Deep learning-based fast time-resolved flame emission spectroscopy in high-pressure combustion environment

Taekeun Yoon\textsuperscript{a,b}, Seon Woong Kim\textsuperscript{a,b}, Hosung Byun\textsuperscript{a,b}, Younsik Kim\textsuperscript{c}, Campbell D. Carter\textsuperscript{d}, and Hyungrok Do\textsuperscript{a,b,*}

\textsuperscript{a} Department of Mechanical Engineering, Seoul National University, Seoul 08826, Republic of Korea
\textsuperscript{b} Institute of Advanced Machines and Design (IAMD), Seoul National University, Seoul 08826, Republic of Korea
\textsuperscript{c} Department of Physics and Astronomy, Seoul National University, Seoul 08826, Republic of Korea
\textsuperscript{d} Air Force Research Laboratory, Wright-Patterson Air Force Base, Ohio 45433, United States

* Corresponding author’s Email address: hyungrok@snu.ac.kr

Abstract

A novel deep learning strategy is developed for fast and accurate gas property measurements using flame emission spectroscopy (FES). Particularly, the short-gated fast FES is essential to resolve fast-evolving combustion behaviors. However, as the exposure time for capturing the flame emission spectrum gets shorter, the signal-to-noise ratio (SNR) decreases, and characteristic spectral features indicating the gas properties become relatively weaker. Then, the property estimation based on the short-gated spectrum is difficult and inaccurate. Denoising convolutional neural networks (CNN) can enhance the SNR of the short-gated spectrum. A new CNN architecture including a reversible down- and up-sampling (DU) operator and a loss function based on proper orthogonal decomposition (POD) coefficients is proposed. For training and testing the CNN, flame chemiluminescence spectra were captured from a stable methane-air flat flame using a portable spectrometer (spectral range: 250–850 nm, resolution: 0.5 nm) with varied equivalence ratio (0.8–1.2), pressure (1–10 bar), and exposure time (0.05, 0.2, 0.4, and 2 s). The long exposure (2 s) spectra were used as the ground truth when training the denoising CNN. A kriging model with POD is trained by the long-gated spectra for calibration and then prediction of the gas properties taking the denoised short-gated spectrum as the input. The measurement or property prediction errors of pressure and equivalence ratio using the new technique were estimated to be 5.7% and 1.5% with 0.2 s exposure, which are exceptionally good and typically not achievable with such low SNR spectrum signals without a signal amplifier.

Keywords: flame emission spectroscopy, chemiluminescence, deep learning, convolutional neural network
Nomenclature

Latin characters

\( I_{\text{dark}} \) dark current
\( L \) loss function
\( N \) number of data pairs
\( N_c \) number of channels
\( N_d \) number of sub-signals
\( N_{\text{electron}} \) number of counted electrons
\( N_k \) number of kernel size
\( N_l \) number of layers
\( N_p \) number of CNN parameter
\( N_R \) read noise
\( P \) pressure
\( T_0 \) cosine annealing cycle step size
\( T_{\text{magn}} \) cosine annealing cycle magnification
\( T_{\text{up}} \) cosine annealing linear warmup step size
\( W \) number of components in input signal
\( x_i \) short-gated spectrum
\( y_i \) long-gated spectrum
\( \hat{y}_i \) denoised short-gated spectrum

Greek symbols

\( \alpha \) constant defined in Eq. (1)
\( \gamma \) cosine annealing decrease rate of max learning rate
\( \eta \) quantum efficiency
\( \eta_{\text{max}} \) cosine annealing max learning rate
\( \sigma_{\text{photon}} \) photon noise
\( \sigma_{\text{dark}} \) dark noise
\( \sigma_{\text{Read}} \) readout noise
\( \tau \) exposure time
\( \phi \) equivalence ratio
\( \phi_p \) photon flux at the CCD

Abbreviations
1. Introduction

Recently developed fast time-resolved optical measurement methods have remarkably extended our understanding of turbulent combustion phenomena. The transient and fast-evolving combustion processes can only be investigated by non-intrusive optical diagnostics with high temporal resolution. For example, several kilohertz (kHz) sampling rates were required to temporally resolve the combustion dynamics in large-scale eddies accompanying micro-vortex structures [1-4]. However, capturing full temporal dynamics in such turbulent combustion environments is challenging even with the most advanced optical diagnostics tools. This is because various gas properties, e.g., temperature, pressure, and species concentrations, and local-bulk velocity profiles that characterize the combustion behaviors, fluctuate within extremely short characteristic time scales; typical integral and Kolmogorov time scales are tens and hundreds of kHz, respectively [5, 6].

Flame emission spectroscopy (FES) is the simplest optical combustion diagnostics method and therefore widely used to monitor the gas properties in the combustion zone. FES utilizes instantaneous combustion chemiluminescence spectra for estimating the gas properties, and the exposure time for capturing the chemiluminescence spectra determines the time resolution of FES. The typical sampling rate of recently developed photon detectors is up to 10 MHz with a complementary metal-oxide-semiconductor camera [7, 8], which implies that the temporal resolution of the emission spectroscopy can significantly be improved using faster and more sensitive detectors. Recall that emission spectroscopy including FES does not require the use of any light sources while laser-aided diagnoses are limited by the repetition rate of the laser systems triggering photon emission or scattering, e.g., laser-induced
fluorescence [9], coherent anti-Stokes Raman scattering [10, 11], and laser-induced breakdown spectroscopy [12].

