CSTNet: A Dual-Branch Convolutional Neural Network for Imaging of Reactive Flows using Chemical Species Tomography

CSTNet: A Dual-Branch Convolutional Neural Network for Imaging of Reactive Flows Using Chemical Species Tomography

Yunfan Jiang, Jingjing Si, *Member, IEEE*, Rui Zhang, Godwin Enemali, *Member, IEEE*, Bin Zhou, Hugh McCann, and Chang Liu, *Senior Member, IEEE*

***Abstract*— Chemical species tomography (CST) has been widely used for *in situ* imaging of critical parameters, e.g., species concentration and temperature, in reactive flows. However, even with state-of-the-art computational algorithms, the method is limited due to the inherently ill-posed and rank-deficient tomographic data inversion and by high computational cost. These issues hinder its application for real-time flow diagnosis. To address them, we present here a novel convolutional neural network, namely CSTNet, for high-fidelity, rapid, and simultaneous imaging of species concentration and temperature using CST. CSTNet introduces a shared feature extractor that incorporates the CST measurements and sensor layout into the learning network. In addition, a dual-branch decoder with internal crosstalk, which automatically learns the naturally correlated distributions of species concentration and temperature, is proposed for image reconstructions. The proposed CSTNet is validated both with simulated datasets and with measured data from real flames in experiments using an industry-oriented sensor. Superior performance is found relative to previous approaches in terms of reconstruction accuracy and robustness to measurement noise. This is the first time, to the best of our knowledge, that a deep learning-based method for CST has been experimentally validated for simultaneous imaging of multiple critical parameters in reactive flows using a low-complexity optical sensor with a severely limited number of laser beams.**

***Index Terms*— Chemical species tomography (CST), convolutional neural network (CNN), deep learning, inverse problem.**

I. Introduction

IN THE past two decades, chemical species tomography (CST) has been widely applied for nonintrusive and sensitive imaging of multiple critical parameters in reactive flows, e.g., gas-phase species concentration [1]–[3], temperature [2]–[3], and velocity [4]. To solve the inverse problem of CST, a variety of computational tomographic algorithms have been developed. Some of them are adapted from previous hard-field tomography modalities and improved subject to the characteristics of the flow field. These algorithms can be mainly categorized as follows:

1) algebraic techniques based on linear back projection, e.g., algebraic reconstruction technique (ART) [5] and Landweber algorithm [1], [6];
2) regularization methods, e.g., Tikhonov regularization [7];
3) global optimization, e.g., simulated annealing [8];
4) statistical inversion, e.g., covariance estimation [9];
5) dimensional reduction techniques, e.g., surrogate functions method [10].

As discussed in [11], these algorithms can, to some extent, mitigate the difficulty of robust image reconstruction in CST. Nevertheless, deficiencies still remain, particularly the appearance of artifacts in the reconstructed images due to the rank-deficient tomographic data inversion and the high computational cost incurred due to the complex mathematical operations. These issues severely hinder the exploitation of CST for applications where high-fidelity performance and real-time image reconstruction are required.

The last ten years have witnessed a boom in the use of learning-based algorithm for bioimaging and medical imaging [12]–[14]. Extreme learning machine (ELM)-based method [15] has been pioneered in the simulative study of CST as it enjoys fast training speed. However, reconstruction accuracy may be limited since only one iteration of learning is performed during ELM training. In contrast, the convolution operation performed in convolutional neural networks (CNNs) [16] can take account of the spatial features and correlations of data and incorporate and learn the inherent domain priors for better image reconstruction. CNN is a good candidate to overcome the above-mentioned issues in CST due to the following properties.

1) Automatic discovery of intricate features. In comparison with the computational CST algorithms with constraints manually imposed on the flow fields, CNN can automatically construct data representations during the learning process [16], enabling end-to-end (i.e., from measurements to reconstructed images) learning of intricate features of the flow fields with superior generalization ability.

Manuscript received October 8, 2020; revised March 27, 2021, June 14, 2021, and September 5, 2021; accepted February 28, 2022. This work was supported by the U.K. Engineering and Physical Sciences Research Council under Grant EP/P001661/1. *(Corresponding author: Chang Liu.)*

Yunfan Jiang, Rui Zhang, Godwin Enemali, Hugh McCann, and Chang Liu are with the School of Engineering, The University of Edinburgh, Edinburgh EH9 3JL, U.K. (e-mail: c.liu@ed.ac.uk).

Jingjing Si is with the School of Engineering, The University of Edinburgh, Edinburgh EH9 3JL, U.K., and also with the School of Information Engineering, Yanshan University, Qinhuangdao 066004, China.

Bin Zhou is with the School of Energy and Environment, Southeast University, Nanjing 210018, China.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TNNLS.2022.3157689.

Digital Object Identifier 10.1109/TNNLS.2022.3157689

2) Accurate image retrieval with strong robustness. CNN can capture and learn distinct features of the flow fields without overfitting the tomographic data, thus yielding strong immunity to measurement noise. As the training sets are established from CST measurements of noise-free target fields, CNN trained under supervision intrinsically addresses the limited sampling of CST systems, thus helping to eliminate artifacts in the tomographic images.

3) Rapid and cost-effective computation, empowered by hardware acceleration, for rapid processing of a large amount of CST data typically acquired. This has the potential to facilitate online imaging and thus real-time flow and combustion diagnosis.

CNN has been demonstrated recently in CST simulations to perform spatially resolved measurements in combustion diagnosis [17], [18]. Previous work employed CNNs in CST and showed that their models could achieve a similar accuracy level as simulated annealing [17] and the reduction in network parameters [18]. In addition, CNN has been applied in a proof-of-concept experiment to reconstruct the 3-D distribution of methane concentration using mid-infrared CST [19]. Although these endeavors are promising for the industrial application of CNN in CST, the following three issues remain to be addressed as a matter of urgency.

1) The properties of the CST measurement system. The only *a priori* information taken into account by the previous work was that pertaining to the attributes of the phantoms themselves, i.e., smooth distributions of species concentration and temperature. However, more in-depth features, e.g., smoothness and centrosymmetry in the CST measurement and sensor layout, were not considered in the learning architectures.

2) Interdependence of species concentration and temperature distributions. The previous work assumed independence between species concentration and temperature distributions, neglecting their internal correlation in combustion processes. Furthermore, these efforts were incapable of multiparameter retrieval. They can only reconstruct either species concentration distribution or temperature distribution with a single neural network model.

3) Practicality for industrial applications. Previous optical sensors used up to six angular views and tens of laser beams per view, greatly assisting the image reconstruction process. However, severely limited optical access with fewer laser beams commonly exists in industrial applications, e.g., for the purpose of reliability maintenance. Such limited projection data place more rigorous challenges on the applicability of CNNs in CST.

To address these issues, we propose here a novel CNN, namely, CSTNet, for high-fidelity, rapid, and simultaneous imaging of species concentration and temperature using CST. It incorporates and learns domain priors of CST measurements via convolution operations. Furthermore, it also consists of a dual-branch decoder with internal crosstalk to integrate correlations between different flow-field parameters, e.g., gas concentration and temperature, such that they can be simultaneously reconstructed with high fidelity. Our contributions are summarized as follows.

1) We incorporate domain priors of CST, for the first time, into the model design, enabling our neural network to learn inherently physical characteristics of CST.
2) To simultaneously reconstruct multiple flow-field parameters, we design a dual-branch decoder with internal crosstalk that considers the natural correlation between these multiple parameters. This architecture enables more reasonable and efficient image reconstruction in practical combustion processes.
3) The proposed CSTNet is both analytically and experimentally proven to be successful in high-fidelity imaging of 2-D H_2O concentration and temperature distributions using two H_2O transitions and a tomographic sensor with 32 laser beams. To the best of our knowledge, this is the first time that a deep learning-based method for CST has been experimentally validated for simultaneous imaging of multiple critical parameters in reactive flows using a low-complexity optical sensor with a severely limited number of laser beams.

