywords: Hydrogen release; Computational fluid dynamics; Autoencoder; Convolutional neural network; Deep learning

1.0 INTRODUCTION

As the global climate and energy concerns continue to escalate, countries worldwide are becoming more focused on researching and utilizing clean energy. In the "14th Five-Year Plan" of China, hydrogen energy has been included in Chapter 9, Section 2 "Forward-looking Planning for Future Industries," and has been jointly planned as part of "organizing and implementing future industry incubation and acceleration plans, planning and laying out several future industries." Guided by policy, China's hydrogen energy industry is rapidly developing, with continuous improvement of hydrogen energy infrastructure, which has triggered a "hydrogen energy boom."

With the ongoing promotion of "hydrogen energy," the development of hydrogen refuelling stations has been rapidly advancing. Unfortunately, accidents can occur, as was the case on May 23, 2019, in Jiangling City, Gangwon Province, South Korea, where a hydrogen storage tank leaked gas, resulting in an explosion that caused two fatalities and six injuries. On June 10, 2019, a fire and explosion occurred at an operational hydrogen refuelling station in Oslo, the capital of Norway. This incident underscores the importance of recognizing the potential dangers associated with the storage and use of hydrogen gas. Despite its many benefits, such as producing zero emissions, hydrogen gas can be highly flammable and explosive under certain conditions, making safety precautions essential to avoid accidents and injuries. Additionally, being prepared for an emergency response and personnel evacuation in case of a hydrogen leak is equally important to minimize the consequences of such occurrences. Thus, developing rapid methods for predicting the outcomes of hydrogen leaks and quickly reconstructing the entire hydrogen flow field is essential. Urgent action is required to enhance safety measures and emergency response protocols for the use of hydrogen in refuelling stations.

The commonly used CFD model can provide accurate representations of the entire flow field distribution following a hydrogen leak. However, CFD modelling can be time-consuming, making it challenging to generate simulation results promptly following a leak. As a result, using CFD methods to restore flow field conditions at the time of an accident during the accident investigation stage may not be feasible.

Recognizing the limitations of CFD methods in reconstructing flow fields, many researchers have turned their attention to artificial intelligence (AI) technology as a promising alternative for achieving rapid and accurate flow field prediction and restoration. In contrast to CFD, AI-based techniques can produce predictions in a much shorter period, making them suitable for scenarios that require the rapid reconstruction of flow fields. Deep learning technology can extract complex data features from large amounts of data, learn from experience or knowledge of the data, and make predictions for relevant complex data 1-4. This method uses the powerful multi-dimensional and complex non-linear data representation capabilities of deep learning technology to directly mine and fit data from the flow field, achieving the prediction of flow field features 5. Many scholars both domestic and international have conducted research on the reconstruction of the consequences of substance leakage and diffusion based on deep learning technology and successfully predicted the concentration distribution after the leakage through deep learning modelling. Na et al. 6 studied the diffusion of toxic gases in urban areas and proposed a non-linear surrogate model based on deep learning to analyze the death rate of gas diffusion in real-time. The model takes several orders of magnitude less time than the CFD model. Shi et al. 7 proposed a mixed probability convolutional-variational autoencoder-variational Bayesian neural network (Conv-VAE-VBNN) based on deep learning technology for real-time prediction of the consequences of leaks and diffusion, to replace CFD models. Zou Jianguo 8 proposed a prediction model called RCL-Learning based on residual network and convolutional LSTM integration, which was used to predict the diffusion trend of regional air pollutants. Ni Jing 9 used field experimental data to construct toxic heavy gas diffusion concentration prediction models based on deep belief networks and convolutional neural networks respectively. The results of the study showed that these prediction models demonstrated significant advantages.

Currently, research on using deep learning methods to construct CFD proxy models to achieve rapid flow field reconstruction and prediction of leakage consequences mainly focuses on toxic gases, methane, and other substances. However, there has been relatively little research on the rapid prediction of consequences resulting from hydrogen leaks. This study aims to create a rapid prediction algorithm for high-pressure hydrogen leakage outcomes in hydrogenation stations using deep learning methods. This algorithm will serve as a basis for guiding emergency response and personnel evacuation decisions.

2.0 HYDROGEN DISPERSION CFD MODEL

According to the relevant regulations, the fire distance between the high-pressure hydrogen storage tank and the firewall is set at 6 meters, and the fire distance between the hydrogen processor and the firewall is set at 10 meters. The minimum fire distance between the high-pressure hydrogen storage tank and the hydrogen station fence is 8 meters, while the minimum fire distance between the hydrogen processor and the fence is 5.3 meters.

In this study, the main parameters of the environmental conditions for CFD simulation are ambient temperature, ambient pressure, ambient wind speed, and wind direction. To facilitate the CFD simulation process, the researchers standardized the local ambient temperature to 300K and set the ambient pressure to the standard atmospheric pressure.

The leak source was chosen to be the high-pressure hydrogen storage tank. The leak point is located at the interface between the top of the high-pressure hydrogen storage tank and the flange, so the modelling of the high-pressure hydrogen storage tank needs to be detailed. According to the relevant requirements of the enterprise standard LNQ019-2021 "Hydrogen Storage Bottle Container Group for Hydrogen Refueling Stations," the high-pressure hydrogen storage tank with a storage working pressure of 45MPa has a model number of a single bottle with an outer diameter of 485mm and a volume of 1000L. Based on the standard model, the high-pressure hydrogen storage tank was modelled, and the entire physical model of the hydrogen refuelling station is shown in Fig. 1.