Nevertheless, as the exposure time decreases, the signal-to-noise ratio (SNR) of the chemiluminescence spectra will decrease, with concomitant lowering of the accuracy of the FES measurements. In general, the SNR of a photon detector is proportional to the square root of the exposure time assuming negligible readout noise. Denoising filters such as Gaussian smoothing that selectively attenuate high-frequency noise [13] can enhance the SNR with reduced exposure. However, the property indicators such as sharp emission peaks and radical band structures in the spectrum will inevitably be blurred, resulting in the loss of critical high-frequency information.

The denoising ability of state-of-the-art machine learning technology for digital image processing is superior to the conventional methods including Gaussian smoothing and wavelet shrinkage [14, 15]. However, these machine learning-based denoising techniques have been mostly developed for qualitative image processing rather than quantitative analyses [16-18]. The introduction of this technique to spectroscopic analyses can make significant impacts on various quantitative optical measurement methods. Several recent studies have employed the technology for spectroscopic analyses in Raman spectroscopy [19], laser absorption spectroscopy [20], electron spectroscopy [21], and angle-resolved photoemission spectroscopy [22], and it was confirmed that this technique can potentially improve the accuracy of any quantitative measurements utilizing spectroscopic analyses.

In this work, we employed a novel data processing architecture utilizing a convolutional neural network (CNN) denoising technique with a new loss function to improve the temporal resolution and accuracy of FES. The proposed method removes noise from the short-gated chemiluminescence spectra with minimal information loss. In addition, a proper orthogonal decomposition (POD) method with a kriging model as a data-driven calibration technique is used for mapping the processed spectra and gas properties. It was confirmed that the proposed method can accurately predict gas properties in broad test ranges, equivalence ratio (0.8 – 1.2) and pressure (1 – 10 bar) of methane-air flame, taking a short-gated flame emission spectrum of low SNR as the input, which enables FES of high accuracy and high temporal resolution.

2. Methodology

2.1. Overall calibration and prediction process

The proposed calibration process and the test procedure of the trained model are summarized in Fig. 1. The calibration process consists of three parts: 1) Acquire training data, 2) Map data, and 3) Train CNN. Then, the prediction procedure is for measurements in arbitrary conditions: 4) Acquire high temporal resolution FES. In the following paragraph, detailed descriptions of the four parts are given.

1) Pairs of high-SNR (long-gated and averaged) and low-SNR (short-gated and instantaneous) spectra are collected in steady-state calibration experiments to train and validate the model. 2) The high-SNR spectra are used to construct
a surrogate model mapping the high-SNR spectrum data to the target gas properties of interest, e.g., pressure (P) and equivalence ratio (ϕ). Any data mapping techniques, e.g., the simplest one-to-one correlation model of emission band intensity ratios [23], partial least square regression [24], and artificial neural network [25], can be used with the novel denoising technique. In this paper, a reduced-order model, POD + kriging is adopted because of its high accuracy and precision [26]. Property-sensitive broadband spectrum features (POD bases) are calculated by the POD that decomposes the high-SNR spectrum data, and the reduced-order model (kriging) trained by the database can accurately predict multiple gas properties taking a high-SNR spectrum as the input. More details regarding the data-mapping technique are in our previous report [26]. 3) Then, the denoising CNN is trained by the high-SNR (label) and low-SNR (input) spectrum pairs to provide a high-SNR spectrum derived from a short-gated spectrum of low-SNR (input). A separate spectrum dataset (validation data) is used for checking the under-fitting or over-fitting of the calibration model. Finally, 4) arbitrary short-gated chemiluminescence spectrum data are collected, and high temporally resolved gas properties data are acquired.

![Figure 1](image-url)  
**Figure 1** The sequence of (a) the calibration process and (b) the prediction procedure of the trained model.

2.2. High-Pressure Flat Flame Burner and Spectrometer
Figure 2 presents a typical flame chemiluminescence image taken on the high-pressure McKenna burner used for collecting the flame emission spectra with varied fuel concentrations and ambient pressure. The flat flame burner was installed in a high-pressure chamber that regulated the chamber pressure stably with electrically controlled mass flow controllers and a choking nozzle at the chamber exit. Fully premixed methane and air flow upward above the burner exit plane surrounded by an air co-flow. The mass flows of the gases are remotely controlled by thermal mass flow controllers (Bronkhorst EL-FLOW, F-211AV for methane, and F-211AC & F-002AV for air). The flow speed matches the laminar flame speed calculated using ANSYS Chemkin with GRI-Mech 3.0 under given gas property conditions. Sufficient diluent air, e.g., 3 times greater than the burner flow, is supplied to always keep the overall fuel concentration in the chamber below the lean flammability limit. A 60.5-mm-diameter sintered brass (porosity = 0.365) plate on the exit plane of the flat burner is water cooled so that the methane-air mixture temperature remains nearly constant.

A portable USB-connected spectrometer (Ocean optics USB 2000+, 600 grooves/mm grating) is used for collecting the flame emission spectra in a broad spectral range from 250 to 850 nm with relatively low spectral resolution (0.5 nm). An optical cable is connected to the entrance of the spectrometer, and a UV-camera lens (UV-Nikkor, f = 105 mm) is used to guide the flame emission toward the other end of the optical cable. The spectra are recorded with varied exposure (gate) times, 0.05, 0.2, and 2 s. The 2 s exposure is sufficiently long to record high-SNR spectrum signals, while the reduced exposure time significantly lowers the SNR blurring characteristic spectral features.