The remainder of this article is organized as follows. Based on the basic formulations of CST, we first introduce the domain priors in CST and the architecture of CSTNet in Section II. Then, the CSTNet model is established using the 32-beam CST sensor and adapting our specific task in Section III. Subsequently, we train the established neural network, examine its performance, and compare with two state-of-the-art methods with simulated test sets in Section IV. The proposed CSTNet is further experimentally validated in Section V. Finally, conclusions are presented in Section VI.

II. Methodology

A. Mathematical Formulation of CST

CST is implemented by multiple line-of-sight tunable diode laser absorption spectroscopic (TDLAS) measurements governed by the Beer–Lambert law [20], [21]. When a laser beam at frequency υ (cm^{-1}) penetrates an absorbing gas sample with a path of length L (cm), the wavelength-dependent absorbance, $\alpha(\upsilon)$, is defined as

$$\alpha(\upsilon) = \ln \frac{I_0(\upsilon)}{I_t(\upsilon)} = \int_0^L P(l)X(l)S(T(l))\phi(\upsilon)dl \tag{1}$$

where $I_0(\upsilon)$ and $I_t(\upsilon)$ are the incident and transmitted laser intensities, respectively, l is the local position along the path, $P(l)$ (atm) is the local pressure, $X(l)$ is the local molar fraction of the absorbing species, $T(l)$ (K) is the local temperature, $S(\cdot)$ ($\text{cm}^{-2}\text{atm}^{-1}$) is the temperature-dependent line strength, and $\phi(\cdot)$ (cm) is the line shape function [11]. Line strength is defined as a temperature-dependent function for each individual transition, which is a discontinuous change of an electron from one energy level to another within an atom. We denote the line strength of transition i at the central wavelength υ_i (cm^{-1}) as S_{υ_i} and the corresponding wavelength-dependent absorbance as $\alpha_{\upsilon_i}(\upsilon)$.

Since the line shape function is normalized to unity, i.e., $\int_{-\infty}^{+\infty} \phi(\upsilon)d\upsilon \equiv 1$, the path integrated absorbance for

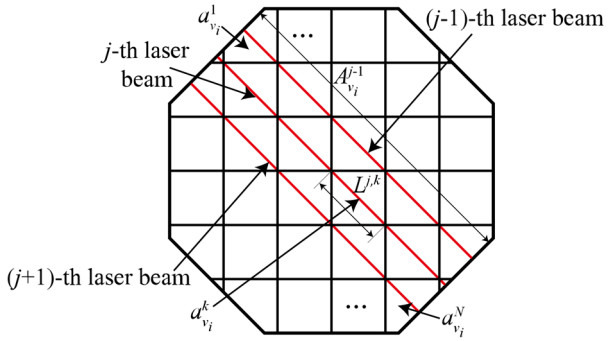


Fig. 1. Geometric description of CST measurements in the discretized RoI.

transition i, A_{υ_i}, can be formulated as

$$A_{\upsilon_i} = \int_{-\infty}^{+\infty} \alpha_{\upsilon_i}(\upsilon)d\upsilon = \int_0^L P(l)X(l)S_{\upsilon_i}(T(l))dl = \int_0^L a_{\upsilon_i}(l)dl \tag{2}$$

where $a_{\upsilon_i}(l)$ is the local density of A_{υ_i}.

The problem of CST is formulated by employing M line-of-sight measurements and discretizing the region of interest (RoI) into N pixels, as shown in Fig. 1. It results in a matrix form

$$\boldsymbol{A}_{\upsilon_i} = \boldsymbol{L}\boldsymbol{a}_{\upsilon_i} \tag{3}$$

where $\boldsymbol{A}_{\upsilon_i} \in \mathbb{R}^{M\times 1}$ denotes the vector of path integrated absorbance obtained from M CST measurements, with its element $A^j_{\upsilon_i}$ representing the path integrated absorbance of the jth beam. Assuming that the CST sensor consists of Q angular views and R parallel projection beams per view, M satisfies the relation of $M = Q \times R$. $\boldsymbol{L} \in \mathbb{R}^{M\times N}$ is the sensitivity matrix with its element $L^{j,k}$ representing the length of the laser path segment for the jth laser beam passing through the kth pixel. $j \in \{1, 2, \ldots, M\}$ and $k \in \{1, 2, \ldots, N\}$ are the indices of laser beams and pixels, respectively. $\boldsymbol{a}_{\upsilon_i} \in \mathbb{R}^{N\times 1}$ is the vector of absorbance density with its elements $a^k_{\upsilon_i} = P^k X^k S_{\upsilon_i}(T^k)$, where P^k, X^k, and T^k are the local pressure, local molar fraction, and local temperature in the kth pixel, respectively.

In common with X-ray CT, the reconstruction of images in CST uses mathematical methods from the field of inverse problems to solve the spatial distribution of molar fraction of a target species and/or the temperature. Concretely, we use CST measurements, i.e., $\boldsymbol{A}_{\upsilon_1}$ and $\boldsymbol{A}_{\upsilon_2}$, from two preselected transitions [11] at central wavelengths υ_1 and υ_2, and the sensitivity matrix $\boldsymbol{L}$, to simultaneously yield images of species molar fraction, i.e., concentration, and temperature in this article. More details regarding the principles, implementation, and application of CST can be found in review papers [11], [22], [23].

B. Domain Priors in CST

1) Smoothness: The optical layout of the CST sensor is given by Q angular views and R parallel laser beams per view, satisfying $Q \times R = M$. Imposed by the adjacent arrangement of laser beams, the path integrated absorbance within the qth angular view θ_q, $\boldsymbol{A}_{\upsilon_i,\theta_q} \in \mathbb{R}^{R\times 1}$, experiences smooth change from beam to beam due to the smooth distributions of flow-field parameters in the RoI

$$\left| A^r_{\upsilon_i,\theta_q} - A^{r-1}_{\upsilon_i,\theta_q} \right| \le \epsilon \quad \forall r \in \{2, 3, \ldots, R\}\ \forall q \in \{1, 2, \ldots, Q\} \tag{4}$$

where ϵ is a threshold and $A^r_{\upsilon_i,\theta_q}$ is the rth element of $\boldsymbol{A}_{\upsilon_i,\theta_q}$. Note that $\boldsymbol{A}_{\upsilon_i,\theta_q} \in \mathbb{R}^{R\times 1}$ represents only a "segment" of the full measurements $\boldsymbol{A}_{\upsilon_i} \in \mathbb{R}^{M\times 1}$. Smoothness heatmaps can be constructed according to Algorithm 1 to assist the learning of such *a priori* information. Learning this domain prior enables speedy convergence with a lower loss during training. Hence, it is incorporated into CSTNet detailed in Section II-C.

