A Novel CNN Modeling Algorithm for the Instantaneous Flow Rate Measurement of Gas-liquid Multiphase Flow

Haifeng Zhang[1], Yiyuan Yang[1], Ming Yang[2], Likun Min[3], Yi Li[1] and Xiangyuan Zheng[1]
[1]Division of Ocean Science & Technology, Tsinghua Shenzhen International Graduate School, China
[2]AI Division, Shenzhen Leengstar Technology Co., LTD, China
[3]Surface engineering department, Jidong Oil Field of Pertrochina, China
+86 18810114818, +86 15927088163, +86 17673166487, +86 18633103380, +86 18028771757, +86 075526036292
zhanghf1986@sz.tsinghua.edu.cn; yangyy19@mails.tsinghua.edu.cn; yangming@leengstar.com; mlk2005@petrochina.com.cn; liyi@sz.tsinghua.edu.cn; zheng.xiangyuan@sz.tsinghua.edu.cn

ABSTRACT
The measurement of the instantaneous flow rate of gas-liquid two-phase flow is a key technology in the industrial production process, and how to build an instantaneous model with long-term cumulative flow labels is also an important technical problem. In order to solve it, we propose a novel CNN (convolutional neural network) modeling algorithm for the instantaneous flow measurement. Firstly, the one-dimensional convolutional neural network is used to build the instantaneous model. Then the long-term flow label slice and average technology are applied for the constraint model. Finally, based on the supervised model, the instantaneous flow model can be trained unsupervised. Test results show that the method can observe instantaneous flow changes and the novel CNN prediction results are generally superior to the other prediction model directly used the average flow samples labels and CNN. The novel CNN modeling algorithm proposes in this paper will have important application value for industrial process measurement.

CCS Concepts
• **Computing methodologies→Machine learning→Learning settings→Semi-supervised learning settings**

Keywords
One-dimensional CNN; Gas-liquid multiphase flow; Instantaneous model; Constraint model

1. INTRODUCTION
Online measurement of gas-liquid two-phase flow is a key technology in many fields such as petroleum, chemistry, pharmaceutical, etc. Mastering the changes in instantaneous gas and liquid during the two-phase mixed-flow process is a problem that the process measurement industry is concerned about. Due to the complicated flow process and difficult mathematical description, it is difficult to accurately measure the flow rate of each phase in the mixed flow process of gas-liquid two-phase flow. Therefore, in recent years, researchers have begun to discuss the measurement models based on machine learning [1]. For example, Fan S used conductivity probe and neural network to perform gas-water two-phase slug flow measurement in a horizontal tube, which implemented the measurement accuracy of gas and liquid flow error less than 10% [2]. Hu D applied a convolutional neural network to predict the flow of gas-liquid multiphase flow in different regions [3]. Zhao C used the microwave time series method to measure the water-liquid ratio of the oil-water-gas three-phase flow and implemented an average absolute error of 5.2% in the moisture content of the sample [4]. However, the flow of gas and liquid is changing continually, so the current measured flow rate is for a period of time such as an average flow of 5 minutes. The error of the instantaneous flow sample's label is considerable, which makes the model training error larger [5]. If we want to get the instantaneous flow label, it requires a higher cost to purchase more advanced equipment or modify the process. However, it is impossible to achieve during the measurement process or experimental engineering production process. At present, the common method still uses the traditional test separation tank to obtain the average/accumulated flow label. However, it cannot achieve the highly accurate measurement of instantaneous measurements.

In order to improve the accuracy of the instantaneous flow measurement. According to the industrial practical application, we improve the structure of the classical convolutional neural network and combines the semi-supervised learning method to build the instantaneous flow measurement model [6,7], which can learn the instantaneous flow change lows in the process under the constraint cumulative flow.

ICMLC 2020, February 15–17, 2020, Shenzhen, China

ACM ISBN 978-1-4503-7642-6/20/02…$15.00
DOI: https://doi.org/10.1145/3383972.3384001

2. TECHNICAL BACKGROUND
2.1 Multiphase Flow Measurement Technology
Multiphase flow online measurement technology has been listed by international energy companies such as BP as one of the five key technologies for determining the success of the future oil and gas industry. Its application covers well test, reserve management, production distribution, flow metering, safety supervision and other processes. Online measurement of oil-water-gas three-phase flow has been a major concern in the field of multiphase flow and

oil production. However, due to the complexity of the multiphase flow process and too many parameters need to be measured, it is difficult to measure accurately.