2.3. Training and test dataset

To train the neural network model, short and long exposure spectrum pairs \((x_i; y_i)^{N_{\text{pairs}}}_{i=1}\) of the chemiluminescence spectra from a stable methane-air flat flame are prepared. Here, \(x_i\) and \(y_i\) represent the short-gated (0.2 s) and long-gated (2 s) spectrum data, respectively. The dark spectra collected without flame are subtracted from the chemiluminescence signals. Then, the chemiluminescence signals are normalized by the average intensity of OH*
bands between 306 and 313 nm, which is a prominent feature in the overall gas property condition tested in this study. The total number of the data pairs used for training and testing the neural network model is 80,000: the combination of 100 low-SNR spectra and 10 high-SNR chemiluminescence signals in each of the 80 different test conditions of varied $\phi$ (0.8 – 1.2) and pressure (1 – 10 bar). As shown in Fig. 3, the training data denoted by green squares are sampled based on the full factorial design of 50 different gas conditions, and the test data denoted by black circles are chosen using Latin hypercube sampling of 30 gas conditions [27].

![Figure 3 Experimental data matrix.](image)

2.4. Neural network architecture

The proposed neural network architecture for denoising the short-gated low-SNR spectra is described in Fig. 4, which combines reversible down- and up-sampling (DU) operators and a CNN. The input signal is a low-SNR emission spectrum recorded for a short exposure time. A reversible down-sampling operator reshaping the input signal vector of $W \times 1$ into a sub-signal tensor of $W/N_d \times N_d$, which is also known as the sub-pixel convolution or pixel unshuffle, is introduced to improve the efficiency of the neural network model [17]. Here, $W$ is the number of components in the input signal, and $N_d$ is the down-sampling parameter determining the number of sub-signals.

The sub-signal tensor then feeds into the CNN. Each layer of the CNN is a combination of three different types of operation: Convolution (Conv), Batch Normalization (BN) [28], and Rectified Linear Units (ReLU) [29]. The first layer combines Conv + ReLU, the middle layers consist of Conv + BN + ReLU, and the last layer needs only Conv. Similar to the architectures of Refs. [16, 17], 1) the number of channels and the kernel size of the filters are set the same for all the layers, 2) the stride (filter movement) is set to 1 without pooling to minimize data loss, and 3) the tensor is zero-padded before each CNN layer to preserve the data size. The critical hyperparameters of the CNN

...
architecture are the depth of layer ($N_l$), the number of channels ($N_c$), and the kernel (filter) size ($N_k$). A combination of $N_l = 7$, $N_c = 32$, and $N_k = 15$ are chosen empirically for minimal complexity with an acceptable performance of the baseline model unless otherwise stated. The performance and efficiency of the neural network model depending on the choice of the hyperparameters will be discussed later in Sec. 3.3 and 3.4.

After the last convolution layer of the last CNN layer, the reversible up-sampling operation, an inverse of the down-sampling operation, is executed to reconstruct the output signal which has the same pixel size as the input signal. Then, the output signal is compared with the high-SNR signal (long-gated spectrum) captured at the same condition, and the performance of the model is evaluated and optimized using a loss function consisting of mean square error (MSE) and POD coefficient error. More details on the loss function are described in Sec 2.5.

An Adam optimizer is used in the optimization process to train the neural network model for 100 epochs with a batch size of 128 [30]. A linear warmup cosine annealing learning rate scheduler is used, where cycle step size ($T_0$), cycle step magnification ($T_{m_{w k}}$), max learning rate ($\eta_{m_{ax}}$), linear warmup step size ($T_{w p}$), and decrease rate of max learning rate by cycle ($\gamma$) are 50, 1, 0.005, 10, and 0.1, respectively [31]. To check the overfitting of the model during the training, 10% of the training data (5 cases, 5000 data pairs) are randomly selected and set aside as the validation data. In addition, the intensity of the signal is randomly re-adjusted between $0.8 \times$ and $1.2 \times$ of the original signal intensity for data augmentation [22].

Figure 4 A schematic diagram of the proposed denoising neural network architecture.

2.5. Loss function

A novel loss function that combines the MSE and the POD coefficient error is proposed as below to improve the calibration accuracy and prevent data overfitting:
\[ L = \frac{1}{N} \sum_{i=1}^{N} \left[ (1 - \alpha)(\hat{y}_i - y_i)^2 + \alpha\left(POD(\hat{y}_i) - POD(y_i)\right)^2 \right] \]  

where \( L, N, \alpha, \hat{y}_i, y_i, POD(\hat{y}_i), \) and \( POD(y_i) \) are respectively the loss function, the number of data pairs, a blending parameter (set to be 0.1 here), a denoised short-gated spectrum, corresponding label spectrum (high-SNR spectrum), and normalized POD coefficients of \( \hat{y}_i \) and \( y_i \). The first term is the MSE loss frequently used for image denoising processes, and the second term is the POD coefficient error loss. The POD using label data can extract property-sensitive and high-dimensional (same dimension as the input spectrum). Each POD coefficient represents the weight of each POD basis to compose a spectrum (e.g., \( \hat{y}_i \) and \( y_i \)); the dot product of the POD coefficients and the POD bases reconstructs the spectrum. The POD coefficient error, therefore, stands for the overall similarity of \( \hat{y}_i \) and \( y_i \), particularly, regarding the property-sensitive parts of the spectrum, which is critical information for improving the performance of the denoising CNN. This will be further discussed in Sec. 3.3.