Algorithm 1 Construction of Centrosymmetry Heatmaps and Smoothness Heatmaps From CST Projection Data

Input: Number of equiangular projections Q, number of equispaced parallel beams R, and CST measurements $\boldsymbol{A}_{\upsilon_1} \in \mathbb{R}^{M\times 1}$ and $\boldsymbol{A}_{\upsilon_2} \in \mathbb{R}^{M\times 1}$ where $M = Q \times R$
Output: Centrosymmetry heatmap $\mathfrak{S} \in \mathbb{R}^{2Q\times R\times 1}$ and smoothness heatmap $\mathfrak{P} \in \mathbb{R}^{R/F\times QF\times 2}$ with F the least prime factor greater than 1 of R
Initialize: Empty $\mathfrak{S}$ and $\mathfrak{P}$
$F \leftarrow$ the least prime factor greater than 1 of R
Construct $\mathfrak{S}$:
  for i in {1, 2} **do**
    $\boldsymbol{A}^{reshape}_{\upsilon_i} \leftarrow$ reshape($\boldsymbol{A}_{\upsilon_i}$, ($Q$, R))
    if υ_i is the counterpart frequency **then**
      $(Q \times (i-1)+1)$-th to $(Q \times i)$-th rows of $\mathfrak{S} \leftarrow$ vertical_flip ($\boldsymbol{A}^{reshape}_{\upsilon_i}$)
    else
      $(Q \times (i-1)+1)$-th to $(Q \times i)$-th rows of $\mathfrak{S} \leftarrow \boldsymbol{A}^{reshape}_{\upsilon_i}$
    end if
  end for
Construct $\mathfrak{P}$:
  for i in {1, 2} **do**
    for j in $\{1, 2, \ldots, Q\}$ **do**
      $\boldsymbol{A}^{patch}_{\upsilon_i} \leftarrow$ reshape($(R \times (j-1)+1)$-th to $(R \times j)$-th elements of $\boldsymbol{A}_{\upsilon_i}$, ($\frac{R}{F}$, F))
      $(F \times (j-1)+1)$-th to $(F \times j)$-th columns of the i-th channel of $\mathfrak{P} \leftarrow \boldsymbol{A}^{patch}_{\upsilon_i}$
    end for
  end for

2) Centrosymmetry: CNN exploits the hierarchical property of images and therefore is superior for recognizing and detecting patterns in the images [16]. This characteristic can assist to perceive the physical fields with CST. When the CST beam array is rotationally symmetric around the center of the RoI, e.g., invariant on 180° rotation, the measurement patterns thus created can be exploited by CNN. We illustrate this property by the following example. Fig. 2(a) and (b) shows two "phantom" temperature distributions. These two identical distributions are centrosymmetric to each other except that they are rotated by 180° around the center of the RoI. Using the 32-beam CST sensor, path integrated absorbance at two transitions υ_1 and υ_2, i.e., $\boldsymbol{A}_{\upsilon_1} \in \mathbb{R}^{32\times 1}$ and $\boldsymbol{A}_{\upsilon_2} \in \mathbb{R}^{32\times 1}$, can be measured to carry out two-line temperature imaging [11]. The centrosymmetry introduced by the CST measurement can be characterized by a pattern, named the centrosymmetry heatmap, which contains information of both $\boldsymbol{A}_{\upsilon_1}$ and $\boldsymbol{A}_{\upsilon_2}$. First, $\boldsymbol{A}_{\upsilon_1}$ and $\boldsymbol{A}_{\upsilon_2}$ are reshaped to

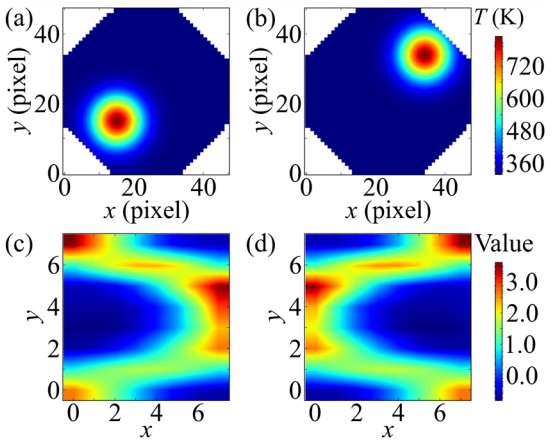


Fig. 2. Illustration of *a priori* information of centrosymmetry in CST. (a) and (b) Two centrosymmetric temperature images and (c) and (d) heatmaps containing $\boldsymbol{A}_{\upsilon_1}$ and $\boldsymbol{A}_{\upsilon_2}$ obtained from the temperature images in (a) and (b), respectively. In (c) and (d), the x-axis and y-axis indicate the horizontal and vertical indices of the heatmap, respectively. The color bar represents the values of path integrated absorbances.

$\boldsymbol{A}_{\upsilon_1}^{\text{reshape}} \in \mathbb{R}^{4\times 8}$ and $\boldsymbol{A}_{\upsilon_2}^{\text{reshape}} \in \mathbb{R}^{4\times 8}$, respectively. Then, the heatmap is constructed by concatenating $\boldsymbol{A}_{\upsilon_1}^{\text{reshape}}$ at the upper half and vertically flipped $\boldsymbol{A}_{\upsilon_2}^{\text{reshape}}$ at the lower half. As shown in Fig. 2(c) and (d), the two phantom temperature distributions lead to heatmaps with identical patterns but opposite orientations. In other words, when a fixed centrosymmetric beam geometry is employed, moving the inhomogeneity to a centrosymmetric location in the RoI is equivalent to reorientating the heatmap. Learning these heatmaps enables the prediction of both the distributions of flow-field parameters and the locations of inhomogeneities in the RoI. Therefore, centrosymmetry is adopted as another *a priori* information in CSTNet.

C. CSTNet Architecture

As shown in Fig. 3, the proposed CSTNet consists of two main parts, i.e., a shared feature extractor and a dual-branch decoder with internal crosstalk. A lambda layer is first used to generate two heatmaps from the projection data of CST. The two heatmaps are then directly learned by the feature extractor [24]. The upper and lower parts of the feature extractor learn the centrosymmetry and smoothness, respectively. Both learning processes can be formulated by

$$\boldsymbol{O} = f(\boldsymbol{W} * \boldsymbol{I} + \boldsymbol{b}) \tag{5}$$

where $\boldsymbol{I} \in \mathbb{R}^{H_I\times W_I\times C_I}$ is the input heatmaps or intermediate feature maps, $\boldsymbol{W} \in \mathbb{R}^{H_W\times W_W\times C_W}$ is the convolution kernel, $\boldsymbol{O} \in \mathbb{R}^{H_O\times W_O\times C_O}$ is the output feature maps, $\boldsymbol{b} \in \mathbb{R}^{C_b\times 1}$ is the bias vector, $f(\cdot)$ is the activation function, and $*$ is the operand for 2-D convolution. $H_I(H_O)$, $W_I(W_O)$, and $C_I(C_O)$ are height, width, and channel of $\boldsymbol{I}$ ($\boldsymbol{O}$), respectively. H_W, W_W, and C_W are the filter height, filter width, and the number of filters of $\boldsymbol{W}$, respectively. C_b is the length of $\boldsymbol{b}$.