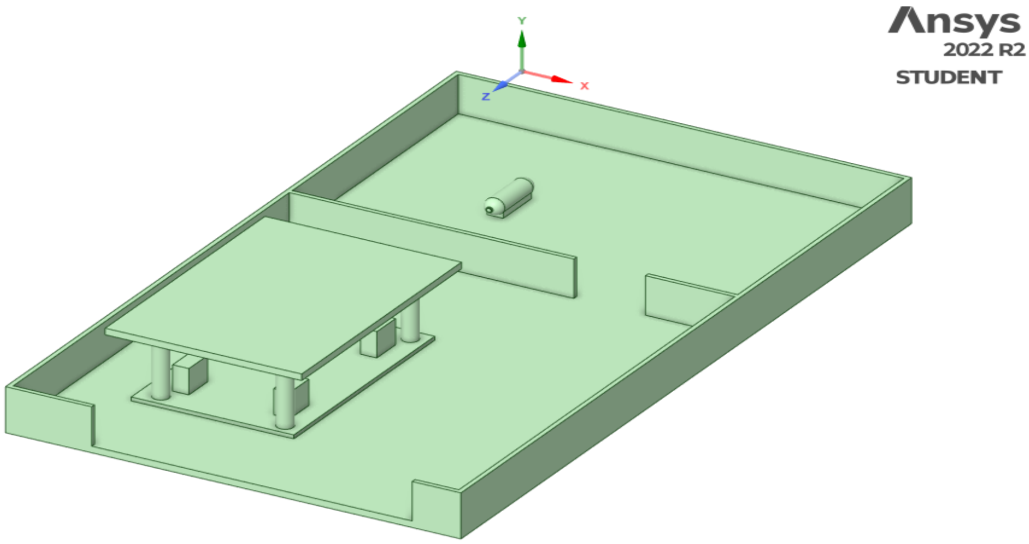

Figure 1. Physical model of hydrogen refuelling station

In this study, 20 different operating conditions of high-pressure hydrogen gas leakage in hydrogen stations were simulated using computational fluid dynamics (CFD). The simulations considered environmental conditions with and without wind, including 5 wind speeds and 2 wind directions. The leakage process was examined for 2 leak sizes and 2 leakage rates. The specific physical conditions for CFD simulations of high-pressure hydrogen gas leakage in hydrogen stations are shown in Table 1, where "+Z" indicates south wind and "-Z" indicates north wind.

Table 1. CFD simulation of physical working conditions

Number	Wind speed (m/s)	Wind direction	Leakage aperture (mm)	Virtual nozzle diameter (mm)	leakage rate (kg/s)	Pressure (MPa)	Temperature (K)
1	0	+Z	10	14	1.85	45	300
2	0	-Z	10	14	1.85	45	300
3	0	+Z	20	28	7.40	45	300
4	0	-Z	20	28	7.40	45	300
5	2	+Z	10	14	1.85	45	300
6	2	-Z	10	14	1.85s	45	300
7	2	+Z	20	28	7.40	45	300
8	2	-Z	20	28	7.40	45	300
9	2.4	+Z	10	14	1.85	45	300
10	2.4	-Z	10	14	1.85	45	300

11	2.4	+Z	20	28	7.40	45	300
12	2.4	-Z	20	28	7.40	45	300
13	3.2	+Z	10	14	1.85	45	300
14	3.2	-Z	10	14	1.85	45	300
15	3.2	+Z	20	28	7.40	45	300
16	3.2	-Z	20	28	7.40	45	300
17	3.6	+Z	10	14	1.85	45	300
18	3.6	-Z	10	14	1.85	45	300
19	3.6	+Z	20	28	7.40	45	300
20	3.6	-Z	20	28	7.40	45	300

ANSYS Fluent 2022R2 was selected to perform CFD simulation, and the simulation results were directly outputted during post-processing after the solution was completed. After the simulation was completed. Following the simulation, the results of the high-pressure hydrogen gas leak simulation at the refuelling station were output in the form of cloud diagrams. To enhance the visualization of the simulation results, a contour plot was created on the plane formed by the centerline position of the hydrogen jet in the ZX plane. This allowed for a clearer representation of the distribution of the high-pressure hydrogen gas leak simulation results.

Given that the explosion limit of hydrogen ranges from 4% to 75%, hydrogen within this concentration range poses a high risk of explosion. This study primarily focuses on the concentration distribution of high-pressure hydrogen leakage within this dangerous range. Fig. 2 shows the concentration distribution range of high-pressure hydrogen leakage in six working conditions within 4%~75% in the cloud map.