2.6. Noise analysis

Instrumental noise sources are quantitatively analyzed to reveal the correlations among the SNR, exposure time, and calibration accuracy. The SNR of a spectrum signal can be estimated using the equation below, Eqn. (2), assuming that the primary noise source is the charge-coupled device (CCD) sensor in the spectrometer; the major components of the CCD noise are dark noise, read noise, and photon noise [32].

\[ SNR = \frac{N_{\text{electron}}}{\sqrt{\sigma_{\text{Photon}}^2 + \sigma_{\text{Dark}}^2 + \sigma_{\text{Read}}^2}} = \frac{\phi_p \eta \tau}{\sqrt{\phi_p \eta \tau + I_{\text{dark}} \tau + N_R^2}} \]  

where, \( N_{\text{electron}}, \sigma_{\text{Photon}}, \sigma_{\text{Dark}}, \sigma_{\text{Read}}, \phi_p, \eta, \tau, I_{\text{dark}}, \) and \( N_R \) are the number of counted electrons, photon noise, dark noise, readout noise, photon flux at the CCD, quantum efficiency, exposure time, dark current, and read noise, respectively.

3. Results and discussion

3.1. Denoising with the proposed CNN

Figure 5 demonstrates the superior denoising capability of the proposed novel CNN architecture. The spectra used here to test the model are from the test dataset that was not used to train the model. This is to confirm the performance of the model taking the input of arbitrary unknown chemiluminescence signals. The chemiluminescence spectra
acquired for 0.2 s and 2 s exposure are shown in Figs. 5 (a) and (d), which are denoted as low SNR (LS) and high SNR (HS), respectively. Characteristic local spectral features from excited molecules are clearly observed in HS spectra (Fig. 5(d)) including OH* at 306.4 nm, CH* at 431.4 nm, C_2* swan bands at 517 nm, and H_2O* at 700 – 850 nm. The strength and shape of the spectral features and the broadband background emission profile over the entire spectrum range are highly dependent on the gas properties in and near the combustion reaction zone and the chemical reactions in progress.

Figure 5 (b) and (c) present the outputs from a typical denoising CNN (N_d is 1) and the proposed CNN (N_d is 16), which are denoted as LS + CNN and LS + DU + CNN, respectively. The two models were trained using the same sets of training and validation data. Obviously, both the denoising CNN architectures could significantly lower the noise level of the spectra compared to the short-gated noisy input signals in Fig. 5 (a). Nevertheless, the relatively weak characteristic spectral features, such as CH* at lean conditions and H_2O* emission bands at low-pressure conditions, that are denoised by the LS + CNN model (Fig. 5 (b)) are different from them in the corresponding HS spectra (Fig. 5 (d)). Recall that the HS spectra are the ground truth signals. Considering that the CH* and C_2* bands are highly sensitive to the fuel concentration (\(\phi\)) and the H_2O* bands are the most prominent pressure indicator, the gas property prediction with the LS + CNN model denoising the LS spectra would be inaccurate.

On the other hand, the spectra denoised by the LS + DU + CNN model are quite well matched with the HS spectra in most details. The noise of the input LS spectra could be successfully suppressed without missing critical spectral features under various gas property conditions regardless of the noise level with the DU operator (N_d = 16). We conjecture that this is because proper down-sampling makes a wider receptive field (RF), which is the number of input pixels involved to produce one output pixel. In other words, each pixel of the output signal is denoised and reconstructed by global spectral features of the input signal with large RF. Therefore, the successful denoising process is enabled by the model decoupling the property-sensitive spectral features from the noise signal even though the noise level changes with the randomly re-adjusted overall signal intensity. This will be further discussed in Sec. 3.3.
3.2. Calibration and prediction of gas properties

To evaluate the validity of the proposed CNN quantitatively, gas properties (P and $\phi$) are predicted by calibrating 1) the raw LS spectrum signals, and the emission spectra denoised-reconstructed by the 2) LS + CNN and 3) LS + DU + CNN architectures (Fig. 6). The x-axis in Fig. 6 denotes the ground truth values which are measured by high-accuracy sensors, and the y-axis denotes the predictions by the short-gated (0.2 s) chemiluminescence spectra (a) without and with the denoising CNN of (b) LS + CNN, and (c) LS + DU + CNN. The error bars indicate the standard deviations of the gas property predictions from 100 short-gated LS chemiluminescence signals, and the uncertainty bands of the sensor-measured P and $\phi$ are also presented in the figures (gray dotted lines), which are calculated using the uncertainty propagation equation by a first-order Taylor series expansion [33]. Even though the LS + CNN could effectively denoise the LS spectrum signals as shown in Fig. 5 (b), the accuracy of the gas property prediction via calibrating the LS + CNN processed spectra is not much improved compared to the prediction by the raw LS signals without the denoising CNN process. However, LS + DU + CNN remarkably improves the accuracy and precision of the property prediction, P and $\phi$ in Fig. 6 (c). This is because LS + DU + CNN can effectively remove the noise from the LS spectrum signals with minimal information loss as confirmed in Fig. 5 (c).

