Fuel xxx (xxxx) xxx

Contents lists available at ScienceDirect

Fuel

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

Full Length Article

Prediction of combustion state through a semi-supervised learning model and flame imaging

Zhezhe Han [a], Jian Li [a], Biao Zhang [a], Md. Moinul Hossain [b], Chuanlong Xu [a,*]

[a] Key Laboratory of Energy Thermal Conversion and Control of Ministry of Education, School of Energy and Environment, Southeast University, Nanjing 210096, China
[b] School of Engineering and Digital Arts, University of Kent, Canterbury CT2 7NT, Kent, UK

ARTICLE INFO

Keywords:
Combustion state prediction
Novel loss function
Denoising autoencoder
Generative adversarial network
Gaussian process classifier

ABSTRACT

Accurate prediction of combustion state is crucial for an in-depth understanding of furnace performance and optimize operation conditions. Traditional data-driven approaches such as artificial neural networks and support vector machine incorporate distinct features which require prior knowledge for feature extraction and suffers poor generalization for unseen combustion states. Therefore, it is necessary to develop an advanced and accurate prediction model to resolve these limitations. This study presents a novel semi-supervised learning model integrating denoising autoencoder (DAE), generative adversarial network (GAN) and Gaussian process classifier (GPC). The DAE network is established to extract representative features of flame images and the network trained through the adversarial learning mechanism of the GAN. Structural similarity (SSIM) metric is introduced as a novel loss function to improve the feature learning ability of the DAE network. The extracted features are then fed into the GPC to predict the seen and unseen combustion states. The effectiveness of the proposed semi-supervised learning model, i.e., DAE-GAN-GPC was evaluated through 4.2 MW heavy oil-fired boiler furnace flame images captured under different combustion states. The averaged prediction accuracy of 99.83% was achieved for the seen combustion states. The new states (unseen) were predicted accurately through the proposed model by fine-tuning of GPC without retraining the DAE-GAN and averaged prediction accuracy of 98.36% was achieved for the unseen states. A comparative study was also carried out with other deep neural networks and classifiers. Results suggested that the proposed model provides better prediction accuracy and robustness capability compared to other traditional prediction models.

1. Introduction

Combustion is an important process in the utilization of fossil fuels and widely exists in large-scale industries such as electric power, chemical industry, and metallurgy, etc. Maintaining an optimal combustion state is crucial to improve combustion efficiency and reduce exhaust pollutants (i.e., NO_x and SO_2) [1]. However, the combustion state often encounters abnormalities and seriously threatening the safety operation due to the dynamic loads, flexible fuels of the combustion processes and harsh operating environments [2]. Therefore, a reliable prediction model of the combustion state is necessary, which would be useful to raise awareness of the drifts or faults of the combustion system in advance. Since the combustion flame has an intrinsically complex and stochastic nature, it is difficult to predict the combustion state accurately. Along with the unceasing development of monitoring techniques, flame imaging incorporating soft-computing algorithms is considered to be a promising method, which has been widely used to determine coal types [3], predict heat output [4], and identify flame structure [5]. Different from the typical monitoring techniques (i.e., pressure and temperature sensors or flame failure detectors), flame imaging-based monitoring techniques can provide more comprehensive measurement information such as temperature distribution, oscillation frequency, etc. [6].

The flame imaging-based monitoring techniques generally involve feature extraction and state prediction steps. Feature extraction from the flame images is an essential step for combustion state prediction and has been studied extensively. The purpose of this step is to represent the raw flame images with lower dimensional data. Yan et al. [7] extracted the flame geometrical and luminous parameters to characterize the combustion state. González-Cencerrado et al. [8] showed that the luminous and spectral parameters are extracted from flame images can characterize the combustion state. Li et al. [9] extracted a set of heterogeneous features such as color, global and local features of the flame image to

* Corresponding author.
E-mail address: chuanlongxu@seu.edu.cn (C. Xu).

https://doi.org/10.1016/j.fuel.2020.119745
Received 12 August 2020; Received in revised form 15 October 2020; Accepted 10 November 2020

Nomenclature

Abbreviations

AE	autoencoder
ANN	artificial neural network
CNN	convolutional neural network
DAE	denoising autoencoder
DL	deep learning
FL	fuel load
GAN	generative adversarial network
GP	gaussian process
GPC	gaussian process classifier
KSVM	kernel support vector machine
LR	logistic regression
LSVM	linear support vector machine
MSE	mean square error
PA	primary air
ReLU	rectified linear unit
RF	random forest
SA	secondary air
SSIM	structural similarity
SVM	support vector machine

Symbols

x_i	the i^{th} original image
x_i	the i^{th} noisy image
z_i	the i^{th} reconstructed image
h_i	deep feature of the i^{th} flame image
y_i	the label of the i^{th} flame image
$h_\star$	deep feature of the test flame image
$y_\star$	the label of the test flame image
F	flattened operation
f	fully connected operation
C	convolution operation
Q	the number of convolution filter
c × c	the size of convolution filter
q	the stride of convolution filter
P	pooling operation
o × o	the size of pooling kernel
l	the stride of pooling kernel
U	upsampling operation
g × g	the size of upsampling filter
H	the labeled image dataset
p	prediction probability
K	covariance function
L_{MSE}	mean square error loss function
L_{SSIM}	structural similarity loss function

Greek letters

δ	corruption ratio
τ	normal distribution random variable
λ	hidden neuron
η	grey level of the image
ε	constant
μ	mean intensity
σ	standard deviation
Σ	diagonal matrix
φ	prediction accuracy
ρ	precision
ν	recall

predict the burning state of rotary kiln combustion. Chen et al. [10] utilized the principal component analysis (PCA) to obtain the principal components or combinational variables of flame images, which are then used to describe the important variations of the oxygen content in the combustion process. Although PCA is recognized as an effective reduction method of data dimensionality, the linear transformation can cause lower accuracy [11]. The kernel principal component analysis (KPCA) was further proposed to extract nonlinear features through nonlinear mapping [12], and its effectiveness is verified through 9 MW heavy oil-fired combustion flames. Hence, extracting representative flame features is the key factor for achieving satisfactory prediction performance. However, these traditional feature extraction methods have some disadvantages, such as (i) require prior knowledge and the selection of hand-crafted features heavily depends on the professional experience; (ii) lack of generalization performance; (iii) poor robustness where shallow features are only suitable for specific diagnostic tasks and (iv) response lag and lower efficiency. Also, traditional methods require more memory and computational power for processing big data. Therefore, a more reliable and advanced data-driven method is further required to extract the flame features accurately.

Deep learning (DL) is one of the breakthroughs and representative techniques in the field of artificial intelligence (AI), which has the potential to overcome the shortcomings of traditional feature extraction methods [13]. Also, the DL is considered as an effective method for discriminative features extraction from a large amount of data through multi-layer nonlinear transformation, which can then be used for accurate classification and recognition. Among the DL techniques, the convolutional neural network (CNN) has been proved to be useful for extracting high-level features due to its unique convolution operation [14]. Especially, the representative features in different hidden layers of CNN are automatically learned by the training process rather than being designed manually, which significantly gets rid of the dependence on expert knowledge. Recently, CNN has been successfully applied in the field of combustion process monitoring. Wang et al. [15] utilized a CNN model to extract flame features for predicting burning states of power plant furnace. Abdurakipov et al. [16] established a CNN model based on labeled flame images, which is used to predict the combustion regimes of a laboratory-scale swirling gas burner. Even though various progress has been made through CNN-based models, one non-trivial problem in these models is that a larger amount of labeled data is needed for training. The prediction accuracy also depends on the scale and quality of the labeled data [17]. Whereas, manual labeling requires a significant human effort, which undoubtedly time-consuming and not cost-effective. Most importantly, human errors can be involved in the process of labeling. Thus, the performance of the prediction model can severely be deteriorated.

The unsupervised learning networks such as the deep belief network (DBN) and autoencoder (AE) can solve the aforementioned problems gradually. Currently, the unsupervised learning networks have been applied for combustion process monitoring. For instance, Liu et al. [18] developed a multi-layer DBN to extract deep features of flame images for predicting the oxygen content of a heavy oil combustion process. Lyu et al. [19] proposed a DBN-based deep learning framework to extract representative features of heavy oil-fired boiler furnace flame images and identified the combustion process faults. Qiu et al. [20] adopted a convolution AE network to extract flame features for the classification of pulverized coal furnace combustion status. Akintayo et al. [21] proposed an end-to-end convolutional selective AE framework to detect the combustion states of laboratory-scale swirl combustor through flame imaging. In general, the unsupervised learning network uses multi-hidden layers to extract the high-level data feature. However, these multi-hidden layers often create training difficulties. This problem can be resolved by improving the loss function rather than simply using the traditional loss function such as mean square error (MSE). Recently, Han

et al. [22] proposed a deep neural network based on the stack sparse autoencoder (SSAE) for feature extraction of flame images, which has been successfully applied to the laboratory-scale ethylene combustion flame. However, the SSAE model is composed of relatively independent two-level networks. The designed structure suffers the global optimization of network parameters.

In recent years, various classification algorithms such as artificial neural networks (ANNs) and support vector machine (SVM), have widely been applied to predict the combustion states through flame imaging. The ANN usually coupled a CNN fully connected layer [23] and the network can predict the combustion states based on the labeled images by fine-tuning. However, the ANN suffers weaknesses such as difficult to determine its architecture and unreliable to deal with small sample cases [24]. The SVM is a statistical learning method with extremely fast learning speed and good generalization ability [25]. Nevertheless, the SVM also has some disadvantages, e.g., high computational complexity with the increase of input variables and the danger of over-fitting with limited dataset [26]. Besides, the hyper-parameters (i.e., the hidden layer number of ANN and punishment coefficient of SVM) of these classification networks have a significant effect on the final prediction performance, whereas considerable skill and experience are required for appropriate parameter selection. Furthermore, these hyper-parameters have internal dependencies, which make them particularly cumbersome for fine-tuning in combination with prediction accuracy [27]. Although grid search (GS) [28], genetic algorithm (GA) [29] and particle swarm optimization (PSO) [30] algorithms opened a solution to obtain the optimal hyper-parameters, they undoubtedly increase the complexity of classifier design. On the contrary, the Gaussian process classifier (GPC) [31] has the potential to overcome these obstacles. It is a non-parametric classifier that does not require complicated hyper-parameters selection process, and it is suitable for dealing with small sample learning problems. More importantly, the GPC is a probabilistic model that provides probabilistic outputs which are valuable to recognize the changes of the combustion state. The Gaussian process regression has been applied for predicting the oxygen content [10] and equivalence ratio estimation [32] through flame imaging. However, the GPC is rarely adopted for classification purposes such as combustion state prediction, so it is worth further exploration.

In this study, a novel semi-supervised learning model is proposed to predict the combustion state. The model is developed through a combination of DAE, GAN and GPC. The DAE network trained by the adversarial learning mechanism of the GAN to extract discriminate features of the flame images. Different from traditional MSE based on the pixel by pixel error, the structural similarity (SSIM) metric, which focuses on image structure information, is utilized as a loss function of the DAE network to improve the feature extraction efficiency. Afterwards, the extracted features are feed into the GPC to predict the combustion state. Experiments were carried on a 4.2 MW heavy oil-fired boiler furnace to capture flame images under different combustion states. The semi-supervised learning model is evaluated and validated through original (seen) and new combustion (unseen) states. Results obtained from the validation and evaluation are presented and discussed.