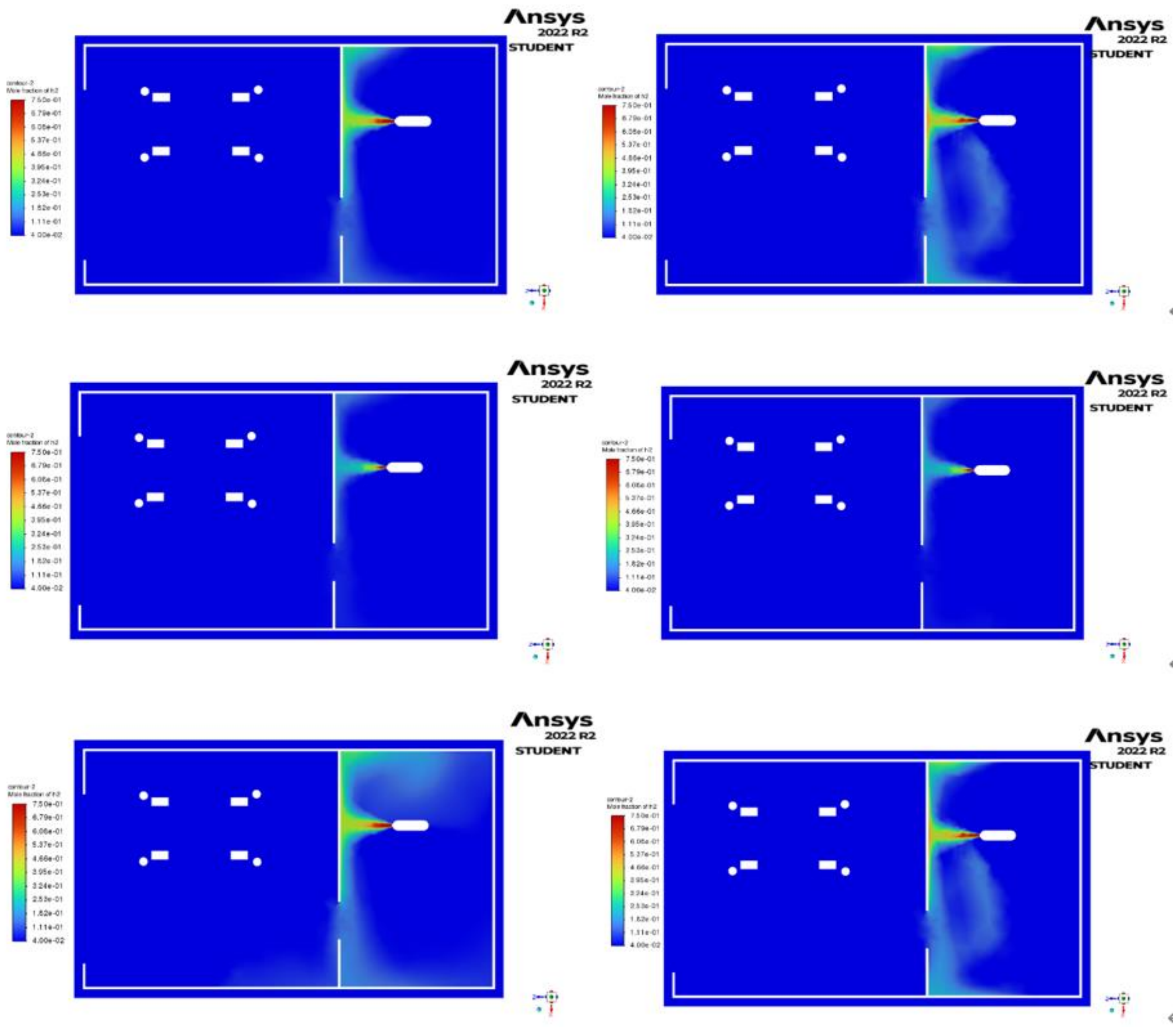

Figure 2. CFD leakage simulation results from high-pressure hydrogen leakage from a hydrogen refuelling station

Based on the nature of simulating the leakage of high-pressure hydrogen gas, it is classified as a transient simulation in the CFD process. To ensure accurate results, a residual convergence criterion of 0.001 is set for each factor. The residuals in the CFD simulation process are shown in Fig. 3.

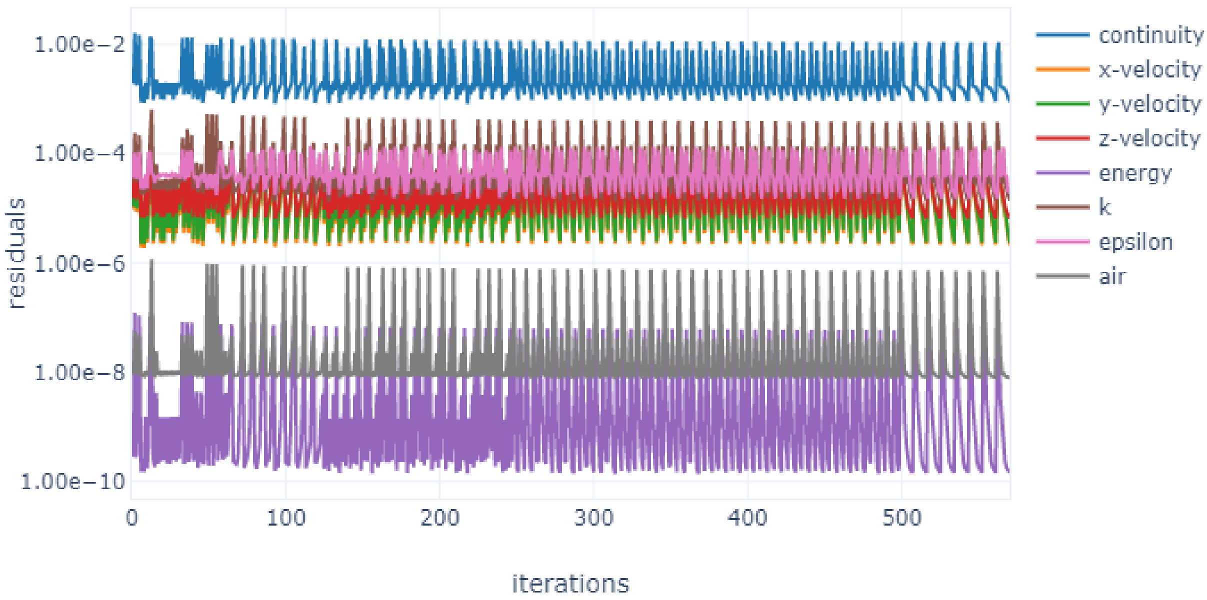

Figure 3. Convergence results of the residuals of each factor in the CFD iterative calculation process

Fig. 3 illustrates that as the number of iterations increases, the residuals for each factor in the CFD simulation process gradually decrease, eventually converging below 0.001. This indicates that the CFD model has accurately calculated the simulation results. The concentration distribution results obtained from the simulation of high-pressure hydrogen gas leakage under 20 different operating conditions in Fluent should be saved and exported. The accumulated data will be used to construct a deep learning model to predict the consequences of hydrogen gas leakage at hydrogen refuelling stations in the future.

3.0 APPLICATION OF CAE-DNN

3.1 Data preparation

To create a proxy model using neural networks, hundreds of sample data points are required for training. However, manually acquiring a large amount of data is expensive. To generate the required data, an automated process is used by linking Python code with Fluent. The operating conditions are used as the input data for the entire model. The input operating parameters and the resulting hydrogen concentration distribution under those parameters are used as the inputs and outputs, respectively, for the deep learning model that was constructed. CFD simulation-generated data on the hydrogen concentration distribution is used as input data for the data reconstruction component. Training the proxy model directly with Fluent's results can be challenging. Therefore, it is necessary to preprocess the data.