Table 1 summarizes the performance of the three property-prediction based on LS, LS + CNN, and LS + DU + CNN. The average relative errors of calibration using the training data (REC), the relative errors of prediction with the test data (REP), and the relative standard deviations of the prediction (RSD) are tabulated, which are to evaluate the calibration accuracy of training data, the prediction accuracy of test data, and the prediction precision of test data, respectively [34]. As expected, LS + DU + CNN exhibits superior prediction accuracy and precision. Particularly, the accuracy (REP) and precision (RSD) of the $\phi$-prediction are only 1.5% and 1.6%, respectively, which are comparable to the sensor measurement uncertainty, commonly not achievable with optical measurements. Considering that the
exposure of the portable spectrometer is only 0.2 seconds without any signal intensifier, the accuracy of the $\phi$-prediction taking the short-gated flame emission spectrum as the sole input is exceptionally high.

![Figure 6](image)

**Figure 6** P and $\phi$-prediction using (a) LS, (b) LS + CNN, and (c) LS+DU+CNN.

| Error | Unit: % | LS | LS + CNN | LS + DU + CNN |
|-------|--------|----|---------|---------------|
| $P$   | REC    | 14 | 8.8     | 2.2           |
|       | REP    | 12 | 9.5     | 5.7           |
|       | RSD    | 14 | 11      | 6.4           |
| $\phi$| REC    | 4.8| 2.0     | 0.56          |
|       | REP    | 4.1| 4.0     | 1.5           |
|       | RSD    | 5.0| 4.8     | 1.6           |

*REC: average relative error of training data in calibration, REP: average relative error of test data in prediction, RSD: average relative standard deviation of test data in prediction.

3.3. Neural network architecture and loss function

Hyperparameter settings for the neural network architecture and the configuration of the loss function are important because they determine the capacity and efficiency of the model. Figure 7 illustrates the REP distribution depending on the choice of hyperparameters and loss function configurations; the REP, prediction error of $P$ and $\phi$, represents the performance of the CNN architecture. The origin of the graph in Fig. 7 means zero prediction error for both $P$ and $\phi$; therefore, the prediction accuracy improves as the markers get closer to the origin. The blue color of the markers in Fig. 7 denotes the prediction accuracy with the newly proposed loss function combining MSE and POD losses, and the green color indicates the REP’s with the conventional loss function considering only MSE loss. The brightness of the color indicates $N_1$, which increases from 2 to 15 as it gets darker, and the shapes of the marker
indicate six different combinations of $N_k$ (3, 15, and 25) and $N_d$ (1 and 16).

In general, the prediction error decreases as the $N_l$ and $N_k$ increase because of its model complexity. More importantly, it is evident that the DU operator ($N_d = 16$) and the new loss function (MSE + POD) remarkably reduce the error. On the other hand, the prediction errors of some cases with the proposed CNN utilizing the conventional MSE loss function are even greater than that of raw LS spectra without any denoising process. This is because the denoising CNN can misinterpret the property information contained in the raw spectrum data when misguided by the MSE loss function. Therefore, the new loss function that additionally considers the POD loss should be used for denoising the short-gated LS spectra, which helps to preserve and reconstruct the property-sensitive spectral features.

Figure 7 indicates that the prediction accuracy can be improved by increasing $N_k$, $N_d$, and $N_l$ which will then significantly increase the RF, i.e., the number of input pixels involved to produce one output pixel. The RF of the proposed CNN is given as below:

$$\text{Receptive field (RF)} = N_d \times (N_l \times (N_k - 1) + 1)$$  \hspace{1cm} (3)

In this study, we found that this is another important control parameter determining the performance of the deep-learning architecture developed for analyzing spectra. Figure 8 describes the non-monotonic trends of REP and RSD as functions of the RF. The $N_l$ (2 – 15), $N_k$ (3 – 45), and $N_d$ (1 – 32) are varied in wide ranges to reveal the impact of RF on the model performance, not the separate influences of the three parameters as in Fig. 7. $N_d$ is 1 for ‘CNN’ and varied from 2 to 32 for ‘DU + CNN’. The REP and RSD are nearly constant with the RF under 100 but decrease rapidly until the RF reaches the pixel number of the input spectrum ($W$, 1696 × 1), and both the performance indicators (REP and RSD) approach the minima at around 2 × $W$. It is noteworthy that each of the output pixels will be
constructed using all the input pixels when the RF is $2 \times W$; therefore, further increase of the RF beyond $2 \times W$ will not help to improve the model performance. In conclusion, the RF of the CNN needs to be set between $W$ and $2 \times W$ to minimize the REP and RSD, or around $W$ to reduce the model complexity with acceptable prediction accuracy and precision.

In previous investigations on denoising CNN architecture for 2D images, the optimal RF size has been discussed, e.g., between $35 \times 35$ and $61 \times 61$ suggested in Ref. [17]. However, the flame emission spectrum has a unique characteristic clearly distinguished from typical 2D images; all the pixels, not only the neighboring ones but also the pixels far separated, are closely related. For example, 1) multiple molecular bands such as OH*, CH*, C2*, and H2O* that are far separated in the spectrum will get stronger at the same time as $\phi$ increases, 2) the widths of the bands will broaden as $P$ rises, and 3) the broadband background emission will become stronger as the gas density and CO2 concentration increase. In short, expanding the RF for exploiting the global contextual information of the emission spectra would improve the prediction accuracy while keeping it below $2 \times W$, because the pixels in the emission spectrum are all related (unlike in common 2D images).