Then, the output feature maps are flattened, concatenated, and fed into the dual-branch decoder for simultaneous imaging of species concentration and temperature. In reactive flows, for example, hydrocarbon combustion processes, the species concentration distribution is generally correlated with the temperature distribution. Therefore, their correlation is incorporated into the dual-branch decoder [25] with internal crosstalk. Both branches consist of G stages and can be simultaneously computed by

$$\boldsymbol{\mathcal{X}}_g = f_g^{\boldsymbol{\mathcal{X}}}\left(h_g^{\boldsymbol{\mathcal{X}}}\left(\boldsymbol{\mathcal{X}}_{g-1}\right) + \boldsymbol{\mathcal{W}}_g^{\mathcal{T}} \odot h_g^{\mathcal{T}}\left(\boldsymbol{\mathcal{T}}_{g-1}\right)\right) \quad g \in \{1, 2, \ldots, G\} \tag{6}$$

and

$$\boldsymbol{\mathcal{T}}_g = f_g^{\mathcal{T}}\left(h_g^{\mathcal{T}}\left(\boldsymbol{\mathcal{T}}_{g-1}\right) + \boldsymbol{\mathcal{W}}_g^{\mathcal{X}} \odot h_g^{\boldsymbol{\mathcal{X}}}\left(\boldsymbol{\mathcal{X}}_{g-1}\right)\right) \quad g \in \{1, 2, \ldots, G\} \tag{7}$$

where $f_g^{\mathcal{X}}(\cdot)$ and $f_g^{\mathcal{T}}(\cdot)$ denote activation functions in the concentration and temperature branches, $h_g^{\mathcal{X}}(\cdot)$ and $h_g^{\mathcal{T}}(\cdot)$ denote the operations before the addition of crosstalk, e.g., linear transformation and batch normalization, and $\boldsymbol{\mathcal{X}}_g \in \mathbb{R}_g^{N\times 1}$ and $\boldsymbol{\mathcal{T}}_g \in \mathbb{R}_g^{N\times 1}$ are the outputs from the gth stage of concentration and temperature branches, respectively. N_g represents the length of the output from the gth stage. $\boldsymbol{\mathcal{W}}_g^{\mathcal{T}} \in \mathbb{R}_g^{N\times 1}$ and $\boldsymbol{\mathcal{W}}_g^{\mathcal{X}} \in \mathbb{R}_g^{N\times 1}$ are the crosstalk weights in the gth stages of the concentration and temperature branches, respectively. $\odot$ denotes the elementwise production.

Equations (6) and (7), visually shown in Fig. 4, describe how the crosstalk is added into the dual-branch decoder. While the left branch in Fig. 4 corresponds to (6), the right branch corresponds to (7). Two cross arrows in Fig. 4 represent the internal crosstalk introduced into the decoder, i.e., the term $\boldsymbol{\mathcal{W}}_g^{\mathcal{T}} \odot h_g^{\mathcal{T}}(\boldsymbol{\mathcal{T}}_{g-1})$ in (6) and the term $\boldsymbol{\mathcal{W}}_g^{\mathcal{X}} \odot h_g^{\mathcal{X}}(\boldsymbol{\mathcal{X}}_{g-1})$ in (7), respectively. $h_g^{\mathcal{X}}(\cdot)$ and $h_g^{\mathcal{T}}(\cdot)$ represent the fully-connected (FC) layer and batch norm layer at the left branch and the right branch in Fig. 4, respectively. In addition, $f_g^{\mathcal{X}}(\cdot)$ and $f_g^{\mathcal{T}}(\cdot)$ represent activations at the left branch and the right branch, respectively. We separate notations for different branches considering that weights are not shared.

III. Model Establishment

A. System Specification

A CST sensor with 32 beams, as shown in Fig. 5(a), is used in this work to generate four equiangular projections, i.e., $Q = 4$, each with eight equispaced parallel laser beams, i.e., $R = 8$ [26]. The angular spacing between projections is 45°. Neighboring beams within each projection are separated by 1.80 cm. The distance between an emitter and a receiver is 36.76 cm. The RoI is defined as the octagonal sensing area with the side length of 12.60 cm. The dimension of each pixel in the RoI is 0.766 cm × 0.766 cm, resulting in 1924 uniformly segmented pixels, i.e., $N = 1924$. Fig. 5(b) shows its practical implementation.

As a principal product of hydrocarbon combustion, water vapor (H_2O) has a strong near-infrared absorption spectrum and therefore is selected as the target absorption species in this work. Two H_2O transitions at $\upsilon_1 = 7185.6\ \text{cm}^{-1}$ and $\upsilon_2 = 7444.36\ \text{cm}^{-1}$ are adopted given their good temperature sensitivity for the target temperature range of 300–1500 K [3]. Using the 32-beam CST sensor, the projection data,

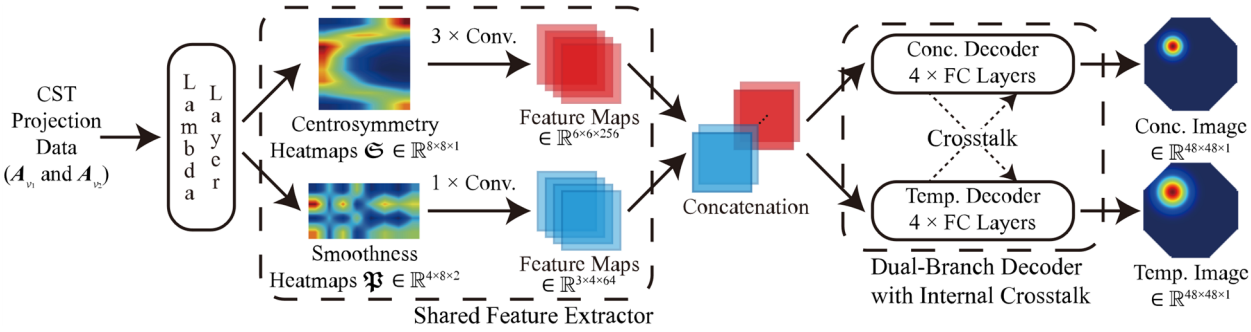


Fig. 3. Overall architecture of CSTNet. Conc.: concentration. Temp.: temperature.

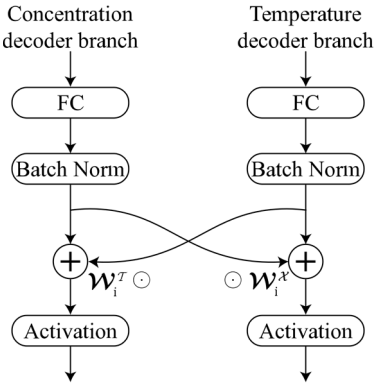
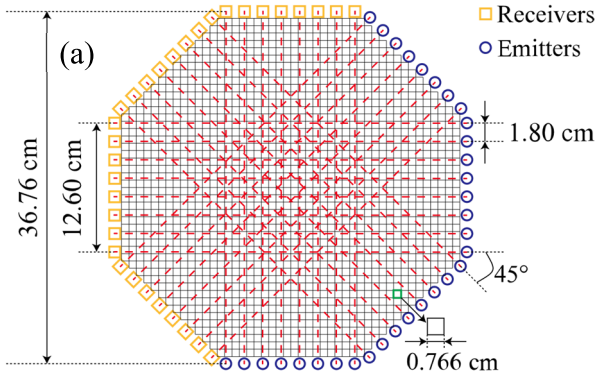


Fig. 4. Stage of crosstalk decoders.



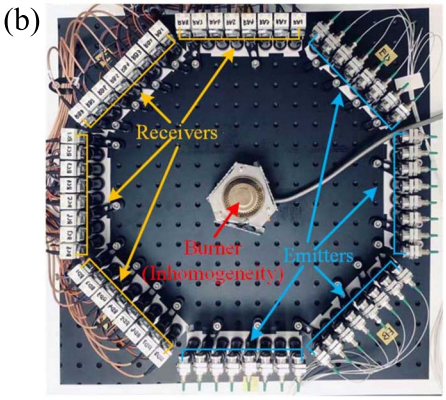


Fig. 5. (a) Schematic layout and (b) photograph of the 32-beam CST sensor.

$\boldsymbol{A}_{v_1} \in \mathbb{R}^{32\times1}$ and $\boldsymbol{A}_{v_2} \in \mathbb{R}^{32\times1}$, are obtained at v_1 and v_2, respectively. The end-to-end CSTNet will be implemented in Section III-B to simultaneously reconstruct the distributions of H_2O concentration and temperature using $\boldsymbol{A}_{v_1}$ and $\boldsymbol{A}_{v_2}$.