The primary variables in the working conditions are ambient wind speed, wind direction, and hydrogen leak rate. The leak rate is associated with the size of the leak orifice. The development of a deep learning surrogate model for high-pressure hydrogen leaks at hydrogen refuelling stations does not incorporate the geometric shape of the station. Instead, the inputs to the deep neural network model consisting of the wind speed (v), wind direction (d), and leak rate (q). Following the application of normalization algorithms, the input data for the regression neural network model is confined to the range of [0, 1] or [-1, 1]. The normalization of the original data features may result in a faster gradient descent speed, leading to faster convergence during the gradient descent optimization process.

Due to the difficulty in collecting data, the amount of data available may be limited, and there may be an uneven distribution of data. To address this issue, we apply data augmentation methods in this paper to increase the amount of training data and enhance its diversity. This approach effectively mitigates the overfitting phenomenon that can occur in models trained on sparse data features and enables the model to generalize better and improve its robustness.

The outcome of CFD simulations aimed at predicting the effects of hydrogen leaks at refuelling stations is primarily impacted by three key parameters: wind speed (v), wind direction (d), and leak rate (q). Predicting the concentration distribution of leaked hydrogen gas is possible by inputting these three operational parameters. The input dataset for the data regression component of the model contains the wind speed (v), wind direction (d), and leak rate (q) parameters under 20 distinct operating conditions. This part of the input is represented as $I \in R^{1\times3}$. The entire input part data set can be expressed as：

$$I_n=\begin{bmatrix} I_{11} & \cdots & I_{13} \\ \vdots & \ddots & \vdots \\ I_{n1} & \cdots & I_{n3} \end{bmatrix},\ n=20$$

Feature mapping of CFD data is predominantly achieved through the parts of data dimensionality reduction and data reconstruction. The input for this part is the CFD output data corresponding to 20 operating conditions after data augmentation. The input for the feature mapping section can be represented as $x \in R^{400\times300\times3}$. The entire input dataset can be represented as：

$$x_n=\begin{bmatrix} x_{11} & \cdots & x_{1j} \\ \vdots & \ddots & \vdots \\ x_{i1} & \cdots & x_{ij} \end{bmatrix}$$

Once the input dataset for the entire algorithm model is established, the training and validation sets are constructed. The data set for dimensionality reduction and data reconstruction is divided into an 8:2 ratio. For data regression, the dataset has been split into a 7:3 ratio, with 70% of the data used for training and the remaining 30% for validation.

3.2 Training and simulation

In traditional autoencoders, the encoder reduces high-dimensional data, such as images, into a low-dimensional feature vector, while the decoder upsamples the compressed data to reconstruct the original image. However, using fully connected linear layers in both the encoder and decoder can be ineffective in representing high-dimensional data, leading to overfitting or underfitting during feature extraction. Therefore, a deep convolutional neural network is used as the encoder instead of linear layers. When it comes to image reconstruction tasks, various methods can be used to upscale image data. Among these methods, transposed convolution is the most effective. Consequently, a transposed convolutional neural network is employed for the upsampling process in the decoder when building a convolutional autoencoder (CAE) model. In addition, mini-batch implementation techniques can be compared for their respective characteristics in training the CAE model. Fig. 4 illustrates the structure of the CAE network model.

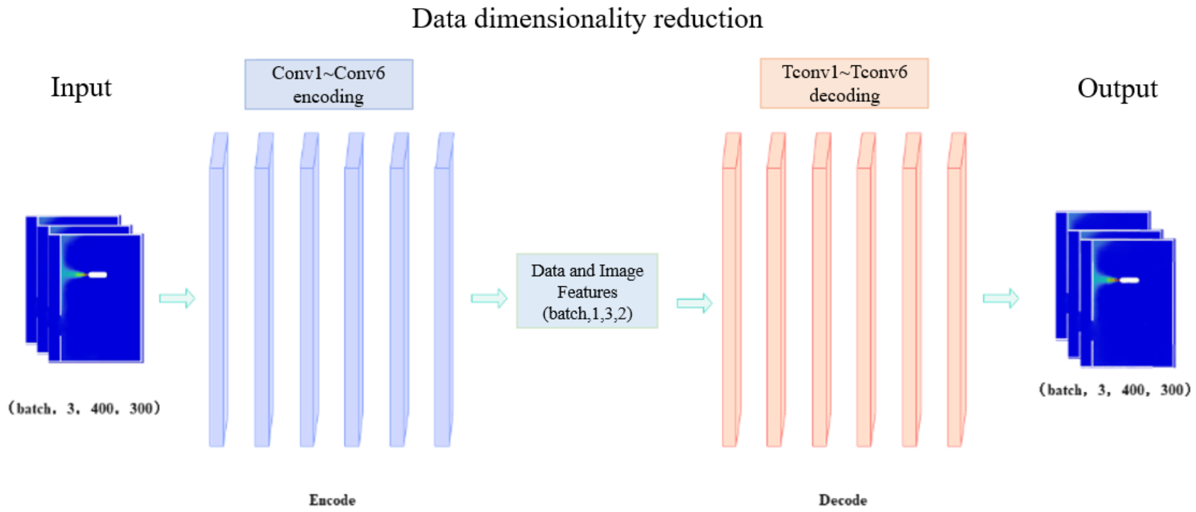

Figure 4. CAE model structure

Since the model is trained using mini-batch stochastic gradient descent, the input data for the CAE model is a three-channel hydrogen leakage concentration distribution map. Batch normalization is applied to each layer of the neural network in the encoder and decoder of the CAE model. The activation function used in the first five layers is ReLU, while the last convolutional layer uses the Sigmoid function for output. In the CAE model, the output single-channel (3x2) image feature data is directly fed into 6 layers of transposed convolutional neural network for upsampling reconstruction, and the result is a three-channel high-dimensional output tensor. This ensures that the input and output have the same size during reconstruction. To perform data regression prediction, a deep neural network consisting of 7 linear layers is constructed. The detailed network structure is illustrated in Fig. 5.