In recent years, many researchers have tried and introduced new technical methods to improve the accuracy and applicability of oil-water-gas three-phase flow such as venturi pressure drop technology. Venturi pressure drop technology has many advantages such as simple structure, good stability and low cost. So it becomes a useful method used by many research institutions and companies at home and abroad such as FMC, Roxar and Agar [8]. The double differential pressure venturi flow sensor uses the differential pressure between the contraction section and the expansion section of the venturi to realize the resolution of the liquid phase and the measurement of the phase separation flow, which provides a simple and efficient measurement method for the online measurement of gas-liquid two-phase flow without separation [9]. The venturi differential pressure flowmeter consists of three parts: the "contraction section", the "throat" and the "diffusion section". Its technical principle is based on the continuity equation and the Bernoulli equation, the fluid will locally compress when the fluid passes through the "contraction section" of the venturi, the flow rate increases, and the static pressure decreases so that a positive pressure difference is formed before and after the contraction section. When the "diffusion section" passes, the flow rate decreases and the static pressure rises, forming a negative pressure difference. It is further discovered through experiments that the venturi sensor has different ratios of gas and liquid contained in the fluid in the "contraction section" and "diffusion section" differential pressure signals, and the signals obtained are not the same, so they can be installed separately. The differential pressure sensor of the "contraction section" and the "diffusion section" obtains two sets of differential pressure signals (dP1, dP2). Besides, pressure (P) and temperature (T) sensors obtain gas density and combine machine learning algorithms to obtain gas and liquid phase flow. A schematic diagram of the venturi sensor structure is shown in Figure 1.

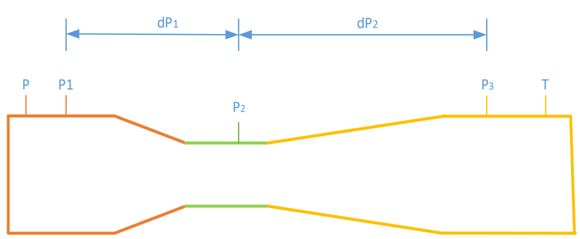


Figure 1. Schematic diagram of the Venturi sensor structure

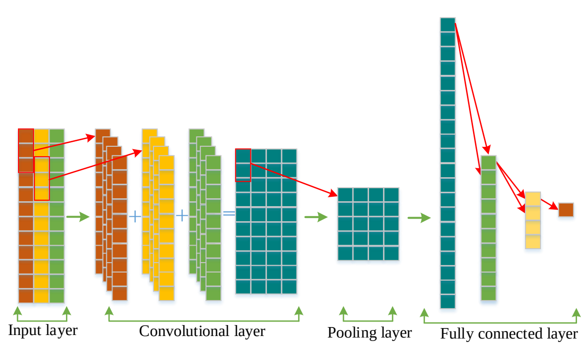


Figure 2 The structure of one-dimensional CNN

2.2 One-dimensional Convolutional Neural Network

Convolutional Neural Network (CNN) is a feedforward neural network with convolutional computation and deep structure. It is one of the representative algorithms of deep learning [10,11]. According to the structure and application fields of convolution kernels in convolutional neural networks, they are also divided into one-dimensional convolutional neural network, two-dimensional convolutional neural network and three-dimensional convolutional neural network. The venturi sensor measurement principle is a one-dimensional data structure for detecting time series such as differential pressure [12]. Therefore, a one-dimensional convolutional neural network structure is used to build an instantaneous flow measurement model.

The one-dimensional convolutional neural network is also composed of input layer, convolutional layer, pooling layer and fully connected layer. Its structure is shown in Figure 2. The following are the detailed steps of calculation.

Input layer. $X = [x_1, x_2, \dots x_t]^T$ is the input layer of the neural network, where $X \in R^{t \times n}$ is the time series data, t is the time series length, n is the dimension of data. x_t is the feature vector of the moment t, $x_t \in R^{1 \times n}$.