B. Implementation of CSTNet

1) Lambda Layer: The lambda layer is used to construct the centrosymmetry heatmap, $\mathfrak{S} \in \mathbb{R}^{8\times8\times1}$, and the smoothness heatmap, $\mathfrak{P} \in \mathbb{R}^{4\times8\times2}$, by combining and rearranging $\boldsymbol{A}_{v_1}$ and $\boldsymbol{A}_{v_2}$ according to Algorithm 1. With the lambda layer, $\boldsymbol{A}_{v_1}$ and $\boldsymbol{A}_{v_2}$ can be directly used as inputs for the shared feature extractor, facilitating end-to-end learning.

2) Shared Feature Extractor: The shared feature extractor takes the output of the lambda layer as inputs and directly learns centrosymmetry from $\mathfrak{S}$ and smoothness from $\mathfrak{P}$ by convolution. Its outputs are shared by two branches in the subsequent decoder.

Table I lists the experimentally determined hyperparameters for our model structure. The shared feature extractor mainly consists of convolutional layers, while the dual-branch decoder mainly consists of FC layers. The first branch that learns the centrosymmetry consists of three convolutional layers using 3×3 filters with strides of 1 for height and width. While inputs to the first two layers are padded by 1 in height and width, input to the last layer is not padded. Each layer adopts PReLU [27] as the activation function. Batch normalization [28] is employed for accelerated training and can largely prevent overfitting. There are two reasons for not adopting pooling layers in CSTNet. First, the use of pooling is detrimental to the pixelwise prediction that we aim to achieve [29]. Second, CSTNet can hardly benefit from the reduced dimensions of feature maps, which are already of small dimensions, by using pooling layers. The forward propagation is computed by

$$\boldsymbol{\mathcal{S}}_i = \begin{cases} \mathfrak{S}, & i = 0 \\ \mathrm{PReLU}_i\left(\mathrm{BN}_{\gamma_i,\beta_i}\left(\boldsymbol{W}_i * \boldsymbol{\mathcal{S}}_{i-1}\right)\right), & i \in \{1,2,3\} \end{cases} \tag{8}$$

where $\boldsymbol{W}_i$ is the convolution kernel for the ith convolution layer, $\boldsymbol{\mathcal{S}}_i$ is the ith intermediate feature maps extracted from $\mathfrak{S}$, $\mathrm{PReLU}_i(\cdot)$ is the PReLU activation function for the ith layer, and $\mathrm{BN}_{\gamma_i,\beta_i}(\cdot)$ is the batch normalization in the ith layer.

TABLE I

DETAILED PARAMETERS FOR IMPLEMENTING CSTNET. THE INPUT DIMENSION OF EACH CONVOLUTIONAL LAYER IS DESCRIBED IN THE FORM OF $H_I \times W_I \times C_I$. THE OUTPUT DIMENSION IS DESCRIBED IN THE FORM OF $H_O \times W_O \times C_O$. THE FILTER SIZES AND STRIDES ARE DESCRIBED IN FORMS OF $H_W \times W_W$ AND (HEIGHT AND WIDTH)

	Shared feature extractor (Convolutional layers)				Dual-branch decoder with internal crosstalk (FC layers)					
	Centro-symmetry extractor	Layer 1	Layer 2	Layer 3	Conc. decoder		Stage 1	Stage 2	Stage 3	Stage 4
Input dim.		8×8×1	8×8×64	8×8×128		Output size	8192	4096	2048	1924
Output dim.		8×8×64	8×8×128	6×6×256						
Filter size		3×3	3×3	3×3						
Stride		(1, 1)	(1, 1)	(1, 1)						
Padding		1	1	0						
	Smoothness extractor	Layer 1			Temp. decoder		Stage 1	Stage 2	Stage 3	Stage 4
Input dim.		4×8×2				Output size	8192	4096	2048	1924
Output dim.		3×4×64								
Filter size		2×2								
Stride		(1, 2)								
Padding		0								

The second branch that learns the smoothness contains a single convolution layer, in which PReLU and batch normalization are adopted as well. To correctly extract the smoothness information, filters in $\boldsymbol{W}$ are intentionally designed with the size of 2×2 and strides of (1, 2) such that they always operate on projection data obtained from adjacent receivers within the same projection angle, i.e., $R = 8$ and, hence, $F = 2$ in Algorithm 1. The forward propagation is computed by

$$\boldsymbol{\mathcal{P}} = \mathrm{PReLU}\big(\mathrm{BN}_{\gamma,\beta}(\boldsymbol{W} * \boldsymbol{\mathfrak{P}})\big) \tag{9}$$

where $\boldsymbol{\mathcal{P}}$ is the extracted feature maps from $\boldsymbol{\mathfrak{P}}$.

As a result, multiscale feature maps, $\boldsymbol{\mathcal{S}}_3 \in \mathbb{R}^{6\times6\times256}$ and $\boldsymbol{\mathcal{P}} \in \mathbb{R}^{3\times4\times64}$, are extracted and further concatenated [30], yielding a vector containing latent features, $\boldsymbol{\mathcal{L}} \in \mathbb{R}^{9984\times1}$.

3) Dual-Branch Decoder With Internal Crosstalk: $\boldsymbol{\mathcal{L}}$ is fed into the dual-branch decoder with crosstalk for simultaneous imaging of H_2O concentration and temperature. In this case, FC layers are employed to fuse $\boldsymbol{\mathcal{L}}$ [31].

With batch normalization and PReLU activation, outputs from each of the first three stages can be computed through

$$\boldsymbol{\mathcal{X}}_i = \begin{cases} \boldsymbol{\mathcal{L}}, & i = 0 \\ \mathrm{PReLU}_i^{\mathcal{X}}\Big(\mathrm{BN}_{\gamma_i^{\mathcal{X}},\beta_i^{\mathcal{X}}}\big(\boldsymbol{W}_i^{\mathcal{X}}\boldsymbol{\mathcal{X}}_{i-1}\big) \\ \qquad +\boldsymbol{\mathcal{W}}_i^{\mathcal{T}} \odot \mathrm{BN}_{\gamma_i^{\mathcal{T}},\beta_i^{\mathcal{T}}}\big(\boldsymbol{W}_i^{\mathcal{T}}\boldsymbol{\mathcal{T}}_{i-1}\big)\Big) \\ \qquad i \in \{1, 2, 3\} \end{cases} \tag{10}$$

and

$$\boldsymbol{\mathcal{T}}_i = \begin{cases} \boldsymbol{\mathcal{L}}, i = 0 \\ \mathrm{PReLU}_i^{\mathcal{T}}\Big(\mathrm{BN}_{\gamma_i^{\mathcal{T}},\beta_i^{\mathcal{T}}}\big(\boldsymbol{W}_i^{\mathcal{T}}\boldsymbol{\mathcal{T}}_{i-1}\big) + \boldsymbol{\mathcal{W}}_i^{\mathcal{X}} \\ \qquad \odot \mathrm{BN}_{\gamma_i^{\mathcal{X}},\beta_i^{\mathcal{X}}}\big(\boldsymbol{W}_i^{\mathcal{X}}\boldsymbol{\mathcal{X}}_{i-1}\big)\Big), \quad i \in \{1, 2, 3\} \end{cases} \tag{11}$$

where $\mathrm{BN}_{\gamma_i^{\mathcal{X}},\beta_i^{\mathcal{X}}}(\cdot)$ and $\mathrm{BN}_{\gamma_i^{\mathcal{T}},\beta_i^{\mathcal{T}}}(\cdot)$ denote batch normalization in H_2O concentration and temperature branches of the ith stage, respectively.