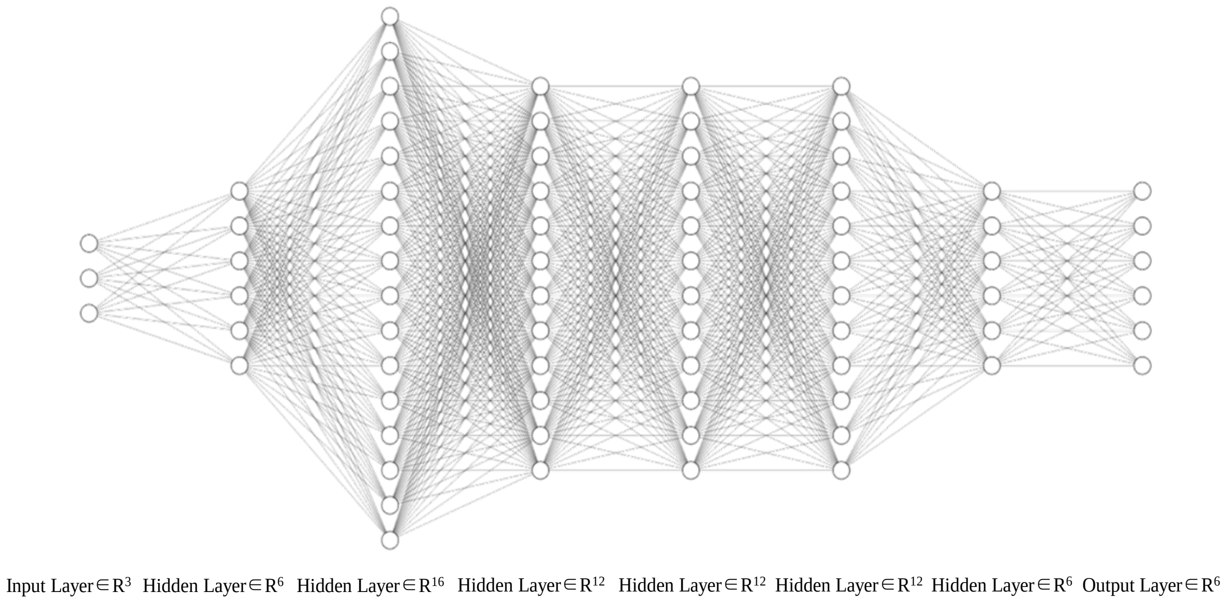

Figure 5. DNN model structure

The DNN model takes as input the dataset I_n, which contains physical operating conditions for hydrogen refuelling stations. Each data record in the dataset is a one-dimensional tensor with a size of 3. The model is trained using a supervised learning approach. The DNN model consists of 7 fully connected layers. To prevent overfitting during the fitting process, Dropout is applied after the fully connected layers in the 3rd, 4th, and 5th layers of the DNN model, and the neuron dropout rate is set to 0.1. After the data features are fitted by the DNN model and the data shape is modified, they can be passed into a data reconstruction network to generate the hydrogen concentration distribution of hydrogen refuelling station leakage consequences as the output.

Upon completion of the training phase, both the CAE model and data regression DNN model were saved. These models were then combined to create the hydrogen refuelling station high-pressure hydrogen leakage consequence prediction CAE-DNN model. The structure of the CAE-DNN model is illustrated in Fig. 6, depicting the encoder, decoder, and data regression components.

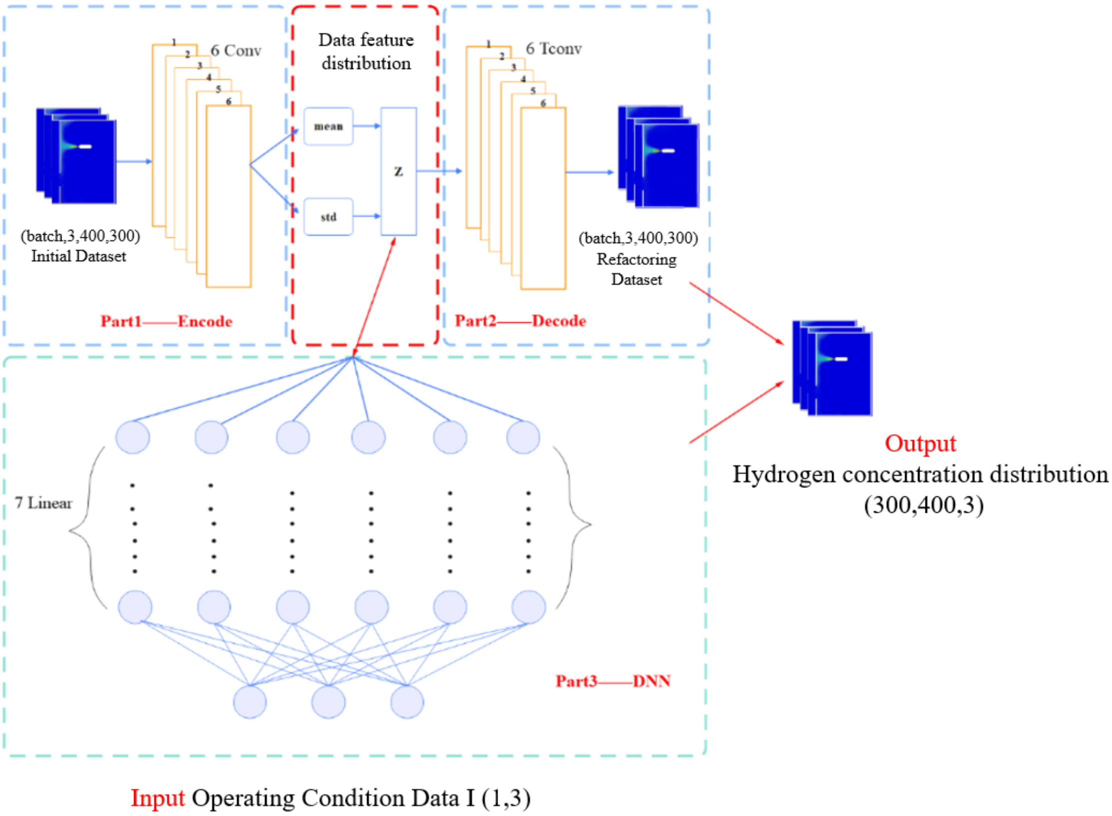

Figure 6. CAE-DNN model

The CAE-DNN model takes in three physical data inputs, namely wind speed, wind direction, and leak rate, to quickly predict the consequences of high-pressure hydrogen gas leaks in hydrogen refuelling stations. The model's input is a parameter of shape (1, 3), and the output is a three-channel image data of shape (1, 3, 400, 300). After restoring the data shape using Torchvision, the final output is a hydrogen distribution cloud map of shape (300, 400, 3).