POD effectively extracts these correlations among the pixels in a broad range of wavelengths that are sensitive to the properties of interest, which can enhance the accuracy and precision of FES [26]. Figure 9 (a) shows the property-sensitive POD bases extracted from the HS training dataset. Basis 1 represents the P-sensitive spectrum component, Basis 2 is the most $\phi$-sensitive, and Basis 3 indicates the correlations among the emission bands. Figure 9 (b) shows that the first 4 – 5 bases contain over 99.7% of the spectrum energy, therefore, any HS spectrum or denoised LS spectrum can be fully interpreted by combining these dominant bases (e.g., dot products of POD coefficients and POD.
bases) that define the relationships of all the pixels in the spectrum. Recall that a reduced order model (kriging) predicts the gas properties based on the denoised LS input spectrum profile being decomposed by the dominant bases; the decomposition process brings out the POD coefficients (weights of the bases) to calculate the gas properties via the kriging model previously trained by the training dataset. Therefore, the new loss function including the POD loss can preserve global features of LS signal by decoupling random noise signal from property-sensitive spectrum information and prevent over-fitting, while the conventional loss functions simply consider the mean square errors (MSE) of the spectra regardless of the sensitivity of the spectrum profile to the properties of interest.

Figure 9 (a) POD bases extracted from the training data (HS), and (b) the energy contained in primary bases

3.4. Computational Efficiency

Computational efficiency is another important parameter that defines the performance of a deep learning-based method. The accuracy and precision of prediction can be improved by increasing RF as shown in Fig. 8; however, the increasing $N_t$ and $N_k$ along with the RF costs the neural network computational complexity to lengthen the forward calculating time (running time shown in Fig. 10) and limits real-time gas property measurements. Nevertheless, increasing the RF via implementing the DU operators and increasing $N_d$ accelerates the running and training of a neural network model [17]. The computation time of the proposed CNN model is estimated in a Python environment on a computer with an NVIDIA GeForce RTX 3090, an AMD Ryzen 5 2600X, and 32GB of RAM as illustrated in Figs. 10 (a) and (b). The running time shown in Fig. 10 (a) is calculated by averaging the time with 5,000 inputs using batch size as 1, and the training time in Fig. 10 (b) is estimated by the time used to train a model for 100 epochs.

The trends observed in Figs. 10 (a) and (b) confirm that the DU operator accelerates the computation under a given RF; widening the RF of the CNN by increasing $N_d$ is much more cost-effective than increasing $N_k$ or $N_t$ while $N_d$, $N_t$, or $N_k$ increases the RF by the same order of magnitude (Eqn. (3)). The DU operator ($N_d > 1$) reduces the computational time because the range of convolution would decrease via reshaping the input pixels into sub-signals. Moreover, the number of training parameters ($N_p$), calculated as Eqn. (4), is always smaller with increased $N_d$ than
that with increased $N_k$ or $N_l$ at a given RF as shown in Fig. 10 (c); increasing $N_p$ adds computational cost, i.e., complexity of the CNN architecture.

$$\# \text{ of parameters } (N_p) = (N_l - 2) \times (N_c^2 N_k + N_c) + 2N_d N_c N_k + (N_c + N_d) \quad (4)$$

$N_p$ increases with all of the three parameters, $N_l$, $N_k$, and $N_d$; however, $N_l$ increases $N_p$ by factor of $N_c$ while $N_l$ and $N_k$ increase $N_p$ by factor of $N_c^2$ ($> N_c$). Thus, $N_d$ less affects the computational cost.

![Graphs showing running time, training time, and $N_p$ versus receptive fields of CNN ($N_d = 1$) and DU + CNN ($N_d = 2 - 32$).]

3.5 Exposure time

To monitor the gas properties in fast-evolving combustion environments using FES, the exposure time should be minimized. Here, the effect of the reduced exposure time on the performance of the calibration model is evaluated by repeating the experiments with varied exposure times, 0.05, 0.2, and 0.4 s. In Table 2, each component of the noise with the three different exposure times at $P = 10$ bar and $\phi = 1$ is estimated. Presumably, the constant read-out noise is responsible for the decreased SNR with the reduced exposure time. Figure 11 (a) confirms that the SNR of the chemiluminescence signal decreases with reduced exposure time; the average SNR of OH* at 308nm is 4.3, 14, and 22 for 0.05, 0.2, and 0.4 s exposure, respectively. Prediction accuracies of LS, LS + CNN, and LS + DU + CNN, defined as 100 minus REP, are presented in Figs. 11 (b) and (c). As expected, the accuracy drops with decreasing exposure time. Again, the LS + DU + CNN model performs the best (the most accurate) among the three calibration methods; P- and $\phi$-prediction accuracies (100 minus REP) are above 80% and 95%, respectively, even with 0.05 s exposure.
Table 2 Instrumental noise analysis of OH* signal at 308 nm at P = 10 bar and ϕ = 1.

| Noise source | Exposure time |
|--------------|---------------|
|               | 0.05 s | 0.2 s | 0.4 s |
| N_{background} | 115   | 463   | 934   |
| σ_{background} | 11    | 22    | 31    |
| σ_{Dark}      | 4.6    | 9.1   | 13    |
| σ_{Read}      | 24    | 24    | 24    |
| SNR           | 4.3   | 14    | 22    |

Figure 11 (a) Typical chemiluminescence signals with the three different exposure times (0.05, 0.2, and 0.4 s) at P = 10 bar and ϕ = 1, and (b) P- and (c) ϕ-prediction accuracy versus exposure time.