2. Technical strategy

2.1. Overall framework of the prediction model

Fig. 1 shows the overall strategy of the proposed semi-supervised learning model for the combustion state prediction. The strategy mainly consists of two stages, i.e., Stage 1: data acquisition, preprocessing and sorting; Stage 2: establishment of the semi-supervised learning model. A brief of each stage is given below:

Stage 1: An imaging system is used to acquire flame images under different combustion states. All acquired images are resized to 256 (H) × 256 (V) and normalized to a range of 0 to 1. After preprocessing, the flame images categorized into original states (seen) and new states (unseen). The images from original states are used to train the DAE-GAN. The proposed semi-supervised learning model, i.e., DAE-GAN-GPC is evaluated through the images of the new states. Detailed descriptions of the training and evaluation process can be found in Section 4.

Stage 2: A semi-supervised learning model combining unsupervised and supervised learning networks is established to extract flame representative features and combustion state prediction. For unsupervised learning, the DAE-GAN is established with parameter initialization and then trained by unlabeled images from original states. The DAE-GAN is then used to extract deep features of labeled images from the original states. In supervised learning, the extracted features are used to train the supervised GPC. Afterwards, the original combustion states are predicted through the GPC. For new combustion states, the trained GPC is fine-tuned with a small portion of labeled images from new states. The tuned GPC is then used to predict the combustion states of the original and new states accurately. Detailed descriptions of the semi-supervised

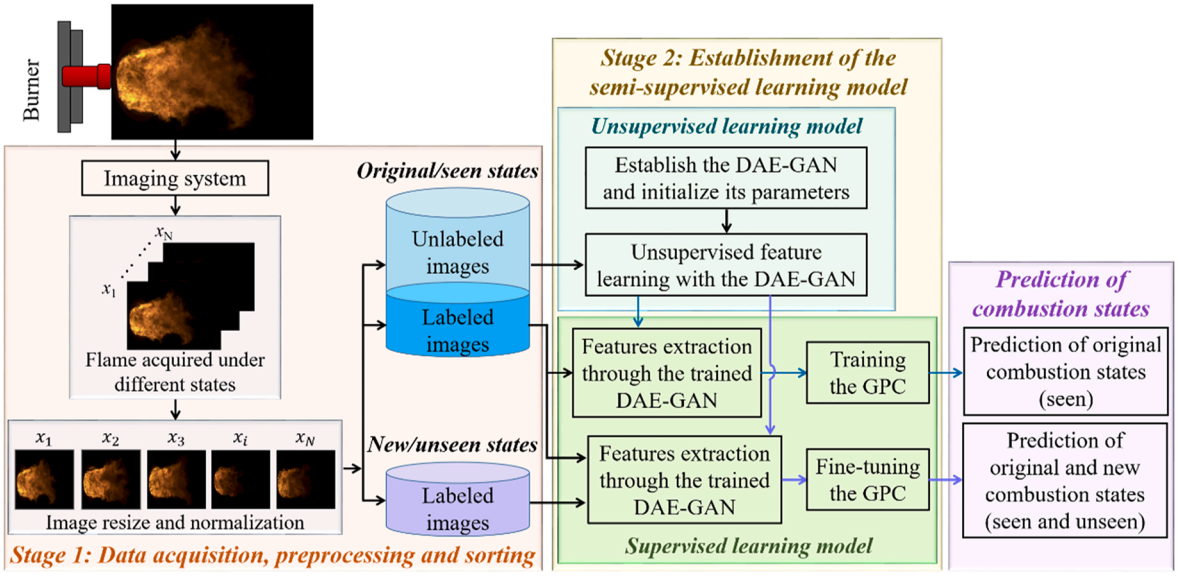


Fig. 1. The overall strategy of combustion state prediction.

learning model can be found in Section 2.2.

2.2. The semi-supervised learning model

Fig. 2 demonstrates the architecture of the semi-supervised learning model, which is consists of DAE-GAN and GPC. The preprocessed flame images are used as the input of the semi-supervised learning model. A multi-layer denoising autoencoder (DAE) is developed to extract discriminative features of the input flame images. To achieve more representative features, the DAE is integrated with the generative adversarial network (GAN). The DAE-GAN is then trained through unlabeled flame images. Once the DAE-GAN training is completed, flame representative features are extracted and these extracted features are used to train the supervised GPC and predict the combustion state. The detailed description of the individual network of the semi-supervised learning model are discussed as follows:

2.2.1. Autoencoder

The autoencoder (AE) is a symmetrical neural network and

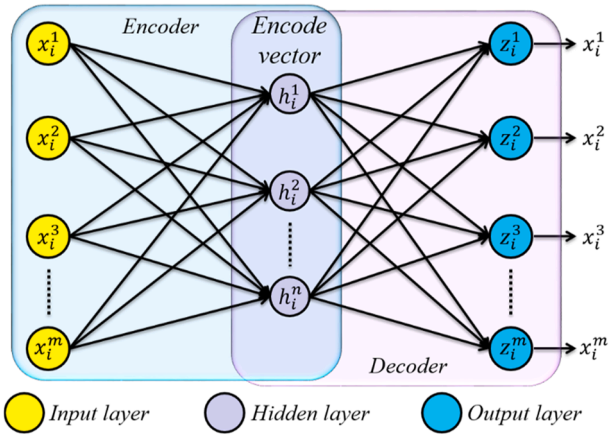


Fig. 3. The basic structure of the autoencoder.

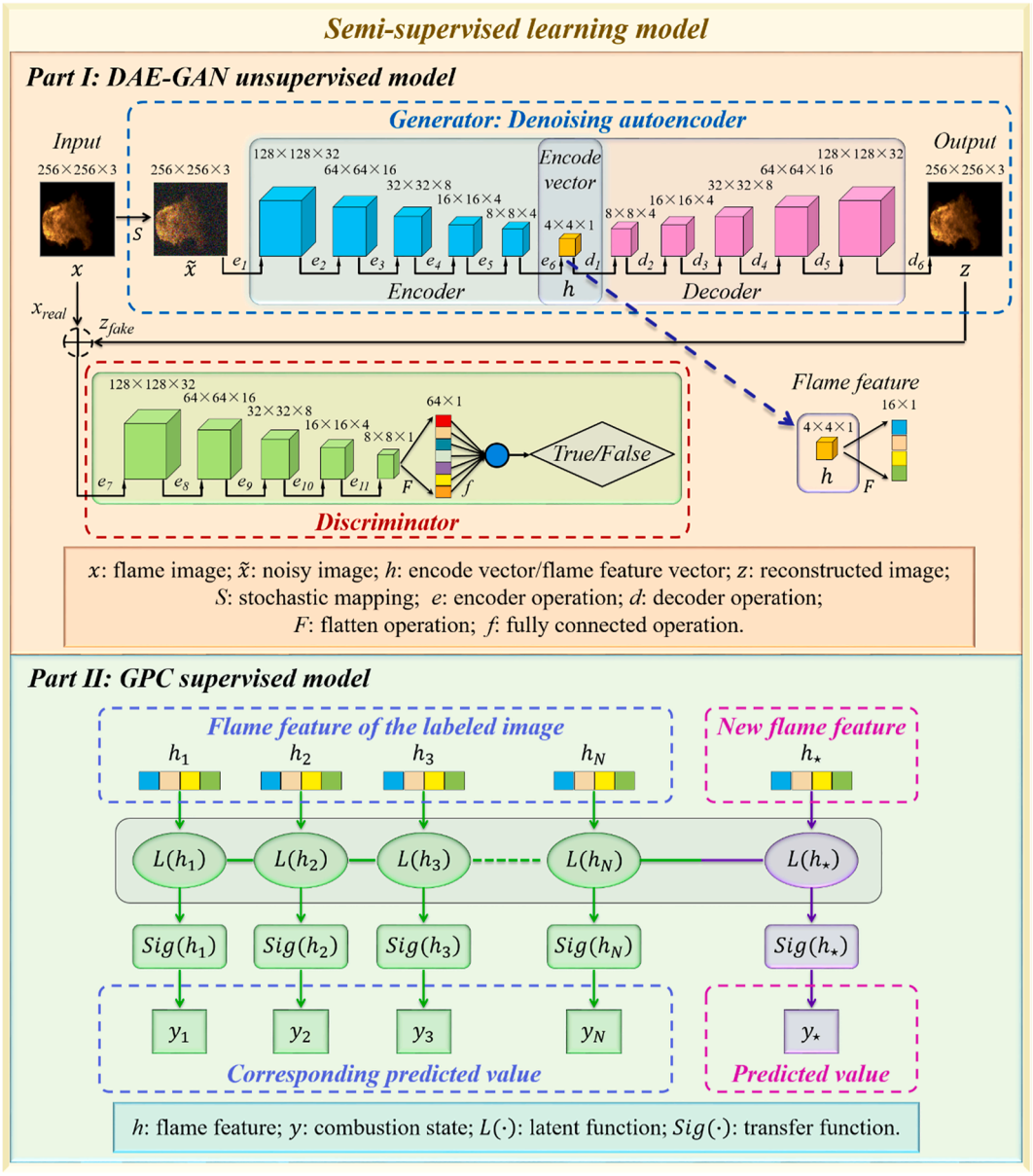


Fig. 2. The structure of the proposed semi-supervised learning model.

composed of fully connected layers. The structure of a basic AE network is shown in Fig. 3. Assuming an unlabeled dataset $\{x_i\}_{i=1}^{M}$ of M samples, the AE network transforms m-dimensional input sample x_i into then-dimensional encode vectorh_i(called feature vector) through encoder operation, which is typically $m > n$. Then, the encode vector h_i is mapped back into the$m$-dimensional output sample z_i through decoder operation. The weight parameters of the AE network can be optimized by minimizing reconstruction error (known as loss function) between inputs and outputs.

2.2.2. Denoising autoencoder

Although the basic AE network can reconstruct the input information correctly, it may simply copy the information from the input layer to the hidden layer [33]. In this case, the effective feature learning ability and robustness of the network cannot be guaranteed. To solve this issue, denoising coding is integrated into the basic AE to form the denoising autoencoder (DAE), whose input samples are corrupted by random noise. The decoder of the DAE network reconstructs the encode vector to obtain the noise-free sample. Fig. 4 illustrates the structure of the DAE, where x_i is the i^{th} noisy sample with the dimension of m.

Unlike the AE network, the DAE network establishes a nonlinear mapping relationship between the reconstructed sample and the noisy sample, rather than the original sample. Due to this constraint, simple replication of input information can be avoided and the robustness of the network will be improved. To mine deeper information of the input data, the DAE network usually uses multi-hidden layers. However, due to a large number of parameters caused by multi-hidden layers, the fully connected DAE network is prone to gradient vanishing or gradient explosion during the training process [34].

2.2.3. Denoising autoencoder-generative adversarial network

To overcome the shortcomings of the DAE network, two improvement strategies can be considered. First, the fully connected operation is replaced by a convolution operation. The convolution operation has the characteristics of sparse connectivity and weight sharing, which greatly reduces the number of network parameters to be learned. The second improvement is that the DAE network is integrated with the GAN. The GAN mainly uses a novel training framework which is consists of a generator and a discriminator. Any discrepant functions such as multilayer perception can be used to represent the generator and discriminator [35]. Therefore, the DAE network can be considered as the generator of the GAN.

Fig. 2 (*PartI*) shows the architecture of the constructed DAE-GAN. The noisy image x is obtained by adding Gaussian noise with a fixed *corruption ratio* δ to the original image x. The process can be defined as:

$$x = x + \delta\tau \tag{1}$$

where τ is a normal random distribution within a range of −2.576 to 2.576 with 99% probability. As shown in Fig. 2 (*PartI*), the noisy image x is mapped to the encode vector h (flame feature vector) through a series of encoders, and then remapped to reconstructed image$z$ through a series of decoders. Finally, the flame image can be represented by the extracted 16-dimensional feature h (4 × 4 × 1). The detailed configuration of the DAE-GAN is illustrated in Table 1.