The last stage, i.e., the output stage, is formed by imposing physical constraints on the H_2O concentration and temperature. Hyperbolic tangent function, tanh(·), is adopted as the output activation considering that flow-field parameters are supposed to be within a physically reasonable range. Therefore, the distributions of H_2O concentration and temperature are finally reconstructed by

$$\boldsymbol{X} = \mathrm{Tanh}\Big(\mathrm{BN}_{\gamma_4^{\mathcal{X}},\beta_4^{\mathcal{X}}}\big(\boldsymbol{W}_4^{\mathcal{X}}\boldsymbol{\mathcal{X}}_3\big) + \boldsymbol{\mathcal{W}}_4^{\mathcal{T}} \odot \mathrm{BN}_{\gamma_4^{\mathcal{T}},\beta_4^{\mathcal{T}}}\big(\boldsymbol{W}_4^{\mathcal{T}}\boldsymbol{\mathcal{T}}_3\big)\Big) \tag{12}$$

and

$$\boldsymbol{T} = \mathrm{Tanh}\Big(\mathrm{BN}_{\gamma_4^{\mathcal{T}},\beta_4^{\mathcal{T}}}\big(\boldsymbol{W}_4^{\mathcal{T}}\mathcal{T}_3\big) + \boldsymbol{\mathcal{W}}_4^{\mathcal{X}} \odot \mathrm{BN}_{\gamma_4^{\mathcal{X}},\beta_4^{\mathcal{X}}}\big(\boldsymbol{W}_4^{\mathcal{X}}\boldsymbol{\mathcal{X}}_3\big)\Big). \tag{13}$$

IV. NETWORK TRAINING AND TESTING

A. Dataset

In the state-of-the-art simulation of laminar flames [5], [6], [8], [15], [17], [18], the cross section temperature distribution has been assumed to be a Gaussian profile or the combination of multiple Gaussian profiles. We follow this assumption and use Gaussian profiles to simulate the laminar flames in our study. However, in turbulent flow fields with more complex temperature distributions, for example, swirl combustion in gas turbine [32], the datasets should be established with more reliable and field-oriented simulated data, instead of simple Gaussian profiles. These data can be possibly obtained from computational fluid dynamics (CFD) simulation [33] and more complex measurement modalities, e.g., planar laser-induced fluorescence (PLIF) [34].

In this work, the constructed dataset includes three categories of 2-D distributions of H_2O concentration and temperature with one, two, and three inhomogeneities. Each inhomogeneity is modeled by a 2-D Gaussian profile. As noted in Section II-C, H_2O concentration distribution is generally well correlated with temperature distribution. In each phantom, the peak locations of the inhomogeneities in H_2O concentration distributions are modeled the same as those in the temperature distributions. In general, the spread of H_2O concentration depends on flow convection, which is slower than heat transfer and dissipation. Therefore, the 2-D Gaussian inhomogeneities in the H_2O concentration distributions are generated with smaller standard deviations than those in the

temperature distributions. To be specific, the distributions of H_2O concentration and temperature are mathematically expressed as

$$X(x, y) = X_{\min} + \sum_{d=1}^{D} \xi_d (X_{\max} - X_{\min}) \times \exp\left(-\frac{\left(x - x_c^d\right)^2 + \left(y - y_c^d\right)^2}{\sigma_X^2} \right) \quad (14)$$

and

$$T(x, y) = T_{\min} + \sum_{d=1}^{D} \xi_d (T_{\max} - T_{\min}) \times \exp\left(-\frac{\left(x - x_c^d\right)^2 + \left(y - y_c^d\right)^2}{\sigma_T^2} \right) \quad (15)$$

where x and y denote the horizontal and vertical coordinates of the RoI, respectively, (x_c^d, y_c^d) is the central position of the dth Gaussian profile. D is the total number of inhomogeneities in the phantom, $X_{\max}$ ($T_{\max}$) and $X_{\min}$ ($T_{\min}$) are the maximum and minimum H_2O concentration (temperature), respectively, $\xi_d \sim U(0.7, 1)$ is a random scaling factor, and σ_X and σ_T are the standard deviations of H_2O concentration and temperature inhomogeneities that satisfy $\sigma_X = \rho\sigma_T$ with $\rho \sim U(1/3, 1)$, respectively.

In this work, we adopt $X_{\min} = 0.01$, $X_{\max} = 0.12$, $T_{\min} = 318$ K, and $T_{\max} = 1300$ K. The dataset is generated with 19 305 independent examples, which are then randomly divided into a training set with 13 440 examples, a validation set with 5760 examples, and a test set with 105 examples. Using the H_2O transitions at $\upsilon_1 = 7185.6\ \text{cm}^{-1}$ and $\upsilon_2 = 7444.36\ \text{cm}^{-1}$, the training and validation sets containing 19 200 sets of $\boldsymbol{A}_{\upsilon_1}$ and $\boldsymbol{A}_{\upsilon_2}$ are generated according to (3) without noise contamination. Six testing sets each with 105 examples are generated on six different levels of noise, with signal-to-noise ratio (SNR) ranging from 20 dB (high noise in CST) to 45 dB (low noise in CST).

Subsequently, path integrated absorbance for the jth beam in the training and validation sets is standardized. The process of standardization has two benefits. First, it can speed up the training process since the averages of input features are moved close to zero and their covariances are kept approximately the same, which balances out the learning speed of weights connected to input nodes [16]. Second, standardization can suppress generalization error during image reconstruction, which is caused by bias and fluctuations in real CST measurements.

B. Training Details

CSTNet is trained through empirical risk minimization (ERM) and structural risk minimization (SRM). ERM corresponds to minimizing a weighted mean-square-error (MSE) loss in terms of the reconstructed and true distributions of H_2O concentration and temperature. SRM corresponds to minimizing an additional L2 regularization term. Hence, the objective function is defined as

$$\min \tau \frac{1}{N} \sum_{k=1}^{N} \left(T_k - \widehat{T_k}\right)^2 + (1 - \tau) \frac{1}{N} \sum_{k=1}^{N} \left(X_k - \widehat{X_k}\right)^2 + \lambda \|\psi\|_2^2 \quad (16)$$

where τ is a hyperparameter to trade off the MSE loss on imaging of H_2O concentration and temperature, $\hat{X}_k$ ($\hat{T}_k$) and X_k (T_k) are the reconstructed and true H_2O concentration (temperature) in the kth pixel, respectively, λ is the weight for L2 weight decay, $||\cdot||_2$ represents the L2 norm, and ψ denotes the trainable weights of the proposed neural network.

Since the value of MSE is the expected value of the squared error between ground truths and the reconstructions, the minimization of MSE loss leads to image reconstruction with higher accuracy. Furthermore, the L2 penalty term serves as a regularization, preventing the proposed model from overfitting.

Adam optimizer [35] is employed with a value of 10^{-3} for learning rate found by the range test [36] and default values for other hyperparameters. λ is set to 2×10^{-6} determined by the Monte Carlo estimation [37]. τ is set to 0.5 to maintain the balance between H_2O concentration and temperature imaging since both have been processed to have the same scale. Training our model takes approximately an hour using a single NVIDIA Tesla P100 graphics processing unit (GPU) with 16-GB memory. It is trained for 350 epochs until convergence.