4. Conclusion remarks

A novel deep-learning architecture was proposed to enable accurate gas property measurements using a short-gated flame emission spectrum under high-pressure conditions. The new architecture utilizes a reversible down- and up-sampling (DU) operator, deep CNN layers, and a newly proposed loss function to train with the pairs of low-SNR (LS) and high-SNR (HS) spectra, i.e., training data, captured with short and long exposure times, respectively. The proposed neural network architecture successfully denoised the LS spectrum signals while conserving the characteristic spectral features in the spectrum. Then, the POD and a reduced order model (kriging) that are trained by the HS spectra from the training dataset were used to predict the gas properties from the denoised short-gated LS spectrum. The prediction errors of the P and ϕ employing the new technique were improved to approximately 5.7% and 1.5% with 0.2 s exposure time (from 11% and 4% without denoising but with the POD and kriging model), and to 18% and 5% with 0.05 s exposure time (from 43% and 16% without denoising), respectively.

It was confirmed that the combination of the proposed CNN model and the new loss function including POD coefficient loss can effectively decouple the noise and signal and selectively suppress the noise that increases with the decreasing exposure time by a data-driven approach. The proper choice of the down-sampling parameter (Nd) in the
DU operator in CNN architecture reduced the model complexity ($N_l$ and $N_k$) and accelerated the calculation with increased receptive field (RF). Another advantage of the technique comes from the use of the proper orthogonal decomposition (POD) method in the denoising and calibration procedure. POD can effectively extract the property-sensitive components from the emission spectra. Therefore, the new loss function including POD loss can better guide the denoising CNN by considering the property-sensitive global features of the emission spectra.

It is evident that the proposed calibration method can further improve the time-resolution and accuracy of FES when using more sensitive and faster photon detectors such as high-speed kHz-framing cameras with a signal intensifier; accurate gas property monitoring at hundreds of kHz sampling rates is possible with the devices currently available in the market. In addition, the newly proposed denoising technique can also be used to denoise any other forms of multi-dimensional data, e.g., 2D/3D images, when reference data pairs of high SNR and low SNR signals are available.

Acknowledgment

This work was supported by the Basic Research Funding of the Air Force Office of Scientific Research [FA2386-20-1-4054]; Korean Agency for Defense Development [16-106-501-035]; and the National Research Foundation of Korea [2021R1A2C2012697, 2021R1A4A1032023].
Reference