In the encoder, firstly, the noisy sample x is processed by a convolution layer in the encoder e_1, which has 32 filters and each filter size is 3 × 3 and a stride of 1. The feature maps are generated by sliding multiple filters over the complete input sequence through a convolution layer. Once the convolution operation is completed, an activation function is utilized to implement nonlinear transformation and to improve the feature representation ability and divisibility [36]. The rectified linear unit (ReLU) function [i.e., $y(\lambda) = max(0,\lambda)$, λ represents hidden neuron] is used as the activation function due to its faster training convergence compared to Sigmoid and TanH functions [37]. However, the Sigmoid function [i.e., $y(\lambda) = 1/1 + \exp(-\lambda)$] is still used in d_6 operation (6th decoder) and f operation to ensure the output range 0 to 1. Afterwards, the pooling operation is carried out to reduce the network parameters and to improve the translation invariance. In this study, the max-pooling method [38] with a size of 2 × 2 and a strider of 2 is used to reduce dimension, and the feature map with the dimension of 128 × 128 × 32 is obtained. Finally, after a series of similar encoder operations, the input flame image x with the dimension of 256 × 256 × 3 is represented by the flame feature h with the dimension of 4 × 4 × 1 (named as a feature extraction process in this study). The extracted flame feature h is further flattened into the 16-dimensional vector to prepare for subsequent combustion state prediction.

In the decoder d_1, the flame feature h is dimensionally extended by the upsampling layer and then processed through the convolution layer and activation function. Finally, after a series of similar decoder operations, the output flame image with the dimension of 256 × 256 × 3 is generated which is defined as an image reconstruction process.

The discriminator operation processes are like an encoder. Except that, the F represents the flattened operation and it is used to convert the dimension of the feature map from 8 × 8 × 1 to 64 × 1, and the f represents the fully connected operation which is used to judge whether the input of the discriminator is the ground-truth image (true) or generated image (false).

During the DAE-GAN training process, in each iteration, the generator DAE produces some fake samples. The discriminator is trained by these fake samples along with a few true samples. Then the generator is rewarded for generating examples to “muddling through” the discriminator [39]. Through the adversarial training mechanism, the generator and discriminator continuously confront each other and optimize themselves until achieving the Nash equilibrium [40]. Specifically, generator (G) and discriminator (D) are competitors in a min–max adversarial game with the following loss function:

$$L_{GAN}(D,G) = \min_{G}\max_{D}\left\{E_{P_{data}(x)}\log D(x) + E_{P_g(x)}\log[1 - D(G(x))]\right\} \tag{2}$$

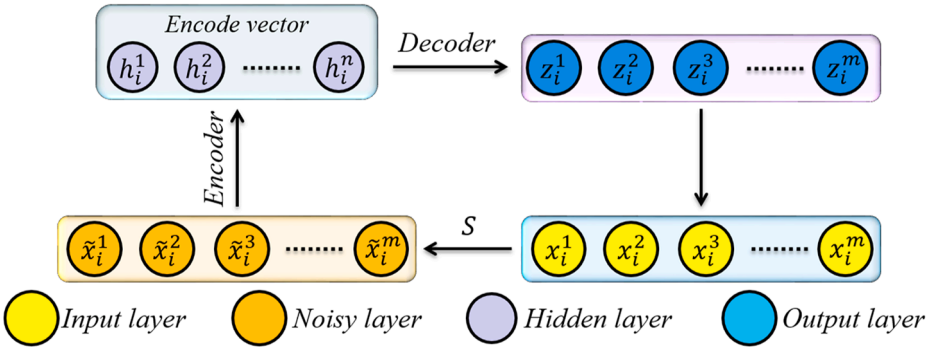


Fig. 4. The structure of the denoising autoencoder. *S*: stochastic mapping.

Table 1
The DAE-GAN configuration.

Items	Type of operation	Upsampling layer	Convolution layer	Activation function	Pooling layer	Output dimension
Input	/	/	/	/	/	256 × 256 × 3
	S	/	/	/	/	256 × 256 × 3
Encoder	e_1	/	$C(32@3\times3+1)$	*ReLU*	$P(2\times2+2)$	128 × 128 × 32
	e_2	/	$C(16@3\times3+1)$	*ReLU*	$P(2\times2+2)$	64 × 64 × 16
	e_3	/	$C(8@3\times3+1)$	*ReLU*	$P(2\times2+2)$	32 × 32 × 8
	e_4	/	$C(4@3\times3+1)$	*ReLU*	$P(2\times2+2)$	16 × 16 × 4
	e_5	/	$C(4@3\times3+1)$	*ReLU*	$P(2\times2+2)$	8 × 8 × 4
	e_6	/	$C(1@3\times3+1)$	*ReLU*	$P(2\times2+2)$	4 × 4 × 1
Decoder	d_1	$U(8\times8)$	$C(4@3\times3+1)$	*ReLU*	/	8 × 8 × 4
	d_2	$U(16\times16)$	$C(4@3\times3+1)$	*ReLU*	/	16 × 16 × 4
	d_3	$U(32\times32)$	$C(8@3\times3+1)$	*ReLU*	/	32 × 32 × 8
	d_4	$U(64\times64)$	$C(16@3\times3+1)$	*ReLU*	/	64 × 64 × 16
	d_5	$U(128\times128)$	$C(32@3\times3+1)$	*ReLU*	/	128 × 128 × 32
	d_6	$U(256\times256)$	$C(3@3\times3+1)$	*Sigmoid*	/	256 × 256 × 3
Output	/	/	/	/	/	256 × 256 × 3
Discriminator	e_7	/	$C(32@3\times3+1)$	*ReLU*	$P(2\times2+2)$	128 × 128 × 32
	e_8	/	$C(16@3\times3+1)$	*ReLU*	$P(2\times2+2)$	64 × 64 × 16
	e_9	/	$C(8@3\times3+1)$	*ReLU*	$P(2\times2+2)$	32 × 32 × 8
	e_{10}	/	$C(4@3\times3+1)$	*ReLU*	$P(2\times2+2)$	16 × 16 × 4
	e_{11}	/	$C(1@3\times3+1)$	*ReLU*	$P(2\times2+2)$	8 × 8 × 1
	F	/	/	/	/	64 × 1
	f	/	/	*Sigmoid*	/	1
Flame feature	F	/	/	/	/	16 × 1

In Table 1: S: stochastic mapping; e_a: the a^{th} encoder; d_b: the b^{th} decoder; F: flattened operation; f: fully connected operation; C: convolution operation; P: pooling operation; U: upsampling operation; '/': no operation. $C(Q@c\times c+q)$ represents the convolution layer and it has Q filters. Each filter scans the input neurons with a fixed size of $c\times c$ and a stride of q. $P(o\times o+l)$ represents the pooling layer that condenses the feature map by selecting a maximum value with an $o\times o$ transformation kernel and a step of l. $U(g\times g)$ represents the upsampling layer that extends the feature dimension to $g\times g$.

where $D(\cdot)$ represents the output of discriminator; $G(\cdot)$ represents the output of the generator; E is the empirical estimate of the expected value of the probability; $P_{data}(x)$ and $P_g(x)$ represent the probability distribution of original image x and noisy image x, respectively. It can be proved that when $P_{data}(x)=P_g(x)$, the new samples generated by the generator perfectly fits the original sample distribution.

The mean square error (MSE) is often used as a loss function of the DAE network and defined as:

$$L_{MSE}=\frac{1}{T}\sum_{t=1}^{T}\|z_t-x_t\|^2 \tag{3}$$

where T represents the number of unlabeled images in the training dataset; z_t is the t^{th} reconstructed image and x_t is the corresponding original image. Although the MSE can quantify the difference between the output and the reference image, it is based on the pixel by pixel error of the image without considering the structure information. To consider the structural information, the structural similarity (SSIM) metric [41] is used as a novel loss function for parameter optimization of DAE. Hence, the loss function L_{SSIM} is defined as:

$$L_{SSIM}=1-\frac{1}{T}\sum_{t=1}^{T}SSIM(x_t,z_t) \tag{4}$$

where $SSIM(x,z)$ is composed of luminance, contrast ratio and structural factor, and it can be expressed as:

$$SSIM(x,z)=\frac{(2\mu_x\mu_z+\theta_1)(2\sigma_{xz}+\theta_2)}{(\mu_x^2+\mu_z^2+\theta_1)(\sigma_x^2+\sigma_z^2+\theta_2)} \tag{5}$$

where μ_x and μ_z are the mean intensity of x and z, respectively; σ_x and σ_z are the standard deviation of x and z, respectively; σ_{xz} is the correlation coefficient between x and z; $\theta_1=(\varepsilon_1\eta)^2$ and $\theta_2=(\varepsilon_2\eta)^2$, where η is the grey level of the image (0–255 for an 8-bit monochrome image); $\varepsilon_1\ll1$ and $\varepsilon_2\ll1$ are constants. The SSIM metric indicates the structural mismatch between the original image and the reconstructed image in the range of 0 to 1, where 0 for entirely dissimilar and 1 for precisely similar. Detailed prediction performance by the novel loss function L_{SSIM} can be found in Section 4.3.2. Finally, the loss function of the proposed DAE-GAN is the combination of Eqs. (2), (3) and (4) and defined as:

$$L_{DAE-GAN}=L_{GAN}(D,G)+L_{MSE}+L_{SSIM} \tag{6}$$

Once the training procedure of the DAE-GAN is completed, the generator DAE can learn the probability distribution of the original data sample effectively and extract the deep features of the flame image accurately.

2.2.4. Gaussian process classifier

A Gaussian process classifier (GPC) is used to predict the combustion state through extracted deep features of the flame image. Like other supervised neural network classifiers such as ANN and SVM, the GPC is established based on the labeled flame images to predict the combustion states of unknown samples. The difference is that the GPC is a non-parametric classifier, no assumption of structural form is required between the input variables and the output prediction [42].

Fig. 2 (*Part II*) shows a structure of the GPC that used for combustion state prediction. For instance, an observation dataset $\{H,Y\}$, where $H=\{h_i\}_{i=1}^{N}$, $Y=\{y_i\}_{i=1}^{N}$, N represents the number of labeled images, h_i represents the flame feature of the i^{th} labeled image, and y_i represent the corresponding combustion state. The GPC is a binary classifier and typically used for two-category classification tasks, i.e.,$y_i\in\{-1,1\}$. Where $y_i=-1$ represents the predicted value is the first category, while $y_i=1$ represents the predicted value is the second category. The probability of the flame feature h_i belongs to the second category can be calculated by:

$$p(y_i=1)=1-p(y_i=-1)=Sig(L_i) \tag{7}$$

where $Sig(\cdot)$ represents transfer function and often used the Sigmoid function; $L_i=L(h_i)$ represents latent function. For a given input dataset H, the Gaussian process (GP) function$\boldsymbol{L}$ can be specified as:

$$\boldsymbol{L}\ GP(\mu(H),K(H,H)) \tag{8}$$

where $\boldsymbol{L}=\{L(h_1),L(h_2),\cdots,L(h_N)\}$is the collection of the latent functions obeying a multivariate Gaussian distribution and $\mu(H)$is the mean function usually set to zero for simplified calculation; $K(H,H)$ is the

covariance function that can be calculated by:

$$K(H,H)=\begin{bmatrix} k(h_1,h_1) & k(h_1,h_2) & \cdots & k(h_1,h_v) & \cdots & k(h_1,h_N) \\ k(h_2,h_1) & k(h_2,h_2) & \cdots & k(h_2,h_v) & \cdots & k(h_2,h_N) \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ k(h_u,h_1) & k(h_u,h_2) & \cdots & k(h_u,h_v) & \cdots & k(h_u,h_N) \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ k(h_N,h_1) & k(h_N,h_2) & \cdots & k(h_N,h_v) & \cdots & k(h_N,h_N) \end{bmatrix} \tag{9}$$

where $k(h_u,h_v)$ is the covariance between two feature variables that determine how the response at one input h_u is affected by the responses at another input h_v. The automatic relevance determination (ARD) is commonly used as a covariance function and defined as [43]:

$$k(h_u,h_v)=\sigma_s^2\exp\left(-\frac{\|h_u-h_v\|^2}{2\ell_f^2}\right) \tag{10}$$

where σ_s^2 is the signal variance; ℓ_f^2 is the characteristic length-scale parameter. The parameters $\vartheta=\{\sigma,\ell\}$ can be determined by using maximum likelihood estimation.