4.0 RESULTS AND DISCUSSION

4.1 Quality analysis of CAE model reconstruction

To validate the data reconstruction performance of the CAE model, a saved CAE model was loaded and the hydrogen concentration distribution data from 20 different operating conditions of CFD simulations without data augmentation were used as the development set. As an example, the hydrogen concentration distribution results of the CAE model were generated for a leakage rate of 7.40 kg/s and a south wind direction, and are shown in Fig. 7.

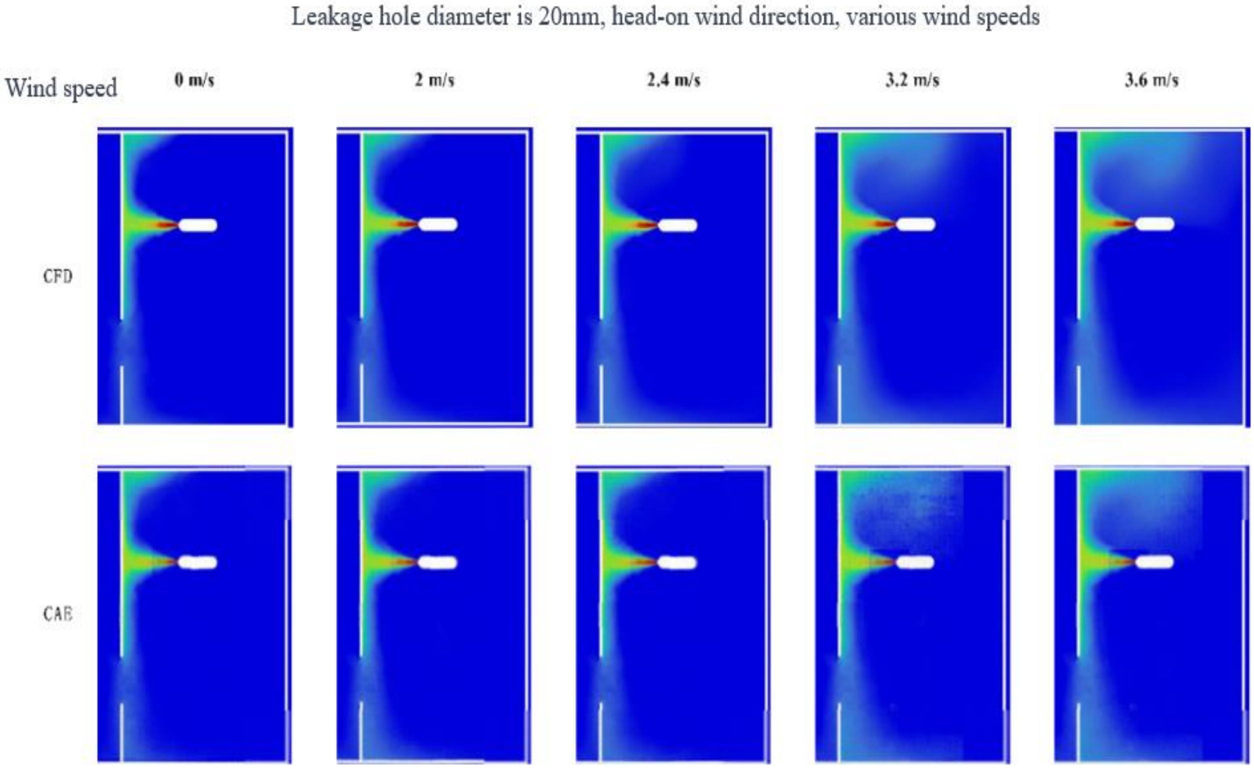

Figure 7. CAE model leakage concentration distribution data reconstruction results

A comparison was made between the reconstructed hydrogen concentration distribution results of the CAE model and the CFD simulation results for a leakage rate of 7.40 kg/s, a head-on wind direction, and various wind speeds. It was observed that the CAE model accurately reconstructed the hydrogen concentration field after the leakage at low wind speeds (0m/s, 2m/s, 2.4m/s). Nevertheless, at wind speeds of 3.2m/s and 3.6m/s, the reconstructed hydrogen concentration distribution images by the CAE model exhibited a significant rise in noise, a sparsely reconstructed distribution at low hydrogen concentrations, and concentration distribution fluctuations at high hydrogen concentrations.

The Peak Signal-to-Noise Ratio (PSNR) was used to quantitatively analyze the quality of the hydrogen concentration distribution images reconstructed by the CAE model. PSNR is a widely used engineering term that measures the quality of a signal's representation by comparing the maximum possible power of the signal to the noise power that can affect its accuracy. As many signals have a very wide dynamic range, PSNR is commonly expressed in logarithmic decibel units. The PSNR was calculated using the following method:

$$\mathrm{PSNR}=10\cdot\log_{10}\left(\frac{\mathrm{MAX}_{\mathrm{I}}^{2}}{\mathrm{MSE}}\right)=20\cdot\log_{10}\left(\frac{\mathrm{MAX}_{\mathrm{I}}}{\sqrt{\mathrm{MSE}}}\right) \tag{1}$$

where MSE is the Mean Squared Error of image data and MAX_I is the maximum value of colour intensity for a pixel in an image.