[1] A.M. Steinberg, J.F. Driscoll, Straining and wrinkling processes during turbulence-premixed flame interaction measured using temporally-resolved diagnostics, Combust. Flame 156 (2009) 2285-2306.
[2] R. Zhang, I. Boxx, W. Meier, C.D. Slabaugh, Coupled interactions of a helical precessing vortex core and the central recirculation bubble in a swirl flame at elevated power density, Combust. Flame 202 (2019) 119-131.
[3] A. Giusti, E. Mastorakos, Turbulent Combustion Modelling and Experiments: Recent Trends and Developments, Flow Turbul. Combust. 103 (2019) 847-869.
[4] E. Karlis, Y.S. Liu, Y. Hardalupas, A.M.K.P. Taylor, Extinction strain rate suppression of the precessing vortex core in a swirl stabilised combustor and consequences for thermoacoustic oscillations, Combust. Flame 211 (2020) 229-252.
[5] V. Sick, High speed imaging in fundamental and applied combustion research, Proc. Combust. Inst. 34 (2013) 3509-3530.
[6] B. Bohm, C. Heeger, R.L. Gordon, A. Dreizler, New Perspectives on Turbulent Combustion: Multi-Parameter High-Speed Planar Laser Diagnostics, Flow Turbul. Combust. 86 (2011) 313-341.
[7] P. Parajuli, T.T. Paschal, M.A. Turner, Y.J. Wang, E.L. Petersen, W.D. Kulatilaka, High-Speed Hydroxyl and Methylidyne Chemiluminescence Imaging Diagnostics in Spherically Expanding Flames, AIAA J. 59 (2021) 3118-3126.
[8] V. Athmanathan, J. Braun, Z.M. Ayers, C.A. Fugger, A.M. Webb, M.N. Slipchenko, G. Paniagua, S. Roy, T.R. Meyer, On the effects of reactant stratification and wall curvature in non-premixed rotating detonation combustors, Combust. Flame 240 (2022).
[9] M. Severin, O. Lammel, W. Meier, Laser diagnostic investigation of a confined premixed turbulent jet flame stabilized by recirculation, Combust. Flame (2022), https://doi.org/10.1016/j.combustflame.2022.112061.
[10] C.D. Slabaugh, C.N. Dennis, I. Boxx, W. Meier, R.R. Lucht, 5 kHz thermometry in a swirl-stabilized gas turbine model combustor using chirped probe pulse femtosecond CARS. Part 2. Analysis of swirl flame dynamics, Combust. Flame 173 (2016) 454-467.
[11] K.A. Rahman, E.L. Braun, M.N. Slipchenko, S. Roy, T.R. Meyer, Flexible chirp-free probe pulse amplification for kHz fs/ps rotational CARS, Opt. Lett. 45 (2020) 503-506.
[12] M. Gragston, P. Hsu, N.B. Jiang, S. Roy, Z.L. Zhang, Emissions in short-gated ns/ps/fs-LIBS for fuel-to-air ratio measurements in methane-air flames, Appl. Opt. 60 (2021) C114-C120.
[13] S.J. Grauer, A. Unterberger, A. Rittler, K.J. Daun, A.M. Kempf, K. Mohri, Instantaneous 3D flame imaging by background-oriented schlieren tomography, Combust. Flame 196 (2018) 284-299.
[14] G. Wang, J.C. Ye, K. Mueller, J.A. Fessler, Image Reconstruction is a New Frontier of Machine Learning, IEEE Trans. Med. Imaging 37 (2018) 1289-1296.
[15] J. Schaeferkoetter, J.H. Yan, C. Ortega, A. Sertic, E. Lechtaen, Y. Eshet, U. Metser, P. Veit-Haibach, Convolutional neural networks for improving image quality with noisy PET data, EJNMMI Res. 10 (2020).
[16] K. Zhang, W.M. Zuo, Y.J. Chen, D.Y. Meng, L. Zhang, Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising, IEEE Trans. Image Process. 26 (2017) 3142-3155.
[17] K. Zhang, W.M. Zuo, L. Zhang, FFDNet: Toward a Fast and Flexible Solution for CNN-Based Image Denoising, IEEE Trans. Image Process. 27 (2018) 4608-4622.
[18] C.W. Tian, L.K. Fei, W.X. Zheng, Y. Xu, W.M. Zuo, C.W. Lin, Deep learning on image denoising: An overview, Neural Networks 131 (2020) 251-275.
[19] B. Manifold, E. Thomas, A.T. Francis, A.H. Hill, D. Fu, Denoising of stimulated Raman scattering microscopy images via deep learning, Biomed. Opt. Express 10 (2019) 3860-3874.
[20] G.S. Zhang, H. Hao, Y.C. Wang, Y. Jiang, J.H. Shi, J. Yu, X.J. Cui, J.S. Li, S. Zhou, B.L. Yu, Optimized adaptive Savitzky-Golay filtering algorithm based on deep learning network for absorption spectroscopy, Spectrochim. Acta A 263 (2021).
[21] T. Miyatake, K. Shiokawa, H. Sakaki, N.P. Dover, M. Nishiuchi, H.F. Lowe, K. Kondo, A. Kon, M. Kando, K. Kondo, Y. Watanabe, Denoising application for electron spectrometer in laser-driven ion acceleration using a Simulation-supervised Learning based CDAE, Nucl. Instrum. 999 (2021).
[22] Y. Kim, D. Oh, S. Huh, D. Song, J. Jeong, J. Kwon, M. Kim, D. Kim, H. Ryu, J. Jung, W. Kyung, B. Sohn, S. Lee, J. Hyun, Y. Lee, Y. Kim, C. Kim, Deep learning-based statistical noise reduction for multidimensional spectral data, Rev. Sci. Instrum. 92 (2021) 073901.
[23] J. Kojima, Y. Ikeda, T. Nakajima, Spatially resolved measurement of OH*, CH*, and C2* chemiluminescence in the reaction zone of laminar methane/air premixed flames, Proc. Combust. Inst. 28 (2000) 1757-1764.

[24] M.M. Tripathi, S.R. Krishnan, K.K. Srinivasan, F.Y. Yueh, J.P. Singh, Chemiluminescence-based multivariate sensing of local equivalence ratios in premixed atmospheric methane-air flames, Fuel 93 (2012) 684-691.

[25] J. Ballester, R. Hernandez, A. Sanz, A. Smolarz, J. Barroso, A. Pina, Chemiluminescence monitoring in premixed flames of natural gas and its blends with hydrogen, Proc. Combust. Inst. 32 (2009) 2983-2991.

[26] T. Yoon, Y. Kang, S.W. Kim, Y. Park, K. Yee, C.D. Carter, S.D. Hammack, H. Do, Proper orthogonal decomposition of continuum-dominated emission spectra for simultaneous multi-property measurements, Energy (2022), https://doi.org/10.1016/j.energy.2022.124458.

[27] J.C. Helton, F.J. Davis, Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems, Reliab. Eng. Syst. 81 (2003) 23-69.

[28] S. Ioffe, C. Szegedy, Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift, Proceedings of the 32nd International Conference on Machine Learning, PMLR (2015), 448-456.

[29] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, Adv. Neural. Inf. Process Syst. 25 (2012) 1097-1105.

[30] D.P. Kingma, J. Ba, Adam: A method for stochastic optimization, arXiv:1412.6980 (2014), https://doi.org/10.48550/arXiv.1412.6980.

[31] I. Loshchilov, F. Hutter, Sgdr: Stochastic gradient descent with warm restarts, arXiv:1608.03983 (2016), https://doi.org/10.48550/arXiv.1608.03983.

[32] T.S. Tkaczyk, Field guide to microscopy, SPIE Press, Wash., USA, 2010.

[33] R.C. Smith, Uncertainty quantification : theory, implementation, and applications, Society for Industrial and Applied Mathematics, Philadelphia, USA, 2013.

[34] J.B. Sirven, B. Bousquet, L. Canioni, L. Sarger, S. Tellier, M. Potin-Gautier, I. Le Hecho, Qualitative and quantitative investigation of chromium-polluted soils by laser-induced breakdown spectroscopy combined with neural networks analysis, Anal. Bioanal. Chem. 385 (2006) 256-262.