Denote $h_\star$ as the deep feature of a flame image in the testing/prediction dataset. The predictive probability $p(y_\star=1)$ means the $h_\star$ belong to the second category of the combustion state, which is monotonically related to the output of the latent function $L_\star$ [$L_\star = L(h_\star)$]. The posterior probability $p(y_\star=1|H,Y,h_\star)$ can be calculated by:

$$p(y_\star=1|H,Y,h_\star)=\int Sig(L_\star)p(L_\star|H,Y,h_\star)dL_\star \tag{11}$$

$$p(L_\star|H,Y,h_\star)=\int p(L_\star|H,h_\star,\boldsymbol{L})p(\boldsymbol{L}|H,Y)d\boldsymbol{L} \tag{12}$$

The probability $p(\boldsymbol{L}|H,Y)$ can be calculated by using the Bayesian inference and defined as:

$$p(\boldsymbol{L}|H,Y)=\frac{p(Y|\boldsymbol{L})p(\boldsymbol{L}|H)}{p(Y|H)} \tag{13}$$

where $p(Y|H)$ represents the marginal likelihood and defined by:

$$p(Y|H)=\int p(Y|\boldsymbol{L})p(\boldsymbol{L}|H)d\boldsymbol{L} \tag{14}$$

where $p(\boldsymbol{L}|H)$ represents the Gaussian prior distribution of $\boldsymbol{L}$ $GP(0,K(H,H))$. The joint likelihood function is calculated by:

$$p(Y|\boldsymbol{L})=\prod_{i=1}^{N}p(y_i=1|L_i) \tag{15}$$

where $p(y_i=1|L_i)$ is the standard density distribution function. Note that neither the posterior probability nor marginal likelihood can be calculated analytically, so an approximation is required. Thus, the expectation propagation (EP) algorithm [31] is applied in Eq. (11) to obtain Gaussian approximation of the posterior distribution. This yields the following result:

$$p(y_\star=1|H,Y,h_\star)=\Phi\left(\frac{k_\star(K(H,H)+\Sigma)^{-1}\mu}{\sqrt{1+k(h_\star,h_\star)-k_\star(K(H,H)+\Sigma)^{-1}k_\star}}\right) \tag{16}$$

where $\Phi(\cdot)$ denotes the cumulative density function of the standard normal distribution; $k_\star=[k(h_1,h_\star),\cdots k(h_N,h_\star)]^{-1}$ collects covariances between training inputs H and $h_\star$;$\mu=\left\{\mu_i\right\}_{i=1}^{N}$, and Σ is a diagonal matrix with $\Sigma_{ii}=\sigma_i^2$, where μ_i and σ_i^2 are site parameter. In general, $p(y_\star=1|H,Y,h_\star)=0.5$ is the classification boundary of the GPC. When $p(y_\star=1|H,Y,h_\star)>0.5$, the input $h_\star$ belongs to the second category.

Furthermore, this binary GPC can solve multi-category prediction tasks. For example, the combustion states can be grouped into $\mathscr{W}$ categories, which is,$y_i\in\{1,2,\cdots,\mathscr{W}\}$, then

$$y_i^w=\begin{cases}1, & if\ y_i=w, w=1,\cdots,\mathscr{W}\\ -1, & otherwise\end{cases} \tag{17}$$

where y_i^w is the binary output and it indicates that the input h_i belongs to the w^{th} category of the combustion state. Define $Y^w=\{y_i^w\}_{i=1}^N$ and then implement the binary GPC individually based on the dataset H and Y^w for $w=1,\cdots,\mathscr{W}$. This would lead to the total number of binary GPCs is $\mathscr{W}$. Finally, the predictive probability of the new flame feature $h_\star$ belongs to the w^{th} category can be achieved by:

$$p(y_\star^w=1|H,Y^w,h_\star)=\int Sig(L_\star)p(L_\star|H,Y^w,h_\star)dL_\star \tag{18}$$

The index of the largest $p(y_z^w=1|H,Y,h_\star)$ estimates the most-likely category of the input flame feature. For example, if $p(y_\star^1=1|H,Y,h_\star)$ is the largest, it implies that the current input flame feature belongs to the first category. In contrast, if $p(y_\star^2=1|H,Y,h_\star)$ is the largest, it implies that the current input flame feature belongs to the second category. Thus, the GPC can provide predictive distributions rather than merely point predictions, which is a great benefit for an in-depth understanding of the probability of the flame feature corresponding to different combustion states. More technical details about the Gaussian process classifier can be found elsewhere in [31].

2.3. Performance evaluation

The prediction performance of the semi-supervised learning model is evaluated through the testing dataset. The prediction accuracy is calculated as follows:

$$\varphi=\frac{\alpha}{\beta}\times 100\% \tag{19}$$

where φ represents the prediction accuracy; α is the number of correctly predicted samples; β is the number of testing samples. Furthermore, three different metrics are also employed to evaluate the prediction performance of the proposed model quantitatively. These metrics are:

$$precision(\wp)=\frac{TP}{TP+FP} \tag{20}$$

$$recall(\wp)=\frac{TP}{TP+FN} \tag{21}$$

$$F_1-score=2\times\frac{\wp\times\wp}{\wp+\wp} \tag{22}$$

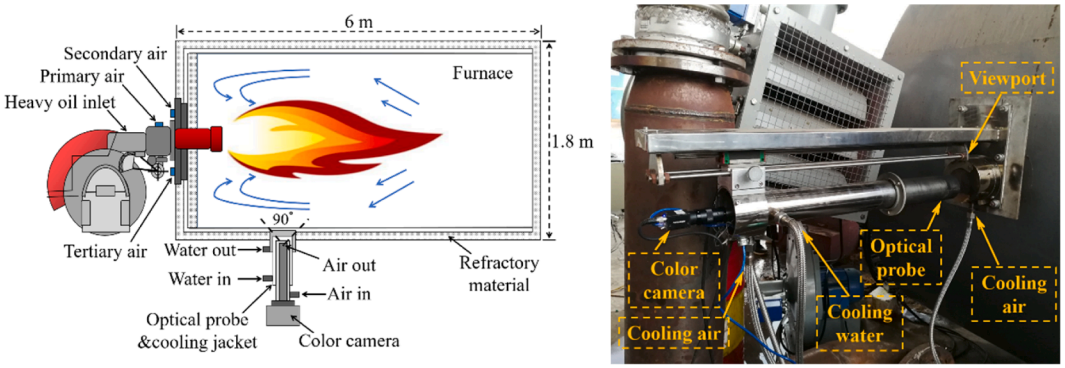
where TP is the number of true positives; FP is the number of false positives; FN is the number of false negatives. F_1-*score* is the harmonic average based on *precision* and *recall*, whose value ranges from 0 to 1 [22]. Notably, if any output of the above evaluation metrics is closed to 1, it indicates that the semi-supervised learning model has a strong prediction ability.

3. Experiments on a 4.2 MW heavily oil-fired boiler furnace

3.1. Experimental setup

To evaluate the effectiveness of the semi-supervised learning model for predicting the combustion state, experiments were carried out on a 4.2 MW heavy oil-fired boiler furnace. Fig. 5 shows the schematic diagram of the furnace and physical implementation of the flame imaging system.

The furnace is equipped with a rotary-cup atomization burner in a horizontal cylindrical combustion chamber with 6 m in length and 1.8 m in diameter. The refractory brick is installed in the furnace to prevent high temperature and heat loss. The 380# heavy oil [44] is preheated to



(a) Schematic diagram of the experimental setup.

(b) The physical implementation of the flame imaging system.

Fig. 5. Experimental apparatus of the heavy oil-fired boiler furnace.

~ 85 °C by electric heating and then injected into the combustion chamber through an oil gun. After being accelerated by the high-speed rotating cup, the heated heavy oil collides with the reverse primary air to achieve rapid atomization. The swirled secondary and tertiary air provide sufficient oxidizer to ensure combustion completion. An airflow control system is used to regulate the air supply by controlling valves.

The flame imaging system is consisting of an optical probe and a color camera (HIKVISION MV-CA003-50GC) with a resolution of 640 × 480 pixels at 150f/s (frames per second). To prevent the images from being too dark and saturated, the exposure time is set to 120 μs. The optical probe is protected by an air–water cooling jacket to ensure the temperature of the probe does not exceed 50 °C. As shown in Fig. 5 (a), "Air in" is the cooling air inlet, and "Air out" is the cooling air outlet. The pressure of the cooling air is always higher than that of the furnace so that the cooling air can flow into the furnace through the optical probe. The cooling air design not only effectively reduces the temperature of the probe but also ensures to keep the window clean for a long time running. "Water in" is the cooling water inlet, and "Water out" is the cooling water outlet. The cooling water design can further improve the cooling effect of the flame imaging system. The physical implementation of the flame imaging system is shown in Fig. 5 (b).

The optical probe is mounted on the viewport at the front of the furnace section, equipped with a 90° angle of objective view lens and visualize the root region of the flame. The root region of the flame is regarded as the primary reaction zone of the combustion process in terms of energy conversion and pollutant emission formation [45].

3.2. Dataset preparation

A wide range of flame images was recorded under different combustion states along with varying fuel loads (FL) and air supply. Table 2 depicts an overview of the flame image dataset obtained from nine different combustion states. Dataset *A* contains three different combustion states under different FLs with a fixed primary air (PA) and secondary air (SA). In dataset *B,* only PA was varied, and SA was varied in dataset *C*. The tertiary air is always kept at a 20% opening for all the experiments. For each combustion state, 4000 RGB (Red, Green, and Blue) flame images are collected. Example of flame images under nine combustion states is shown in Fig. 6.

Table 2
Overview of the dataset.

Dataset	States	FL (%)	PA (%)	SA (%)	Total images
A	S1	20	20	50	4000
	S2	30	20	50	4000
	S3	40	20	50	4000
B	S4	60	20	50	4000
	S5	60	35	50	4000
	S6	60	50	50	4000
C	S7	40	20	15	4000
	S8	40	20	35	4000
	S9	40	20	65	4000

4. Model establishment

4.1. Data preprocessing

To eliminate the influence of different image sizes and accelerate the convergence speed of the semi-supervised learning model, the flame images are resized to 256 (H) × 256 (V) and normalized to 0 to 1 by using the min–max scale [46]. Fig. 7 shows the overall structure of the dataset used for training and validation the semi-supervised learning model. To verify the generalization ability of the DAE-GAN, only unlabeled images of dataset *A* are used to train DAE-GAN. As a result, the dataset *A* is regarded as the original state (seen) and the dataset *B* and *C* are treated as the new states (unseen). In the first part, 80% of data from dataset *A* is randomly selected to form training dataset *A1,* and the remaining 20% data is used for validation (dataset *A2*). 94% of data (*A3*) from dataset *A1* is used to train the DAE-GAN, and the rest of the 6% (dataset *A4*) is used for the GPC training.

The second part is mainly the prediction of the unseen combustion state. In this part, the trained DAE-GAN is directly used without further training, while the trained GPC is further fine-tuned. The structure of dataset *B* and *C* is like dataset *A*. A small portion of data, i.e., 6% of data (*B4* and *C4*) from datasets *B1* and *C1* are used respectively for fine-tuning the GPC.