A larger PSNR value corresponds to a smaller Mean Squared Error (MSE), indicating a higher image quality. A larger PSNR value corresponds to a smaller Mean Squared Error (MSE), indicating a higher image quality. PSNR evaluation image quality standards are shown in Table

2.

The PSNR value of the reconstructed hydrogen concentration data after high-pressure hydrogen leakage at the refuelling station output by the CAE model was calculated, and the quality of the reconstructed images was analysed. Table 3 shows the PSNR calculation results in 20 different operating conditions in the development set.

Table 2. PSNR evaluation criteria

PSNR Range	Image quality evaluation results	Level
PSNR ≥40dB	Very good image quality (i.e. very close to the original image)	A
30dB≤PSNR<40dB	Good image quality (i.e. distortion is perceptible but acceptable)	B
20dB≤PSNR<30dB	Poor image quality	C
PSNR<20dB	Unacceptable image quality	D

Table 3. PSNR calculation results

Number	PSNR	Level	Number	PSNR	Level
1	29.3945	B	11	28.2572	B
2	31.4305	A	12	29.7720	B
3	28.9533	B	13	29.3500	B
4	30.4884	A	14	30.0688	A
5	28.8098	B	15	30.4391	A
6	30.3140	A	16	27.7089	B
7	28.3754	B	17	31.1062	A
8	30.7570	A	18	28.6609	B
9	28.1060	B	19	29.4399	B
10	28.8508	B	20	28.5324	B

Table 3 shows that the CAE model was effective in reconstructing the hydrogen concentration distribution results when the wind speed was low and the wind direction was north (opposite to the leakage direction), and when the data noise quality was good. However, for other operating conditions, the reconstructed data had poor noise quality. The average PSNR value across all operating conditions was 29.4408.

4.2 Quality analysis of CAE-DNN model reconstruction

CAE-DNN model takes as input three physical operating values of the leak scenario in the refuelling station: wind speed, wind direction, and leak rate, i.e., input data $I_n \in R^{1\times3}$。 After undergoing DNN calculations in the model, the input data outputs hidden layer data $H_n \in R^{1\times6}$。 Based on the twenty operating condition data used in the test experiment, the hidden layer feature data is reshaped and input into the CAE generator for result generation. Finally, the CAE-DNN model outputs a hydrogen concentration distribution image in RGB format. For instance, the results obtained by the CAE-DNN model for the condition with a leak rate of 7.40 kg/s are compared to those generated by CFD, as shown in Fig. 8.

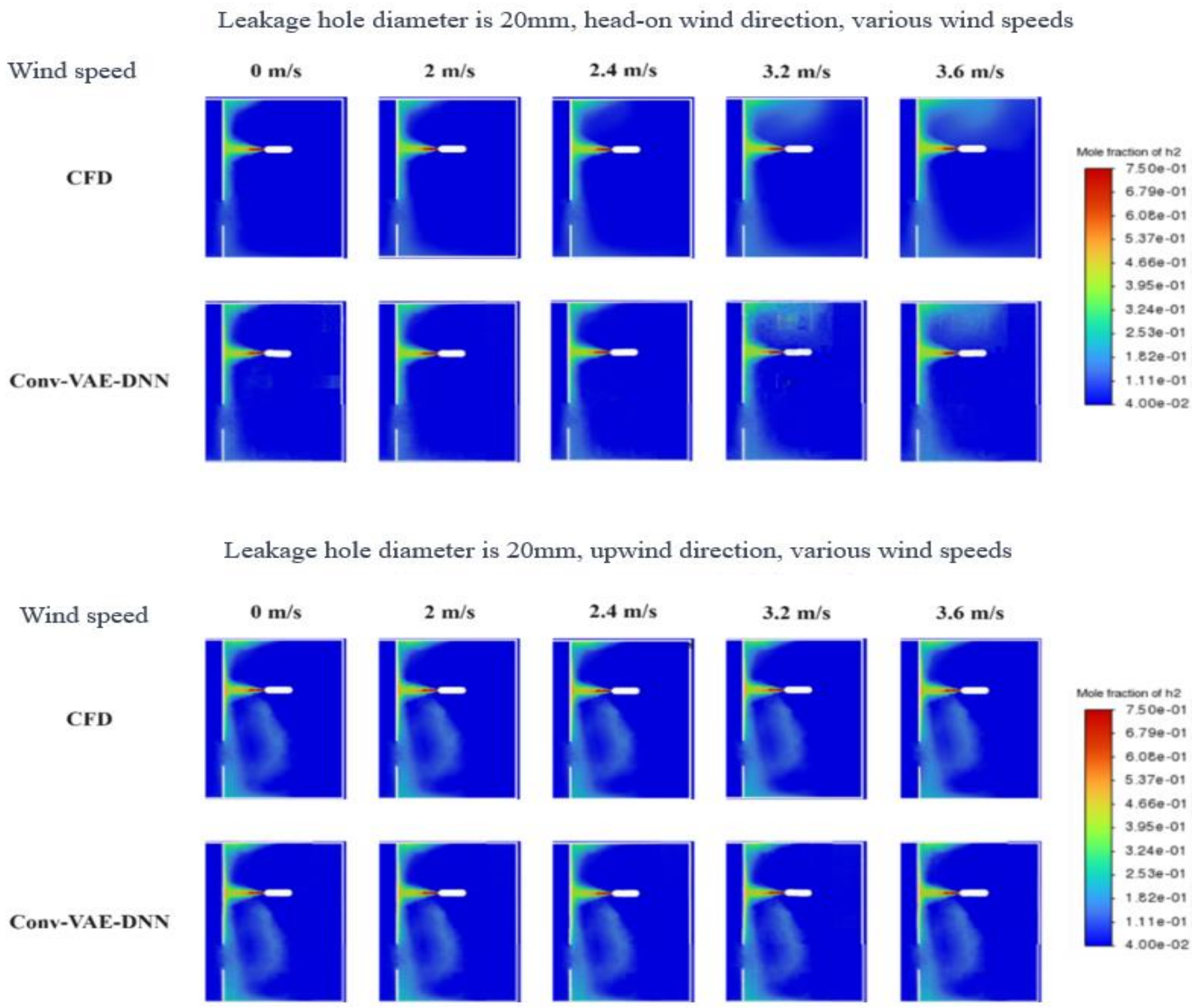

Figure 8. CAE-DNN hydrogen refuelling station high-pressure hydrogen leak consequence prediction