C. Test Results

The established CSTNet is trained for three times with different randomness, yielding an ensemble of three different sets of model weights, i.e., $\boldsymbol{\Psi} = \{\psi_1, \psi_2, \psi_3\}$. Furthermore, we compare our method with two state-of-the-art methods. One is the model used in [17] and [18], denoted as HuangNet, which directly takes CST measurements as input without adopting any domain priors in CST. The other is the ELM-based method used in [15], denoted as ELM. In addition, to test the contribution of the dual-branch decoder and the internal crosstalk, we introduce two ablative versions of CSTNet, denoted as CSTNet (single), each with a single-branch decoder either for concentration or temperature reconstructions. The same training sets are used for all models. Note that HuangNet, ELM, and the ablative CSTNet (single) have to be trained twice using different labels (either true H_2O concentration distributions or temperature distributions) to retrieve these two flow-field parameters.

The proposed CSTNet is compared with these methods by computing the image errors of H_2O concentration and temperature imaging at different SNRs

$$\text{IE}_{\text{conc.}} = \frac{1}{H} \sum_{h=1}^{H} \frac{\left\| \boldsymbol{X}_h - \widehat{\boldsymbol{X}_h} \right\|_2}{\left\| \widehat{\boldsymbol{X}_h} \right\|_2} \quad (17)$$

and

$$\text{IE}_{\text{temp.}} = \frac{1}{H} \sum_{h=1}^{H} \frac{\left\| \boldsymbol{T}_h - \widehat{\boldsymbol{T}_h} \right\|_2}{\left\| \widehat{\boldsymbol{T}_h} \right\|_2} \quad (18)$$

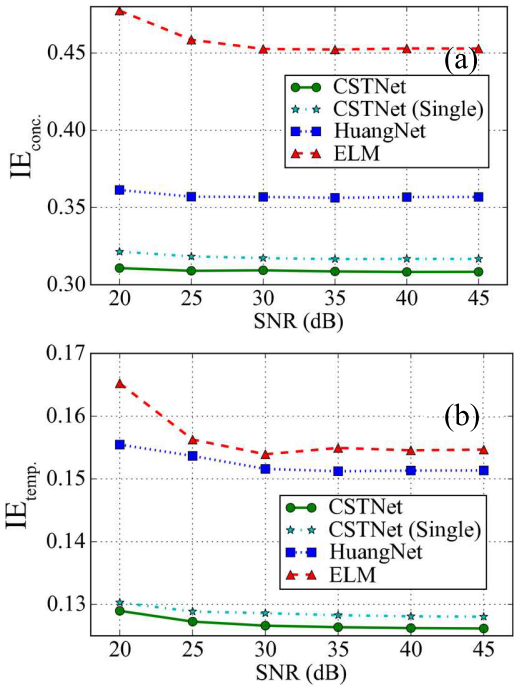


Fig. 6. Comparison between the proposed CSTNet, an ablative version with a single-branch decoder, denoted as CSTNet (single), and two state-of-the-art methods, i.e., HuangNet and ELM. (a) and (b) Image errors for the reconstructed distributions of H_2O concentration ($IE_{conc.}$) and temperature ($IE_{temp.}$) at different SNRs, respectively.

where h and H denote the index and the total number of test examples, respectively, and $\boldsymbol{X}_h$ ($\boldsymbol{T}_h$) and $\widehat{\boldsymbol{X}}_h$ ($\widehat{\boldsymbol{T}}_h$) are the hth true and reconstructed H_2O concentration (temperature) distributions, respectively. The proposed CSTNet can simultaneously reconstruct the two images, ($\widehat{\boldsymbol{X}}_h$, $\widehat{\boldsymbol{T}}_h$). However, HuangNet, ELM, and the ablative CSTNet (single) have to reconstruct them one after another, resulting in doubled inference cost.

Fig. 6(a) and (b) shows the image errors for the reconstruction of H_2O concentration ($IE_{conc.}$) and temperature ($IE_{temp.}$), respectively. The x-axis represents different levels of noise ranging from high noise (SNR = 20 dB) to low noise (SNR = 45 dB). Both $IE_{conc.}$ and $IE_{temp.}$ decrease as SNR increases. The proposed CSTNet is always the best at all levels of noise, while ELM is the worst. With a practical SNR in real applications at approximately 35 dB, $IE_{conc.}$ and $IE_{temp.}$ for CSTNet are 0.3087 and 0.1264, respectively, which are lower than those obtained using the previous state-of-the-art methods, e.g., $IE_{conc.}$ = 0.3563 and $IE_{temp.}$ = 0.1512 for HuangNet and $IE_{conc.}$ = 0.4523 and $IE_{temp.}$ = 0.1549 for ELM, demonstrating the capability of achieving high-fidelity reconstruction and the accurate retrieval of the true images using the proposed CSTNet. Furthermore, $IE_{conc.}$ and $IE_{temp.}$ for CSTNet increase by 0.79% and 2.21% when the SNR varies from 45 to 20 dB. In contrast, for HuangNet and ELM, $IE_{conc.}$ increases by 1.41% and 5.58%, while $IE_{temp.}$ increases by 2.8% and 7.35%, respectively. Therefore, the proposed CSTNet is also more robust for noise-contaminated measurements.

In the ablation study, at the same noise level (SNR = 35 dB), $IE_{conc.}$ and $IE_{temp.}$ for CSTNet (single) are 0.3167 and 0.1283, respectively. Compared with HuangNet that does not consider any domain priors in CST, the lower image errors demonstrate the benefits of incorporating domain priors in CSTNet. Furthermore, $IE_{conc.}$ and $IE_{temp.}$ obtained using the full CSTNet model with dual-branch decoder are lower than those obtained using the CSTNet (single), denoting that correlations between multiple flow-field parameters, e.g., temperature and gas concentration, are helpful for better-fidelity retrieval. The correlations are learned through the internal crosstalk introduced by the dual-branch decoder.

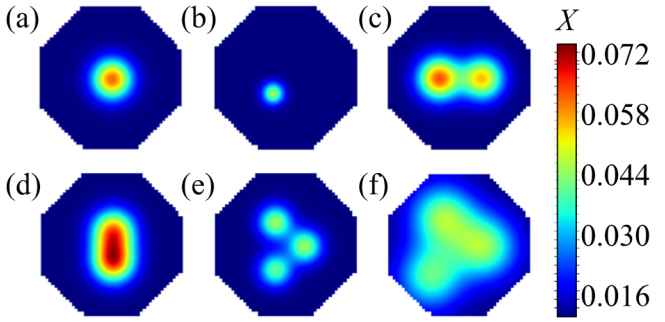


Fig. 7. True H_2O concentration distributions with (a) and (b) one inhomogeneity, (c) and (d) two inhomogeneities, and (e) and (f) three inhomogeneities.

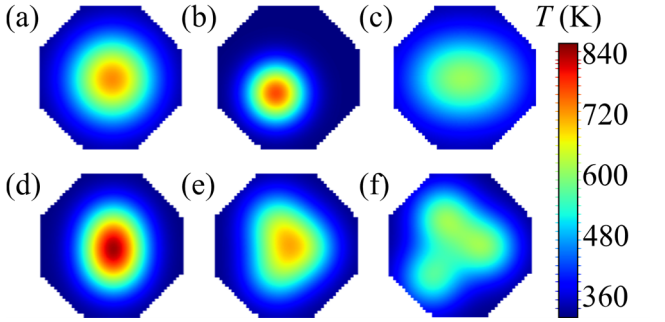


Fig. 8. True temperature distributions with (a) and (b) one inhomogeneity, (c) and (d) two inhomogeneities, and (e) and (f) three inhomogeneities.