Convolutional layer. Sequence X is mapped by a one-dimensional convolution operation can be expressed as $a_c^j = f_r\left(X * W_c^j + b\right)$. $*$ is represented as a one-dimensional convolution operation. a_c^j is j^{th} feature map generated by a convolution kernel W_c^j, $j \in [1, n_c]$ and n_c is the number of convolution kernels. The convolution kernel $W_c^j \in R^{m \times n}$ is the matrix of weights. And m is the size of the convolution kernel, b is the offset. $f_r(\cdot)$ is the activation function, which can provide nonlinear modeling capabilities of the network and realize the nonlinear mapping learning ability of the deep neural network. The common activation functions are "relu", "sigmoid", "tanh", etc.

Pooling layer's main function is feature extraction, dimension reduction, eliminate over-fitting and improve the fault tolerance of the model. The most common pooling operations are average pooling and maximum pooling. Average pooling: Calculate the average of the selected area as the pooled value of the area. Maximum pooling: Calculate the maximum value of the selected area as the pooled value of the area.

Fully connected layer plays the role of classification or regression in the whole convolutional neural network. Its network structure is consistent with the traditional neural network structure. It consists of multiple hidden layers. The fully connected layer further abstracts the global time series features. The combination is finally classified and output through the "softmax" activation function or the regression output is performed via the "relu" activation function. In this paper, the instantaneous flow prediction of gas-liquid two-phase flow is modeled, which is a regression problem, so the "relu" activation function is adopted.

3. MODEL

3.1 Data Acquisition

In order to verify the feasibility of the gas-liquid instantaneous flow model based on the cumulative flow sample label, we carry out a series of experiments based on the actual situation of the oilfield site in the multi-phase flow engineering laboratory platform of Tsinghua University (Figure 3). The experimental procedure is single phase fluid (gas/liquid) - turbine flow meter (acquisition of sample label) - mixer (gas-liquid mixing state) - multiphase flow meter (acquisition of measurement signals) for testing experiments. Multiphase flowmeter measurement data: venturi front differential pressure (contraction section, dP1), venturi differential pressure (diffusion section, dP2), operating pressure(P), fluid temperature(T) total of 4 groups of observation signals. Sample labels are gas-phase flow and liquid phase flow parameters. During the measurement process, since the data acquisition process is a continuous dynamic measurement, the single-phase flow passes through the turbine flow meter and then flows through the multi-phase flowmeter. The measurement data has a time difference, and the gas-liquid pass through the mixer to form a mixed fluid with different transient volume gas content (GVF), thus there is an error between the measurement tag and the fluid flow label that actually flows through the three-phase flowmeter. Therefore, the cumulative flow rate (average flow rate during the conversion period) of each group of sample data is continuously collected for 5 mins to reduce the sample data label's error. In this paper, the experimental samples are divided into the training set, verification set and test set with the ratio of 8:1:1. A specific experimental scheme is shown in Table 1.

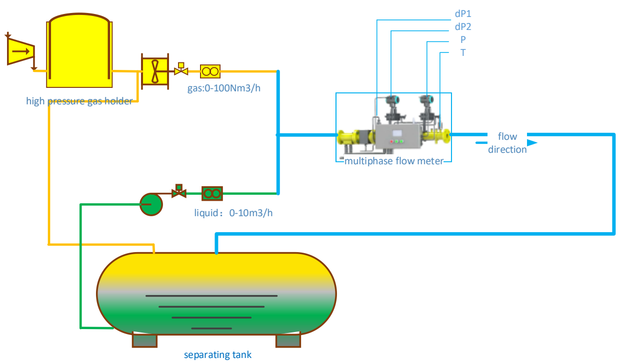
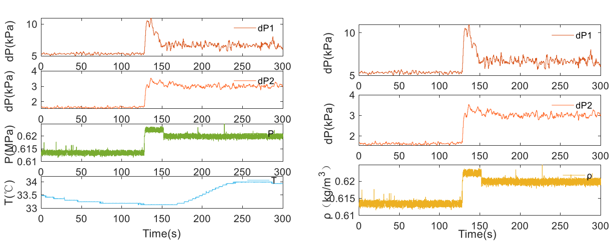