4.2. Training process

In the unsupervised training process, all weights of the DAE-GAN are initialized by Gaussian distribution with a standard deviation (STD) of 0.02. The DAE-GAN is trained using the dataset *A3*. To strengthen the robustness of the model, the dataset *A3* is corrupted by the *corruption ratio* $\delta = 0.2$. The *corruption ratio* is determined via cross-validation in comparison with other *ratios* such as 0.1, 0.3, 0.4 and 0.5 [refer to Fig. 11]. The parameters of the DAE-GAN are then updated via back-propagation using stochastic gradient descent method [47]. Fig. 8 depicts the training and validation progress of the DAE-GAN which performed through the proposed loss function $L_{DAE-GAN}$ [refer to Eq. (6)] with a learning rate of 0.01 and a batch size of 100.

It can be seen that the training time is increased linearly with the number of iterations, which takes 10,400 s under 300th iteration. It can also be seen that the training and validation losses are decreased rapidly in the first 90th iteration and then converged gradually. As the iteration

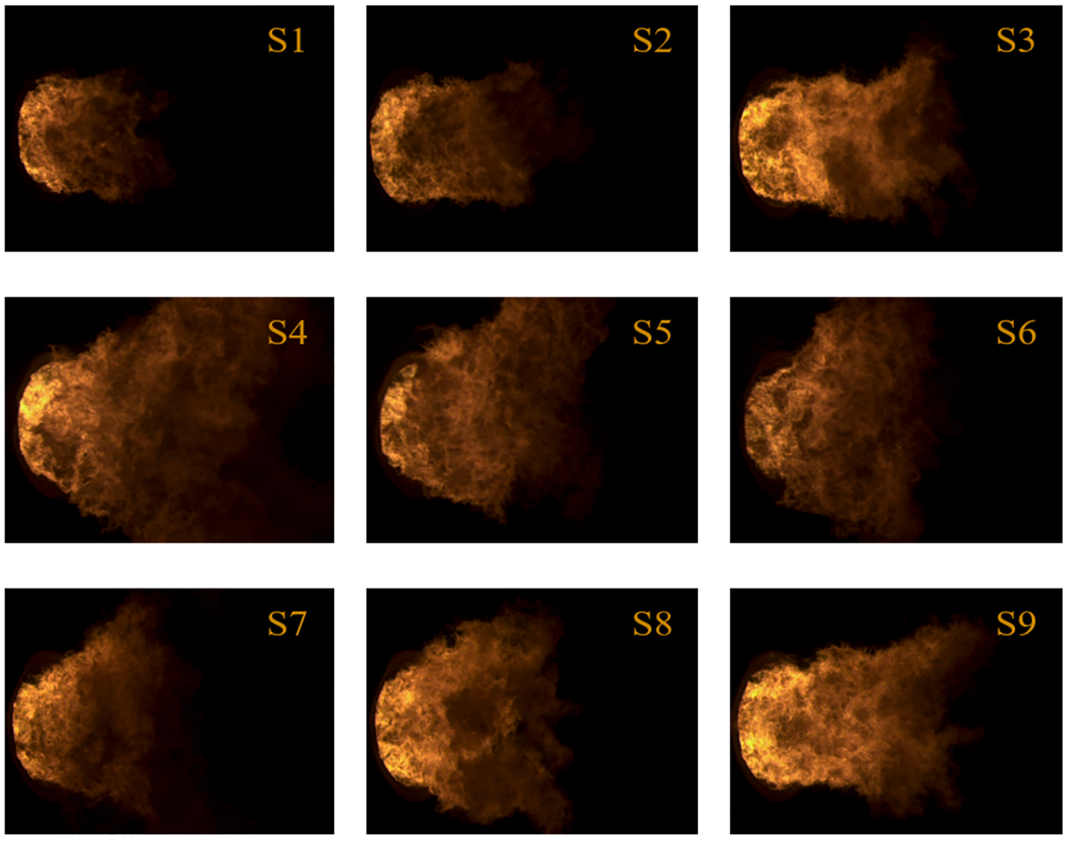


Fig. 6. Example of flame images under nine different combustion states.

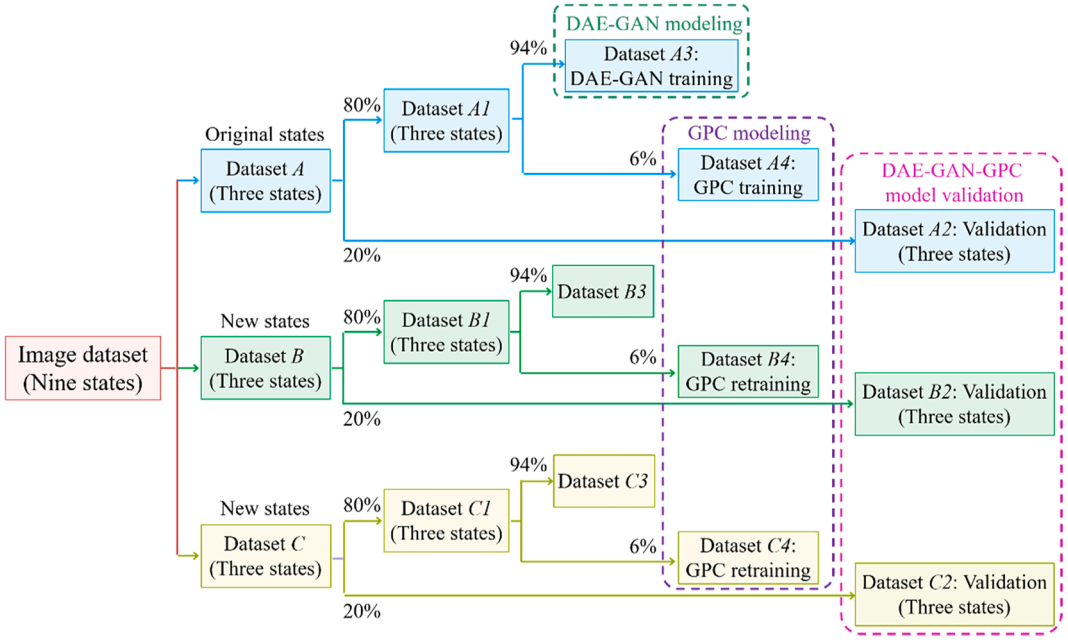


Fig. 7. The overall structure of the flame image dataset used for training and validation the semi-supervised learning model.

reached 180th, the validation loss increased slightly which eventually degraded the performance of the DAE-GAN. It indicates that excessive training is not always the best choice where it may encounter over-fitting. Therefore, the optimum iteration of the DAE-GAN is set to 150 by considering the cost of computation time and to prevent over-fitting.

In this study, all the computations were implemented in Anaconda software with Python programming language, and run on a processor (computing system) with an Intel i9-9900K CPU, 64 GB RAM and GeForce RTX 2080 GPU.

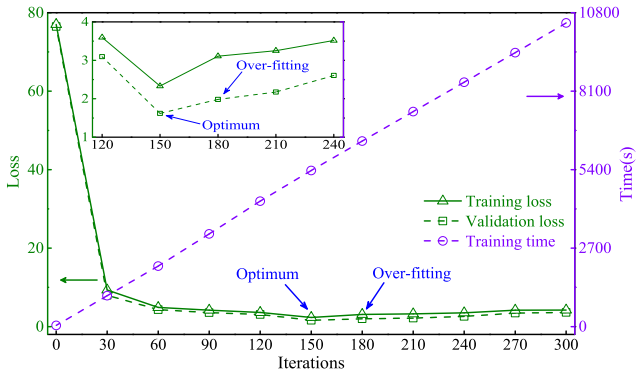


Fig. 8. Training and validation progress of the DAE-GAN with different iterations.

4.3. Performance analysis

To investigate the prediction accuracy and robustness of the semi-supervised learning model, the feature extraction network DAE-GAN is evaluated under different training and validation datasets, loss functions and corrupted ratios and classifiers. The detailed description of each evaluation and their effects on the prediction accuracy of the proposed model are illustrated in the following sections:

4.3.1. Performance analysis of semi-supervised learning model under different training datasets

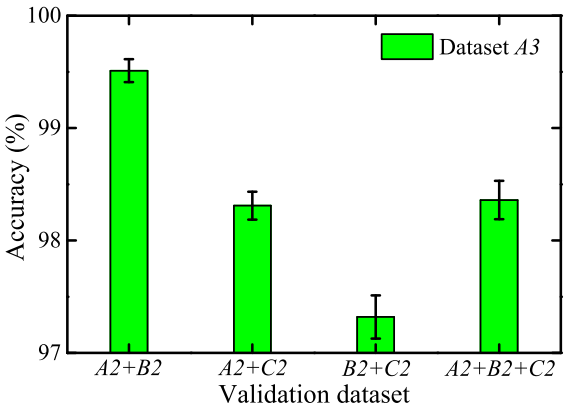
It is worth to investigate the effect of different training datasets on the feature extraction network DAE-GAN. The feature extraction network DAE-GAN was trained only using the dataset *A3* as shown in Fig. 7, whereas datasets *B3* and *C3* were not considered. Therefore, the prediction accuracy of the semi-supervised learning model is investigated by training the DAE-GAN through the datasets *A3*, *B3* and *C3* successively. After training the feature extraction network DAE-GAN, different combinations of datasets such as $A2 + B2$, $A2 + C2$, $B2 + C2$ and $A2 + B2 + C2$ were considered for testing the performance of the semi-supervised learning model. Fig. 9 shows the average prediction accuracy obtained from ten trials under different training datasets and their STDs.

It has been found that no matter which training dataset is used to train the DAE-GAN, the testing accuracy almost 98% and stable for dataset $A2 + B2 + C2$. The maximum STD is 0.23, further indicates that reliability of the prediction accuracy under different test methods. Thus, it is proved that the proposed feature extraction network DAE-GAN has an excellent generalization capability. It is also ensured that the high-precision prediction accuracy of the combustion state can be achieved through the semi-supervised learning model. Besides, by comparing the overall prediction accuracy of $A2 + B2$, $A2 + C2$ and $B2 + C2$, it can be seen that the testing accuracy is reduced to some extent for the dataset $B2 + C2$, in all cases. It is mainly due to the influence of dataset *C*. It has been identified that the flame images in dataset *C* are difficult to differentiate in structure due to a small-scale change in secondary air under a higher fuel load, i.e., 40%.

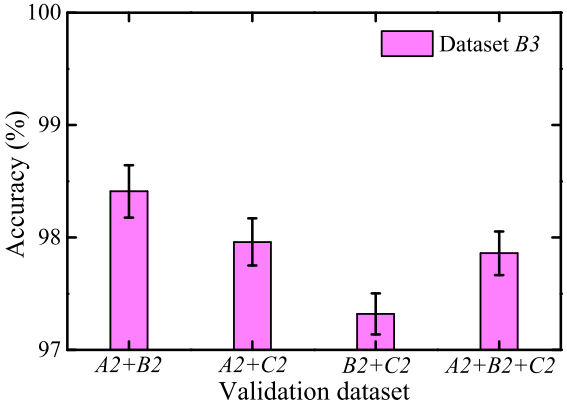
4.3.2. Effect of the proposed novel loss function

The SSIM metric is commonly used to evaluate the image quality but rarely participates in the training process of deep neural network. Therefore, it is an attempt to study whether the loss function with L_{SSIM} proposed in Eq. (6) has a positive impact on the feature extraction network DAE-GAN as well as on the prediction accuracy of the semi-supervised learning model. A comparative study is carried out with the proposed loss function and the loss function without L_{SSIM}. The loss function without L_{SSIM}can be defined as:

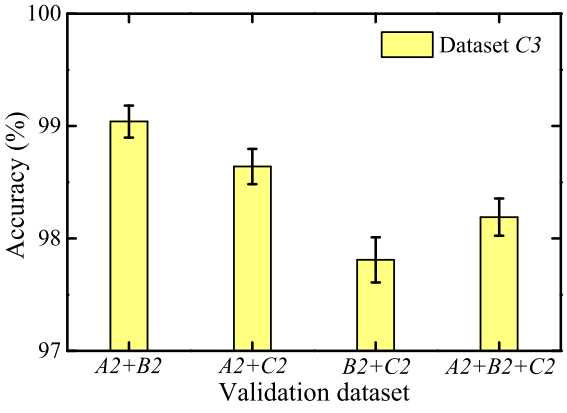
$$L_{NON-SSIM} = L_{GAN}(D, G) + L_{MSE} \tag{23}$$



(a) Accuracy under different validation datasets when dataset *A3* is used to train the DAE-GAN.