The accuracy of the results generated by the CAE-DNN deep learning surrogate models and the CFD method is compared using the Structure Similarity Index Measure (SSIM) concept. Additionally, an analysis of the accuracy of the output results produced by the deep learning model is performed. The calculation method for SSIM can be found in Equation (2):

$$\mathrm{SSIM}(x, \hat{x})=\frac{(2\mu_x\mu_{\hat{x}}+c_1)(2\sigma_{x\hat{x}}+c_2)}{(\mu_x^2+\mu_{\hat{x}}^2+c_1)(\sigma_x^2+\sigma_{\hat{x}}^2+c_1)} \tag{2}$$

where μ is the mean value of the data and σ is the standard deviation of the data, and c is a constant.

Table 4 shows the SSIM and computation time results for the CAE-DNN model.

Table 4. Evaluation of CAE-DNN model output results

Number	SSIM	TIME (s)	Number	SSIM	TIME (s)
1	0.8538	3.400786	11	0.8555	3.400126
2	0.8342	3.400197	12	0.8583	3.300252

3	0.8465	3.300108	13	0.8509	3.400197
4	0.8751	3.100194	14	0.8649	3.300157
5	0.8476	3.400220	15	0.8278	3.200092
6	0.8375	3.300156	16	0.8425	3.200235
7	0.8705	3.300108	17	0.8251	3.200188
8	0.8749	3.300251	18	0.8407	3.200092
9	0.8570	3.400173	19	0.8028	3.400269
10	0.8358	3.300204	20	0.8445	3.200212

According to Table 4, the average similarity of the CAE-DNN model is 0.8473, and the average prediction time is 3.3 seconds. The analysis results indicate that under the same hardware conditions when the accuracy of the CAE-DNN model is 84.73% compared to the CFD model, the processing time is only 3.3 seconds. In contrast, the CFD simulation method takes around 4.5 hours to complete under the same hardware and environmental conditions. These findings strongly suggest that the CAE-DNN model can effectively serve as a surrogate model for the CFD model, providing fast and accurate predictions of the outcomes of high-pressure hydrogen gas leaks at hydrogen refuelling stations.

5.0 CONCLUSIONS

This study is based on CFD simulation data and builds a deep-learning model for the rapid prediction of consequences after high-pressure hydrogen leaks in hydrogen refuelling stations. It achieves a fast and accurate prediction of hydrogen concentration distribution after high-pressure hydrogen leaks. By reconstructing the hydrogen concentration distribution flow field, it provides an important decision-making basis for hydrogen safety management in refuelling stations and timely emergency response after hydrogen leaks.

The optimized convolutional variational autoencoder model performed better than the convolutional autoencoder model in generating data quality. By building a deep regression neural network model, the CAE-DNN deep learning model was constructed. This model can rapidly generate concentration distribution results of hydrogen leakage consequences at hydrogen refuelling stations by inputting operating condition data. The main conclusions are as follows:

(1) A physical model for simulating hydrogen leakage in hydrogen refuelling stations was established. The under-expanded jet of high-pressure hydrogen was simulated under 20 different operating conditions, and hydrogen concentration distribution data within the hydrogen explosion limit along the centerline direction of the leakage jet were obtained for different physical conditions.

(2) Based on the steps of data dimension reduction, data regression, and data reconstruction, a CAE-DNN deep learning model was developed to predict the consequences of high-pressure hydrogen leaks in hydrogen refuelling stations. This model can take in operational condition parameters and predict the corresponding distribution of hydrogen leak concentration. Compared with CFD output results, the CAE-DNN model constructed in this study had an average accuracy of 84.73% of the CFD model, with an average processing time of only 3.3s, while the CFD simulation using grid calculation under the same conditions took 4.5 hours. The resulting CAE-DNN algorithm model is nearly 5000 times faster than the CFD method, making it a highly efficient and accurate tool for predicting the consequences of high-pressure hydrogen leaks in hydrogen refuelling stations.

REFERENCES

1. Cheng, R., A summary of the development of artificial intelligence technology, *Chinese security and safety*, **10**, 2020, pp. 81-83.
2. QiaoJin, G., Jie, H., Shijie, G., Overview of the development of deep learning computing platforms, *Informationization research*, **45**, No. 3, 2019, pp. 1-7.
3. Shuai, X., Overview of Deep Learning and Deep Reinforcement Learning, *Information technology and informatization*, No. 5, 2020, pp. 225-227.
4. Xiaofeng, T., Guangshuai, L., Overview of the Development of Deep Learning, 15th National Symposium on Geophysics for National Security, Gansu Province, 2019.
5. Yingxing, W., Han, R., Ziyang, L., Research progress on deep learning-based modelling techniques for fluid dynamics, *Journal of Aeronautics*, **42,** No. 4, 2021, pp. 225-244.
6. Na, J., Jeon, K., Lee, W.B., Toxic gas release modelling for real-time analysis using variational autoencoder with convolutional neural networks, *Chemical Engineering Science*, **181**, 2018, pp. 68-78.
7. Shi, J., Li, J., Usmani, A.S., et al., Probabilistic real-time deep-water natural gas hydrate dispersion modelling by using a novel hybrid deep learning approach, *Energy*, **219**, 2021, pp. 119-572.
8. Guojian, Z., Research on Prediction of Regional Air Pollutant Dispersion Trend Based on Spatio-temporal Feature Learning, Shanghai Normal University, 2020.
9. Jing, N., Real-time prediction of toxic heavy gas dispersion concentration based on deep learning, Soochow University, 2019.