Figure 3. experimental platform

Table 1. experimental scheme

Name	parameters
medium	gas: air
	liquid: an oil-water mixture
flow range	gas： 0-100 Nm^3/h
	liquid： 0-10 m^3/h
measurement parameter	double differential pressure, pressure and temperature
Sampling frequency	measurement data:10Hz, label data:1Hz
Sample duration	single sample duration: 5mins; total sample duration: 125hr



(a)Original data **(b) Processed data**

Figure 4. measurement data of venturi sensor

3.2 Modeling Process

A random set of venturi measurement signals for the 5mins gas-liquid two-phase flow is shown in the Figure (4), and Figure 4(a) shows the four original signals (dP1, dP2, P, T) measured, then the working condition density of the gas(ρ) is calculated by P and T is shown in Figure.4(b), the method could reduce the input data and improve the computational accuracy of the gas flow rate. As can be seen from the changes in dP1 and dP2 in Figure (4), the flow of gas and liquid is constantly changing in 5mins.

An instantaneous gas-liquid flow rate identification model (named: GL_model) uses a two-layer one-dimensional convolution layer and four-layer full connection layer in this article. Figure 5. describes the 1min instantaneous flow calculation model structure. The data input dimension is 600×3 where the 600 represents the length of the signal in the time range and 3 represents the type of signal, respectively dP1, dP2 and ρ. According to actual requirements, the time length can be set to the different scales, e.g., the input signal can be set to 600×3, for the one second instantaneous gas-liquid flow rate identification model.

In GL_model, Conv1D represents the convolution layer, MaxPooling1D represents the pooling layer, and Dense represents the full connection layer, here the Conv1D (32,3,relu) is expressed as one-dimensional convolution, the number of convolution kernel is 32, the size of convolution sum is 3, and "relu" is used as the activation function. Maxpooling1D (2) is expressed in the pooling layer, the maximum value in the two adjacent regions is calculated as the pooled value. Dense(512,relu) means that 512 neurons are output in the fully connected layer, and the "relu" is adopted as the activation function. This GL_model also adopts Flatten to convert the output multidimensional data of the pooling layer into one-dimensional data input to the full connection layer. Besides, dropout regularization rule is also used in the model to solve the overfitting and gradient vanishing problems of deep neural network.

The output of last full connection layer of this GL_model is the 1mins instantaneous flow rates of gas and liquid, but in actual production, there is no instantaneous flow rate label (such as the 1mins flow rate label), but the long term accumulated or average flow label, such as the 5mins accumulated flow rate label, hence, it is impossible to train the GL_model. To solve this problem, the constraint model (named: Ave_Model) of realizing instantaneous flow model training by using the long-term average flow rate labels is shown in Figure 6. The Model is divided into an input layer, slice layer (Lambda), shared layer (GL_model) and Average output layer. In the input layer, sequence data with parameters of 3000×3 is input where 3000 represents the length of the data for 5mins. In the slice layer, the input layer data is sliced into an equal length of time. In this case, input layer data is sliced into 5 parts and the length of time is 1mins. In shared layer, the 5 parts slice data simultaneously calls the 1mins instantaneous model （GL_model）in Figure 5 to calculate the instantaneous flow rate in 1 minute, and in this layer, the weights of each slice call model are shared. In the Average output layer, the five results output from the above shared layer are averaged to output, and the output value corresponds to the average flow rate label of the 5mins. In this model, it is necessary to convert the accumulated flow into the average flow rate during the 5mins.

The above designed instantaneous model and constraint model need to be combined for model training. In the model training process, the constraint model (Ave_Model) uses supervised learning to ensure the accuracy of the model's average flow calculation, and the instantaneous model (GL_model) adopts unsupervised learning to learn the change law of instantaneous flow autonomously and determines the flow rate distribution of the cumulative flow in the instantaneous process [13]. So the GL_model can be constricted intelligently based on the above method. The above model construction method proposed in this paper can also build different instantaneous models according to time scales, such as the 1s, 5s, 10s, 30s, or a shorter time model.

The Keras deep learning framework is adopted in the overall training process of the model. The relevant training parameters are batch size=128, loss='mse', learning rate=0.0001, epoch=3000, optimizer='Adam' and GPU is'2080ti'.