(b) Accuracy under different validation datasets when dataset *B3* is used to train the DAE-GAN.



(c) Accuracy under different validation datasets when dataset *C3* is used to train the DAE-GAN.

Fig. 9. Average prediction accuracy and standard deviation of the semi-supervised learning model under different training and validation datasets.

Fig. 10 illustrates the testing accuracy and STD of the semi-supervised learning model under different iterations and loss functions (i.e., $L_{DAE-GAN}$ and $L_{NON-SSIM}$) of DAE-GAN. It can be seen that the fluctuation of all STDs is relatively low, indicating the reliability of the model performance can be obtained. Under the effect of the $L_{DAE-GAN}$, the testing accuracy is initially increased and then decreased with the increasing of iterations. This phenomenon is mainly due to the successive occurrence of under-fitting and over-fitting. However, under the effect of the $L_{NON-SSIM}$, the testing accuracy is relatively low compared to the $L_{DAE-GAN}$. The testing accuracy of dataset $A2 + B2 + C2$ under the

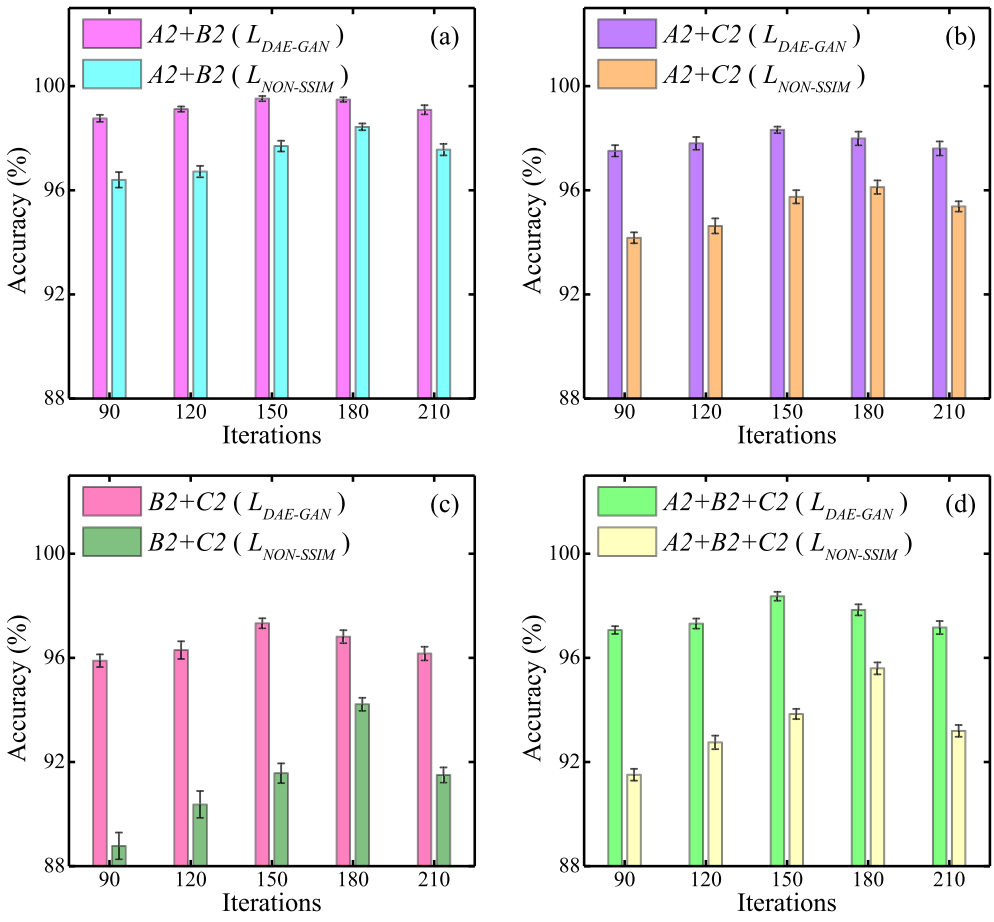


Fig. 10. Accuracy and standard deviation of the semi-supervised learning model under different loss functions and iterations.

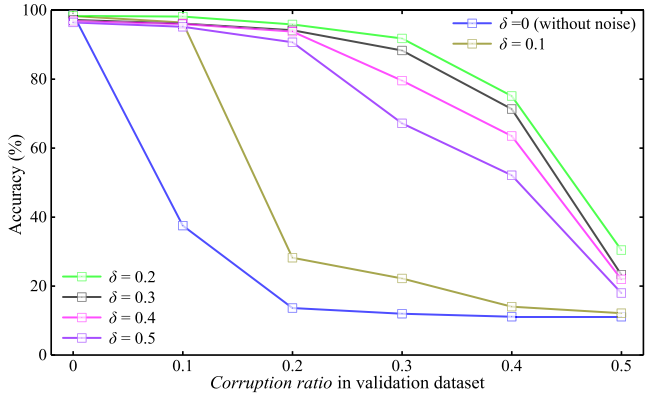


Fig. 11. Accuracy of the semi-supervised learning model under different *corruption ratios* in the training and validation datasets.

effect of the $L_{DAE-GAN}$ reaches its maximum of 98% at 150th iteration, while the testing accuracy under the effect of the $L_{NON-SSIM}$ is a maximum of 95% at 180th iteration. Even if the iteration is increased to 210th, there is no significant increment of the testing accuracy. Therefore, it is suggested that the loss function L_{SSIM} can improve the feature extraction efficiency through the DAE-GAN and thus improve the prediction accuracy of the combustion state with fewer iterations. Finally, the optimum iteration is set to 150, which helps to save computational cost and restrict the degradation of prediction performance.

4.3.3. Effect of the corruption ratio in training and validation dataset

In the heavy oil combustion, oil mist or smoke can easily be produced and affects the flame image quality severely. Therefore, it is necessary to develop a robust prediction model that can predict accurately even with poor quality images. To investigate the robustness of the semi-supervised learning model, Gaussian noise with different *corruption ratios* δ [determined through the Eq. (1)] is added to the training dataset *A3*. The *corruption ratio* δ in the training dataset is changed from 0 to 0.5 with a step size of 0.1. The DAE-GAN is then trained using the corrupted dataset *A3*. Similarly, Gaussian noise with different *corruption ratios* is also added to the validation datasets *A2*, *B2* and *C2* to investigate the robust performance of the semi-supervised learning model. The *corruption ratio* in the validation dataset is also changed from 0 to 0.5 with a step of 0.1. In each testing phase, ten trials were carried out and their averaged accuracy is shown in Fig. 11.

It can be seen that the prediction accuracy decreases gradually with the increase of *corruption ratios*. The highest prediction accuracy of the semi-supervised learning model is achieved by 99% (blue line in the figure) when $\delta = 0$. Whereas 98% (green line in the figure) is achieved with $\delta = 0.2$. The results also illustrate that even the accuracy drops rapidly with the increase of the *corruption ratio* in the validation dataset, the $\delta = 0.2$ provides still the highest accuracy. The prediction accuracy can remain above 91% even with the *corruption ratio* in the validation dataset of 0.3. Therefore, it is suggested that a better prediction accuracy and robustness capability can be achieved through the semi-supervised learning model.

4.3.4. Effect of different feature extraction networks

A comparative study is carried out with other deep neural networks such as DAE network and AE-GAN to verify the feature extraction performance of the DAE-GAN. The Gaussian noise with a *corruption ratio* of 0.2 is added in the validation datasets *A2*, *B2* and *C2* to investigate the

performance of these networks. The structure of the DAE network is the same as the generator of DAE-GAN. Whereas, the structure of the AE-GAN is also like the DAE-GAN, but stochastic mapping operation is not considered in this network [i.e., $\delta = 0$, refer to Eq. (1)]. Table 3 illustrates the *F_1-score* of the different deep neural networks. It can be seen the *F_1-score* of the DAE-GAN is higher than the DAE and AE-GAN in all combustion states. The DAE-GAN is better than the AE-GAN and indicates that appropriate denoising coding can improve the robustness of the semi-supervised learning model. The DAE-GAN also shows a better performance than the DAE network and demonstrates that the adversarial learning mechanism of the GAN further enhances the feature learning capacity of the semi-supervised learning model.

4.3.5. *Effect of different classifiers and proportions of labeled data*

Although the unsupervised network can extract representative features from unlabeled images, a certain amount of labeled data is essential to establish and validate the prediction model. It is known that neural network classifiers demand different proportions of the labeled data due to their different architectures. Thus, careful consideration is required to select the portion of the labeled data. It is therefore desirable to adopt an optimal classifier which can provide better performance with a minimum labeled data. In this study, the performance of the proposed GPC is compared with Random Forest (RF), Logistic regression (LR), ANN, Linear SVM (LSVM), and Kernel SVM (KSVM). Different proportions of the labeled image from datasets *A4*, *B4* and *C4* are used to investigate the performance of these classifiers, and the proportion of labeled data is varied from 1% to 8% with a step size of 1%.

Fig. 12 shows the prediction accuracy of the various classifiers under different proportions of labeled data. Although the hyper-parameters of these classifiers have been optimized, such as RF changes the number of decision trees [48], LR selects different penalty terms [49], ANN chooses a different number of hidden layers and neurons [50], LSVM and KSVM adjust the punishment coefficient [51], the GPC can provide better performance in all cases without further modification. It is also noted that the prediction accuracy is also improved rapidly as the proportions of labeled data increases from 1% to 6%. For the GPC, the testing accuracy reaches 98% even with 4% of labeled data and shows great potential to meets the requirements where the labeled data is limited.

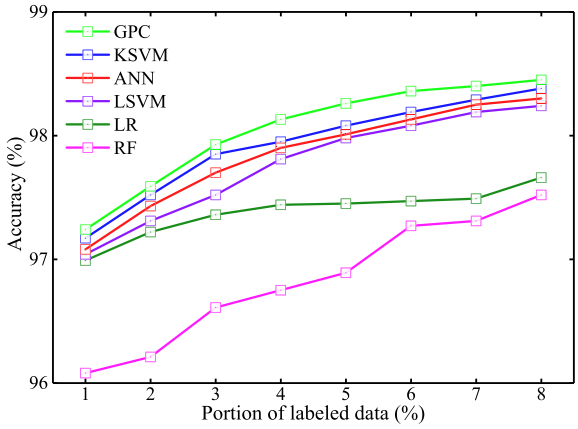


Fig. 12. Accuracy of different classifiers under different portions of labeled images.

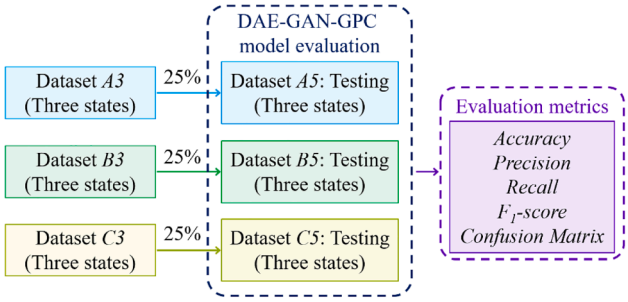


Fig. 13. Structure of the testing dataset for the semi-supervised learning model.

5. Model evaluation and discussions

5.1. *Combustion state prediction*

To evaluate the performance of established semi-supervised learning model, 25% from datasets *A3*, *B3* and *C3* are randomly chosen to form the testing datasets *A5*, *B5* and *C5*, as shown in Fig. 13. All the evaluation trials were repeated ten times to avoid the particularity and contingency of a single result, and the average testing results are summarized in Table 4.

It is evident that the testing accuracy of dataset *A5* reaches 99.83% with the *F_1-score* of 0.99 and indicates that the semi-supervised learning model exhibits a better prediction performance for different FLs. The testing accuracy of dataset *B5* and *C5* is achieved above 96.87% and suggesting that the better prediction can be achieved for the new combustion states by simply fine-tuning the GPC without retraining the DAE-GAN. For the combined dataset $A5 + B5 + C5$, the prediction accuracy reaches 98.36% and demonstrates the effectiveness of the proposed semi-supervised learning model for combustion state prediction with a large number of unlabeled and a small number of labeled data. The results also demonstrate that the DAE-GAN provides a better generalization capability. Even if the new combustion states are not used to train the DAE-GAN, it can still extract essential features accurately and provide acceptable prediction accuracy.

Table 4
Testing accuracy for different combustion states.

Testing dataset	Accuracy	*Precision*	*Recall*	*F_1-score*
A5	99.83%	0.99	0.99	0.99
B5	99.08%	0.99	0.99	0.99
C5	96.87%	0.97	0.97	0.97
$A5 + B5 + C$	598.36%	0.98	0.98	0.98

5.2. *Model examination for all combustion states*

To investigate the prediction accuracy in detail, the confusion matrix of the nine combustion states is calculated. Fig. 14 illustrates the confusion matrix of the datasets *A5*, *B5* and *C5* obtained through the semi-supervised learning model. In this figure, rows correspond to the predicted category, and columns represent the actual category. The

Table 3
F_1-score of different feature extraction networks.

Feature extraction network	dataset *A2*			dataset *B2*			dataset *C2*		
	S1	S2	S3	S4	S5	S6	S7	S8	S9
AE-GAN	0.10	0.10	0.34	0.10	0.22	0.10	0.13	0.21	0.10
DAE	0.96	0.95	0.96	0.90	0.93	0.93	0.88	0.85	0.95
DAE-GAN	**0.98**	**0.97**	**0.98**	**0.95**	**0.97**	**0.96**	**0.93**	**0.91**	**0.97**

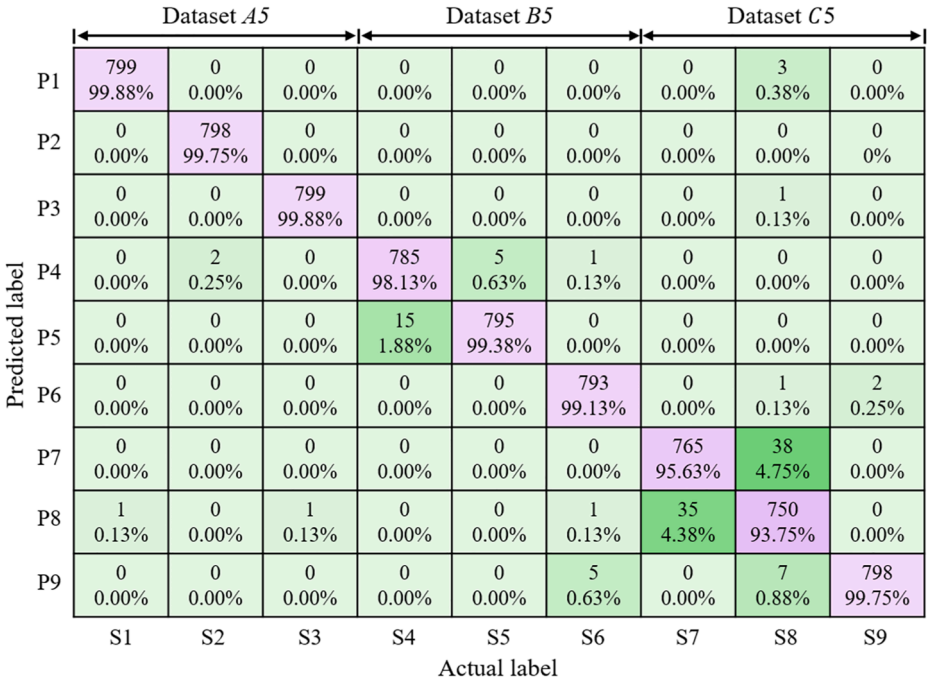


Fig. 14. The confusion matrix of nine combustion states.

diagonal cells show the number of samples which are correctly estimated, and the off-diagonal cells display where the classifier has made mistakes.

It can be seen that the test samples in dataset *A5* are classified accurately. Larger misclassification can be seen for the combustion states 7 and 8 compared to the other combustion states. Among them, the prediction accuracy of state 8 is the worst with 93.75% that only 750 samples are correctly classified out of 800 samples, and the rest of them are erroneously divided into state 7. This is because of almost similar flame structures produced by a small-scale change in secondary air (such as 15% and 35%) under a higher fuel load, i.e., 40%. Although there are inevitably misclassifications, most samples can be predicted accurately with a success rate of 98.36%. Overall, the proposed semi-supervised learning model provides the acceptable accuracy and reliable prediction of combustion states.

5.3. Feature distribution visualization of flame images

To represent the discriminative features of the nine combustion states, the image features extracted by the DAE-GAN are visualized through the t-Distributed Stochastic Neighbor Embedding (t-SNE) technique [52]. The t-SNE method is an effective data visualization technique for high-dimensional data representation. This technique can convert the 16-dimensional features to a two-dimensional (2-D) map with a maximally retained data structure. The 2-D feature map of the testing datasets *A5*, *B5* and *C5* are illustrated in Fig. 15.

The result demonstrates that the image features extracted from the same states are clustered, whilst image features under different states are separated well. It can be seen that although flame images from different combustion states are indistinguishable, their extracted deep features are divisible. In addition, there are some mixing points in the feature space, which are mainly concentrated in combustion states 4 and 5, 7 and 8. This result is consistent with that given by the confusion matrix in Fig. 14, whereas misclassification is prone to occur under these combustion states. It can be concluded that if the images from different combustion states cannot be separated in the feature space, subsequently, the established classifier is unable to classify correctly, which will inevitably reduce the prediction accuracy. Therefore, it is suggested

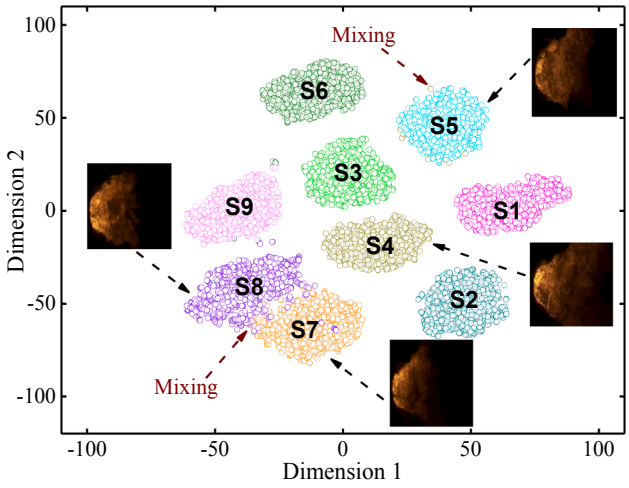


Fig. 15. Visualization of flame image features under nine different combustion states.

that the representative features of flame image play a decisive role for reliable and accurate prediction of the combustion state.

6. Conclusions

This paper reported a novel semi-supervised learning model based on denoising autoencoder, generative adversarial network and Gaussian process classifier for combustion state prediction. The effectiveness of the model is evaluated through 4.2 MW heavy oil-fired boiler furnace flames captured under different combustion states. The deep features of the flame are extracted automatically by using the denoising autoencoder improved through the generative adversarial network. The extracted image features are then used to predict the combustion state through the Gaussian process classifier. The proposed method can predict new combustion states (unseen) by simply fine-tuning the Gaussian process classifier with a small portion of labeled images and thus save the computational cost significantly. The semi-supervised learning

model is validated, and generalization and robustness capability are also verified through the new combustion states. The main outcomes drawn from this study are summarized as follows:

(a) The denoising autoencoder-generative adversarial network can extract the representative features from the unlabeled flame images automatically with an improved anti-noise ability. The proposed model overcomes the shortcomings of the traditional techniques, such as dependence on prior expert knowledge, large demand for labeled data, and complicated parameter optimization.
(b) To establish a high-precision prediction model, this study conducts a detailed exploration including a novel loss function, reasonable training iteration, and optimum *corruption ratio*. The proposed loss function of the deep neural network greatly improves the training efficiency and prediction accuracy.
(c) The proposed semi-supervised learning model provides better prediction accuracy compared with different feature learning methods and classifiers. The testing accuracy is achieved around 98.36% for heavily oil-fired combustion states, indicates that the model is feasible for industrial combustion applications.
(d) The feature extraction network based on denoising autoencoder-generative adversarial network shows strong robustness and generalization capabilities for new combustion states, and it can easily be fine-tuned to predict the combustion stability and exhaust emissions.
(e) Overall, the proposed semi-supervised learning model shows a promising tool for the combustion state prediction, and it can be suitable for other combustion processes such as coal combustion, biomass co-combustion and even in the combustion engine without significant changes.

The future work will be focused on tailoring the proposed model for predicting the oxygen content and NOx emission of different combustion applications.

CRediT authorship contribution statement

Zhezhe Han: Conceptualization, Methodology, Software, Data curation, Validation, Writing - original draft. **Jian Li:** Validation, Formal analysis, Investigation. **Biao Zhang:** Software, Visualization, Data curation. **Md. Moinul Hossain:** Supervision, Validation, Formal analysis, Writing - review & editing. **Chuanlong Xu:** Supervision, Resources, Writing - review & editing, Project administration.

Declaration of Competing Interest

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

Acknowledgements

This work was supported by the National Natural Science Foundation of China [grant number 51976038]; the Natural Science Foundation of Jiangsu Province, China for Young Scholars [grant number BK20190366]; and the Fundamental Research Funds for the Central Universities [grant number 2242019k1G018].