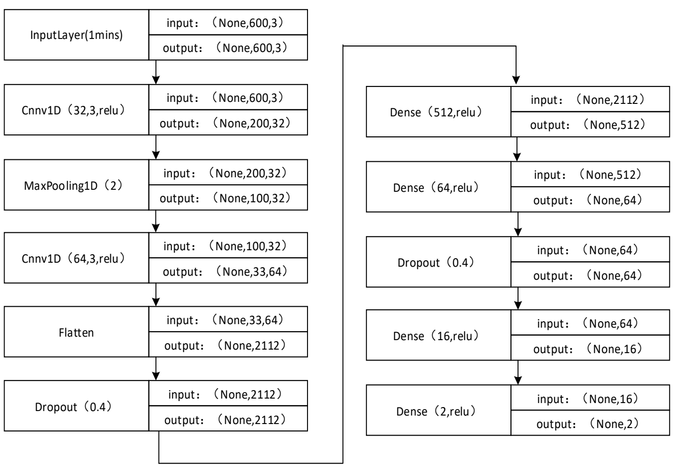


Figure 5. 1mins measurement model

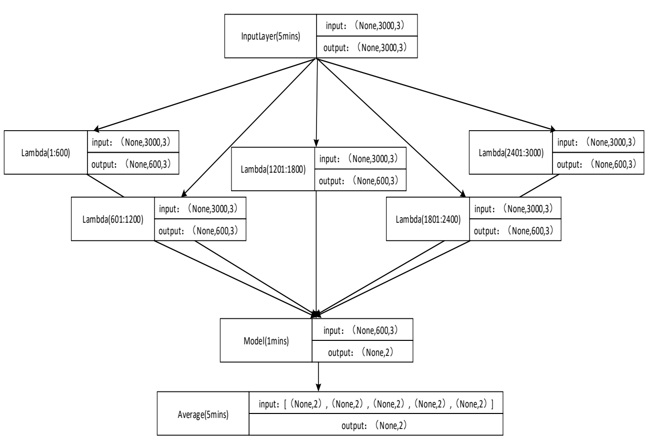


Figure 6. Constraint model

3.3 Model Evaluation

3.3.1 Feasibility evaluation of instantaneous model

Since the experiment only has accurate sample labels of the average flow rate of 5mins, how to verify the accuracy of the instantaneous flow model? The sample processing method is adopted in the article. Firstly, 5 sets of 5mins flow samples are spliced into a 25mins sample, using the average flow rate of 25mins as the model training sample labels. Secondly, we use the method in Figure 5 and Figure 6 to train a 5mins gas phase and liquid phase flow rate model. Finally, the trained 5mins GL_model is tested and evaluated using the accurate 5mins test sample, and the evaluation index using MAPE (mean absolute proportional error). The prediction results and relative error range of 5mins GL_model are shown in Figure 7.

As can be seen from Figure 7, the result is that MAPE = 7.6% for the 5mins liquid phase and MAPE = 4.8% for the 25mins liquid phase average flow rate, MAPE = 18.7% for the 5mins gas phase, and MAPE = 13.4% for the 25mins gas-phase average flow rate. It can be seen from the above test results that the MAPE of the 5mins instantaneous GL_model is higher than the MAPE with 25mins. The reason is that the training process of the 5mins model is an unsupervised learning process, and the comparison results are understandable. However, the 5mins model also achieves a good effect. It is a very important application value for the attention of the instantaneous flow of gas and liquid.