References

[1] Bai X, Lu G, Hossain MM, Szuhánszki J, Daood SS, Nimmo W, Yan Y, Pourkashanian M. Multi-mode combustion process monitoring on a pulverised fuel combustion test facility based on flame imaging and random weight network techniques. Fuel 2017;202:656–64. https://doi.org/10.1016/j.fuel.2017.03.091.
[2] Ballester J, García-Armingol T. Diagnostic techniques for the monitoring and control of practical flames. Prog Energy Combust Sci 2010;36(4):375–411. https://doi.org/10.1016/j.pecs.2009.11.005.
[3] Zhou H, Tang Qi, Yang L, Yan Y, Lu G, Cen K. Support vector machine based online coal identification through advanced flame monitoring. Fuel 2014;117:944–51. https://doi.org/10.1016/j.fuel.2013.10.041.
[4] Toth P, Garami A, Csordas B. Image-based deep neural network prediction of the heat output of a step-grate biomass boiler. Appl Energy 2017;200(15):155–69. https://doi.org/10.1016/j.apenergy.2017.05.080.
[5] González-Cencerrado A, Gil A, Peña B. Characterization of PF flames under different swirl conditions based on visualization systems. Fuel 2013;113:798–809. https://doi.org/10.1016/j.fuel.2013.05.077.
[6] Hernández R, Ballester J. Flame imaging as a diagnostic tool for industrial combustion. Combust Flame 2008;155(3):509–28. https://doi.org/10.1016/j.combustflame.2008.06.010.
[7] Yan Y, Lu G, Colechin M. Monitoring and characterisation of pulverised coal flames using digital imaging techniques. Fuel 2002;81(5):647–55. https://doi.org/10.1016/S0016-2361(01)00161-2.
[8] González-Cencerrado A, Peña B, Gil A. Coal flame characterization by means of digital image processing in a semi-industrial scale PF swirl burner. Appl Energy 2012;94:375–84. https://doi.org/10.1016/j.apenergy.2012.01.059.
[9] Li W, Wang D, Chai T. Flame image-based burning state recognition for sintering process of rotary kiln using heterogeneous features and fuzzy integral. IEEE Trans. Ind. Inf. 2012;8(4):780–90. https://doi.org/10.1109/TII.2012.2189224.
[10] Chen J, Chan LLT, Cheng Y-C. Gaussian process regression based optimal design of combustion systems using flame images. Appl Energy 2013;111:153–60. https://doi.org/10.1016/j.apenergy.2013.04.036.
[11] Chen J, Hsu T-Y, Chen C-C, Cheng Y-C. Monitoring combustion systems using HMM probabilistic reasoning in dynamic flame images. Appl Energy 2010;87(7):2169–79. https://doi.org/10.1016/j.apenergy.2009.11.008.
[12] Sun D, Lu G, Zhou H, Yan Y. Condition monitoring of combustion processes through flame imaging and kernel principal component analysis. Combust Sci Technol 2013;185(9):1400–13. https://doi.org/10.1080/00102202.2013.798316.
[13] Hinton G, Salakhutdinov R. Reducing the dimensionality of data with neural networks. Science 2006;313(5786):504–7. https://doi.org/10.1126/science.1127647.
[14] Gu J, Wang Z, Kuen J, Ma L, Shahroudy A, Shuai B, Liu T, Wang X, Wang G, Cai J, Chen T. Recent advances in convolutional neural networks. Pattern Recogn 2018;77:354–77. https://doi.org/10.1016/j.patcog.2017.10.013.
[15] Wang Z, Song C, Chen T. Deep learning based monitoring of furnace combustion state and measurement of heat release rate. Energy 2017;131:106–12. https://doi.org/10.1016/j.energy.2017.05.012.
[16] Abdurakipov SS, Gobyzov OA, Tokarev MP, Dulin VM. Combustion regime monitoring by flame imaging and machine learning. Optoelectron Instrument Proc 2018;54(5):513–9. https://doi.org/10.3103/S875669901805014X.
[17] Wu Y, Misra S, Sondergeld C, Curtis M, Jernigen J. Machine learning for locating organic matter and pores in scanning electron microscopy images of organic-rich shales. Fuel 2019;253:662–76. https://doi.org/10.1016/j.fuel.2019.05.017.
[18] Liu Yi, Fan Yu, Chen J. Flame images for oxygen content prediction of combustion systems using DBN. Energy Fuels 2017;31(8):8776–83. https://doi.org/10.1021/acs.energyfuels.7b00576.
[19] Lyu Y, Chen J, Song Z. Image-based process monitoring using deep learning framework. Chemom Intell Lab Syst 2019;189:8–17. https://doi.org/10.1016/j.chemolab.2019.03.008.
[20] Qiu T, Liu M, Zhou G, Wang L, Gao K. An unsupervised classification method for flame image of pulverized coal combustion based on convolutional auto-encoder and hidden Markov model. Energies 2019;12(13):2585. https://doi.org/10.3390/en12132585.
[21] Akintayo A, Lore K, Sarkar S, Sarkar S. Early detection of combustion instabilities using deep convolutional selective autoencoders on hi-speed flame video. Int J Prognostics Health Monitor 2016. https://doi.org/10.1145/1235.
[22] Han Z, Hossain MM, Wang Y, Li J, Xu C. Combustion stability monitoring through flame imaging and stacked sparse autoencoder based deep neural network. Appl Energy 2020;259:114159. https://doi.org/10.1016/j.apenergy.2019.114159.
[23] Zhu X, Cai Z, Wu J, Cheng Y, Huang Q. Convolutional neural network based combustion mode classification for condition monitoring in the supersonic combustor. Acta Astronaut 2019;159:349–57. https://doi.org/10.1016/j.actaastro.2019.03.072.
[24] Liukkonen M, Heikkinen M, Hiltunen T, Hälikkä E, Kuivalainen R, Hiltunen Y. Artificial neural networks for analysis of process states in fluidized bed combustion. Energy 2011;36(1):339–47. https://doi.org/10.1016/j.energy.2010.10.033.
[25] Wang Yu, Yu Y, Zhu X, Zhang Z. Pattern recognition for measuring the flame stability of gas-fired combustion based on the image processing technology. Fuel 2020;270:117486. https://doi.org/10.1016/j.fuel.2020.117486.
[26] Zhai Y, Ding X, Jin X, Zhao L. Adaptive LSSVM based iterative prediction method for NOx concentration prediction in coal-fired power plant considering system delay. Appl Soft Comput 2020;89:106070. https://doi.org/10.1016/j.asoc.2020.106070.
[27] Sun W, Shao S, Zhao R, Yan R, Zhang X, Chen X. A sparse auto-encoder-based deep neural network approach for induction motor faults classification. Measurement 2016;89:171–8. https://doi.org/10.1016/j.measurement.2016.04.007.
[28] Duan H, Huang Y, Mehra RK, Song P, Ma F. Study on influencing factors of prediction accuracy of support vector machine (SVM) model for NOx emission of a hydrogen enriched compressed natural gas engine. Fuel 2018;234:954–64. https://doi.org/10.1016/j.fuel.2018.07.009.

[29] Hao Z, Kefa C, Jianbo M. Combining neural network and genetic algorithms to optimize low NOx pulverized coal combustion. Fuel 2001;80(15):2163–9. https://doi.org/10.1016/S0016-2361(01)00104-1.
[30] Wang J, Zhai Z(, Jing Y, Zhang C. Particle swarm optimization for redundant building cooling heating and power system. Appl Energy 2010;87(12):3668–79. https://doi.org/10.1016/j.apenergy.2010.06.021.
[31] Rasmussen C, Williams C. Gaussian processes for machine learning. MIT Press 2005. https://doi.org/10.7551/mitpress/3206.001.0001.
[32] Ögren Y, Tóth P, Garami A, Sepman A, Wiinikka H. Development of a vision-based soft sensor for estimating equivalence ratio and major species concentration in entrained flow biomass gasification reactors. Appl Energy 2018;226:450–60. https://doi.org/10.1016/j.apenergy.2018.06.007.
[33] Vincent P, Larochelle H, Lajoie I, Bengio Y, Manzagol P. Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion. J Mach Learn Res 2010;11:3371–408. https://doi.org/10.1016/j.mechatronics.2010.09.004.
[34] Larochelle H, Bengio Y, Louradour J, Lamblin P. Exploring strategies for training deep neural networks. J Mach Learn Res 2009;1(10):1–40. https://doi.org/10.1145/1577069.1577070.
[35] Wang Z, Wang J, Wang Y. An intelligent diagnosis scheme based on generative adversarial learning deep neural networks and its application to planetary gearbox fault pattern recognition. Neurocomputing 2018;310:213–22. https://doi.org/10.1016/j.neucom.2018.05.024.
[36] Jiao J, Zhao M, Lin J, Zhao J. A multivariate encoder information based convolutional neural network for intelligent fault diagnosis of planetary gearboxes. Knowl-Based Syst 2018;160:237–50. https://doi.org/10.1016/j.knosys.2018.07.017.
[37] Han Y, Tang B, Deng L. Multi-level wavelet packet fusion in dynamic ensemble convolutional neural network for fault diagnosis. Measurement 2018;127:246–55. https://doi.org/10.1016/j.measurement.2018.05.098.
[38] Zhou D, Yao Q, Wu H, Ma S, Zhang H. Fault diagnosis of gas turbine based on partly interpretable convolutional neural networks. Energy 2020;200:117467. https://doi.org/10.1016/j.energy.2020.117467.
[39] Liu H, Zhou J, Xu Y, Zheng Y, Peng X, et al. Unsupervised fault diagnosis of rolling bearings using a deep neural network based on generative adversarial networks. Neurocomputing 2018;315:412–24. https://doi.org/10.1016/j.neucom.2018.07.034.
[40] He Yu, Song K, Dong H, Yan Y. Semi-supervised defect classification of steel surface based on multi-training and generative adversarial network. Opt Lasers Eng 2019;122:294–302. https://doi.org/10.1016/j.optlaseng.2019.06.020.
[41] Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. IEEE Trans. on Image Process. 2004;13(4):600–12. https://doi.org/10.1109/TIP.2003.819861.
[42] Lunderman S, Fioroni GM, McCormick RL, Nimlos MR, Rahimi MJ, Grout RW. Screening fuels for autoignition with small-volume experiments and gaussian process classification. Energy Fuels 2018;32(9):9581–91. https://doi.org/10.1021/acs.energyfuels.8b02112.
[43] Zhang C, Wei H, Zhao X, Liu T, Zhang K. A Gaussian process regression based hybrid approach for short-term wind speed prediction. Energy Convers Manage 2016;126:1084–92. https://doi.org/10.1016/j.enconman.2016.08.086.
[44] Zhou H, Ren T, Yang Y. Impact of OFA on combustion and NOx emissions of a large-scale laboratory furnace fired by a heavy-oil swirl burner. Appl Therm Eng 2015;90:994–1006. https://doi.org/10.1016/j.applthermaleng.2015.07.076.
[45] Lu G, Yan Y, Colechin M. A digital imaging based multifunctional flame monitoring system. IEEE Trans. Instrum. Meas. 2004;53(4):1152–8. https://doi.org/10.1109/TIM.2004.830571.
[46] Tang Z, Zhang Z. The multi-objective optimization of combustion system operations based on deep data-driven models. Energy 2019;182:37–47. https://doi.org/10.1016/j.energy.2019.06.051.
[47] Qian Qi, Jin R, Yi J, Zhang L, Zhu S. Efficient distance metric learning by adaptive sampling and mini-batch stochastic gradient descent (SGD). Mach Learn 2015;99(3):353–72. https://doi.org/10.1007/s10994-014-5456-x.
[48] Lei C, Deng J, Cao K, Ma Li, Xiao Y, Ren L. A random forest approach for predicting coal spontaneous combustion. Fuel 2018;223:63–73. https://doi.org/10.1016/j.fuel.2018.03.005.
[49] Bielza C, Robles V, Larrañaga P. Regularized logistic regression without a penalty term: An application to cancer classification with microarray data. Expert Syst Appl 2011;38(5):5110–8. https://doi.org/10.1016/j.eswa.2010.09.140.
[50] Adewole BZ, Abidakun OA, Asere AA. Artificial neural network prediction of exhaust emissions and flame temperature in LPG (liquefied petroleum gas) fueled low swirl burner. Energy 2013;61:606–11. https://doi.org/10.1016/j.energy.2013.08.027.
[51] Smrekar J, Potočnik P, Senegačnik A. Multi-step-ahead prediction of NOx emissions for a coal-based boiler. Appl Energy 2013;106:89–99. https://doi.org/10.1016/j.apenergy.2012.10.056.
[52] Laurens V, Hinton G. Visualizing data using t-SNE. J Mach Learn Res 2008;9(2605):2579–605. https://doi.org/10.1007/s10846-008-9235-4.