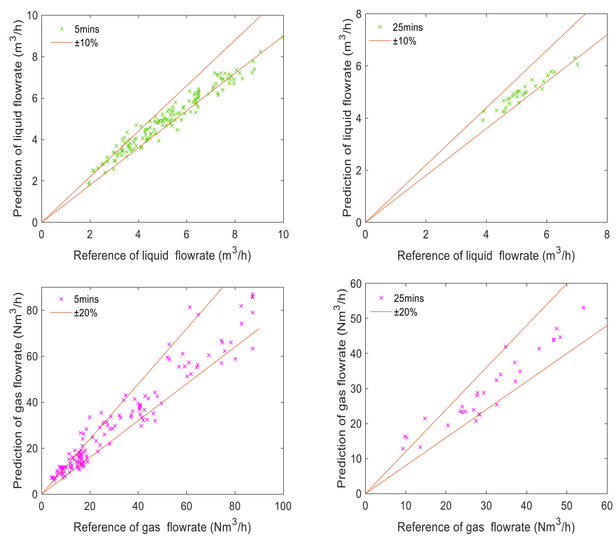
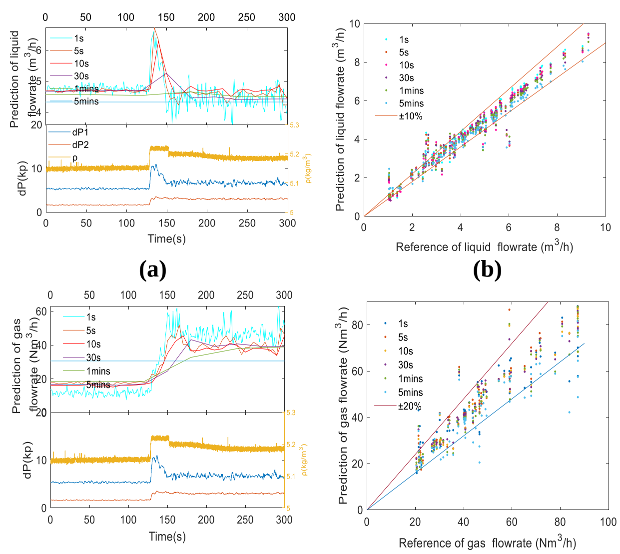


Figure 7. Test results and relative error of 5mins GL_model

3.3.2 Accuracy evaluation of instantaneous model construction

Under the above conditions to verify the feasibility of the instantaneous model, we further analyze the minimum time resolution and accuracy of the instantaneous GL_model. The 5mins time length sample labels are used as the average flow rate to training the different time scale GL_models of 1s, 5s, 10s, 30s, 1mins and 5mins respectively. For the post-training model are also tested and evaluated using test samples of the 5mins average flow rate with the Total parameter (Total number of model parameters), MAPE (mean absolute proportional error) and MAE (mean absolute error). The prediction and comparison results are shown in Figure 8. A set of liquid and gas-phase predictions using different transient models(1s、5s、10s、30s、1mins、5mins) over a 5mins flow variation range are given in Figure 8(a)and Figure 8(c) respectively. From the changes of dp1 and dp2 in the figure, and it is known that the internal fluid has fluctuated greatly during these 5mins. Predicted results for six time-varying instantaneous flow models (1s, 5s, 10s, 30s, 1mins, 5mins) explain that the proposed GL_model method (such as 1s model) can accurately capture the flow changes of fluid, the 5mins model can't achieve. The evaluation results for different time resolution models are shown in Table 2. For Total parameter, the smaller the model parameters, the better the model training and calculation, and the 1mins model parameters are only 1/137 of the 5mins model. For the liquid phase calculation results, the MAPE and

MAE evaluation results of 1s, 5s, 10s, 30s, and 1mins belong to the same level, and the results of these GL_models are better than the 5mins model contracted directly use the original 5mins data samples. For gas-phase instantaneous flow, MAPE and MAE of 1s, 5s, 10s, 30s, 1mins GL_model is also better than 5mins measurement data to build the model. However, for the 1s GL_model, the instantaneous measurement result has a lower accuracy than other instantaneous GL_model for the reason that the sample of 5mins length needs to be sliced into 300 copies to build a 1s model. The excessive number of slices also bring a difficulty to the unsupervised learning for the transient model. Hence, the evaluation index of the time resolution and accuracy of the comparison model need to be considered comprehensively. An exhaustive comparison of the instantaneous GL_model designed in this article, the 5s model is the most suitable for the measurement of instantaneous flow rate in liquid-gas phases, the model has excellent performance in accuracy and can also observe the change of fluid instantaneous flow in a shorter time. In view of the above research, the construction of the instantaneous flow measurement model of gas-liquid two-phase flow proposed in this paper is feasible and has practical application value.



(a) (b) (c) (d)

Figure 8. Instantaneous GL_model prediction results

Table 2. Evaluation results

Model	Total parameter	MAPE: Liquid (%)	MAE: Liquid (m^3)	MAPE: Gas (%)	MAE: Gas (Nm^3)
1s	20,065	7.54	0.251	12.76	5.464
5s	61,025	7.16	0.232	9.25	3.704
10s	101,985	7.37	0.238	8.30	3.568
30s	290,401	7.16	0.230	8.60	3.714
1mins	560,737	7.29	0.236	9.39	4.267
5mins	2,748,001	9.58	0.341	15.31	7.452

4. CONCLUSION

In this paper, an instantaneous flow rate measurement model for gas-liquid two-phase flow based on novel 1D-CNN is proposed. Firstly, a two-layer one-dimensional convolutional layer and four-layer fully connected layer are used to build the instantaneous flow rate measurement model. Then, the long-term average flow label is used to build the constraint model. Finally, the instantaneous model is trained using unsupervised learning with the constraint model. Testing evaluation results are as follows:

(1) Using the method proposed in this paper, the instantaneous flow rate results at different time resolutions can be learned autonomously through long-time cumulative flow labels, and the instantaneous flow changes during the flow can be visually observed.

(2) Instantaneous flow rate model measurements (AMPE, MAE) are better than the 5mins model contracted directly use the original 5mins data sample (shown in Table 2).

(3) The construction method of the instantaneous model proposed in this paper can be applied in other fields and has important application value for industrial process measurement.

5. ACKNOWLEDGMENTS

Our thanks to the support of real-time online three-phase flowmeter platform construction project of CNPC (C11)

6. REFERENCES

[1] Yan Y, Wang L, Wang T, et al. Application of soft computing techniques to multiphase flow measurement: A review[J]. Flow Measurement and Instrumentation, 2018, 60: 30-43.

[2] Fan S, Yan T. Two-phase air–water slug flow measurement in horizontal pipe using conductance probes and neural network[J]. IEEE Transactions on Instrumentation and Measurement, 2013, 63(2): 456-466.

[3] Hu D, Li J, Liu Y, et al. Flow Adversarial Networks: Flowrate Prediction for Gas-Liquid Multiphase Flows Across Different Domains[J]. IEEE transactions on neural networks and learning systems, 2019: 1-11

[4] Zhao C, Wu G, Zhang H, et al. Measurement of water-to-liquid ratio of oil-water-gas three-phase flow using microwave time series method[J]. Measurement, 2019, 140: 511-517.

[5] Chanklan R, Kaoungku N, Suksut K, et al. Runoff Prediction with a Combined Artificial Neural Network and Support Vector Regression[J]. International Journal of Machine Learning and Computing, 2018, 8(1).

[6] Zheng H, Wu C. Predicting Personality Using Facebook Status Based on Semi-supervised Learning[C] //Proceedings of the 2019 11th International Conference on Machine Learning and Computing. ACM, 2019: 59-64.

[7] Kim B, Ye J C. Multiphase Level-Set Loss for Semi-Supervised and Unsupervised Segmentation with Deep Learning[J]. arXiv preprint arXiv:1904.02872, 2019.

[8] Ismail I, Gamio J C, Bukhari S F A, et al. Tomography for multi-phase flow measurement in the oil industry[J]. Flow Measurement and Instrumentation, 2005, 16(2-3): 145-155.

[9] Zhang Q, Xu Y, Zhang T. Wet gas metering based on dual differential pressure of long throat Venturi tube[J]. Journal of Tianjin University, 2012, 45(2): 147-153.

[10] Hinton G E, Salakhutdinov RR. Reducing the dimensionality of data with neural networks[J]. science, 2006, 313(5786): 504-507.

[11] Silver D, Huang A, Maddison C J, et al. Mastering the game of Go with deep neural networks and tree search[J]. nature, 2016, 529(7587): 484-492.

[12] Wang W, Zhu M, Wang J, et al. End-to-end encrypted traffic classification with one-dimensional convolution neural networks[C]//2017 IEEE International Conference on Intelligence and Security Informatics (ISI). IEEE, 2017: 43-48.

[13] Yang N, Zheng Z, Wang T. Model Loss and Distribution Analysis of Regression Problems in Machine Learning[C]//Proceedings of the 2019 11th International Conference on Machine Learning and Computing. ACM, 2019: 1-5.