Six representative results from CSTNet with one, two, and three inhomogeneities are selected. As shown in Figs. 7 and 8, phantoms (a) and (b) include single inhomogeneity with different sizes and locations. Figs. 7 and 8(c)–(f) have more inhomogeneities with different orientations, mutual distances, and sizes. As shown in Figs. 9 and 10, the reconstructions of these phantoms indicate that the trained CSTNet can clearly distinguish the number of inhomogeneities, precisely locate the inhomogeneities, and accurately retrieve the true images. The proposed CSTNet contributes to supremely good quality of the tomographic images with no artifacts. Using compute unified device architecture (CUDA) [38], the well-trained CSTNet achieves simultaneous imaging with an average frame rate of 3134 frames/s, providing great potential for speedy and real-time multiparameter imaging in industrial applications.

V. Experiments

Laboratory experiments were carried out to further validate the proposed CSTNet model. The CST sensor was built in the optical layout shown in Fig. 5. More details of the optics, e.g., lasers and detectors, electronics, e.g., data acquisition and

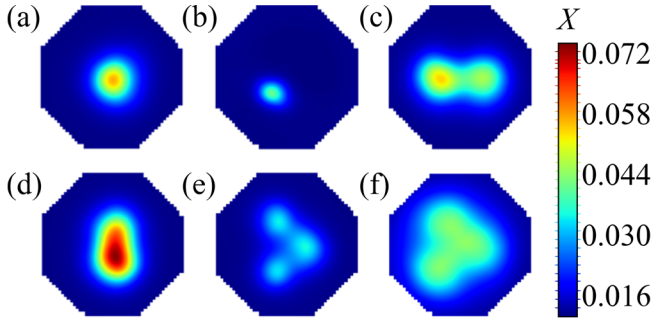


Fig. 9. Reconstructed H_2O concentration distributions under an SNR of 35 dB for the corresponding phantoms shown in Fig. 7 with (a) and (b) one inhomogeneity, (c) and (d) two inhomogeneities, and (e) and (f) three inhomogeneities.

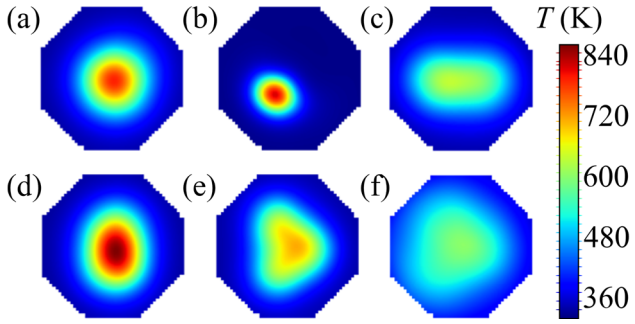


Fig. 10. Reconstructed temperature distributions under an SNR of 35 dB for the corresponding phantoms shown in Fig. 8 with (a) and (b) one inhomogeneity, (c) and (d) two inhomogeneities, and (e) and (f) three inhomogeneities.

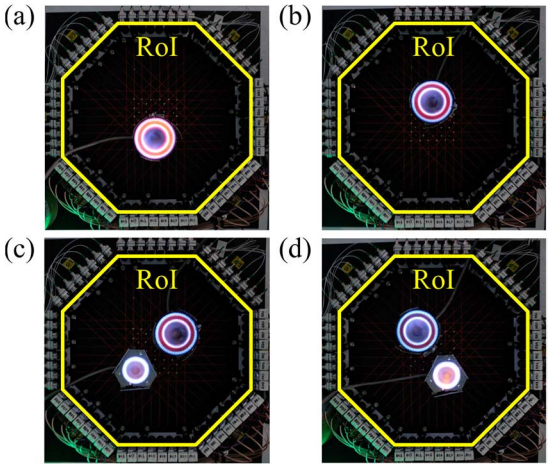


Fig. 11. Four reactive flow fields generated in the experiments with (a) and (b) single flame and (c) and (d) two flames with different sizes.

signal processing system, and the parameter settings in wavelength modulation spectroscopy have been described in [39].

As shown in Fig. 11, four cases with different distributions of H_2O concentration and temperature are demonstrated in the experiments. The first two cases, shown in Fig. 11(a) and (b), contain a single flame located at the lower center and upper center of the RoI, respectively. To consider more complex phantoms, two flames with different sizes and locations are generated in the other two cases shown in Fig. 11(c) and (d).

The tomographic images of H_2O concentration and temperature for the four cases are shown in Figs. 12 and 13, respectively. Retrieved peak values of the inhomogeneities in

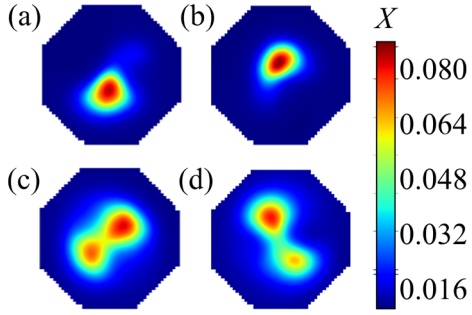


Fig. 12. Reconstructed H_2O concentration distributions for the four cases in Fig. 11 with (a) and (b) single flame and (c) and (d) two flames with different sizes.

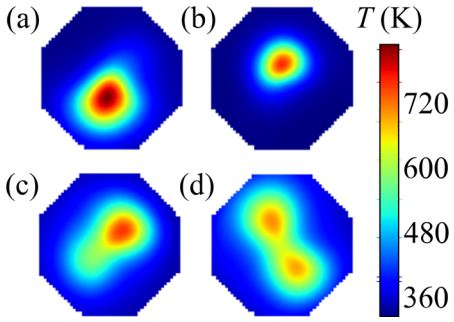


Fig. 13. Reconstructed temperature distributions for the four cases in Fig. 11 with (a) and (b) single flame and (c) and (d) two flames with different sizes.

TABLE II

RETRIEVED PEAK VALUES FOR H_2O CONCENTRATION AND TEMPERATURE IMAGING. CONC.: H_2O CONCENTRATION. TEMP.: TEMPERATURE. L.: LARGER FLAME. S.: SMALLER FLAME

	Peak conc.	Peak temp. (K)		Peak conc.		Peak temp. (K)	
				L.	S.	L.	S.
(a)	0.084	828	(c)	0.082	0.073	755	600
(b)	0.088	761	(d)	0.080	0.065	708	695

the tomographic images are listed in Table II. For the single-flame cases, locations of the flames in the tomographic images reconstructed by CSTNet agree well with original ones. As the same flame is used in both cases, similar retrieved peak values indicate good consistency between the reconstruction and the truth. For the dual-flame cases, the reconstructions not only precisely localize the two inhomogeneities but also reveal their relative sizes. Last but not least, artifacts are significantly limited in all cases, demonstrating that the proposed CSTNet is strongly robust for image reconstruction even with a severely limited number of laser beams. The results are very promising for industry-oriented CST, mostly implemented in harsh environments, with low-complexity optical sensors. Aided by GPUs and CUDA [38], CSTNet completes high-fidelity imaging within several milliseconds. The very short inference time enables online CST and further stimulates the industrial implementation of the proposed CSTNet on embedded devices, such as field-programmable gate arrays (FPGAs) [40], [41].

VI. CONCLUSION

In this article, we developed a novel CNN named CSTNet for simultaneous imaging of the distributions of species

concentration and temperature in reactive flows using CST. The inherently physical characteristics of the CST are learned by a feature extractor, which incorporates the domain *a priori* information of smoothness and centrosymmetry. To simultaneously reconstruct the distributions of species concentration and temperature, a dual-branch decoder with internal crosstalk is designed in CSTNet. It considers the natural correlation between species concentration and temperature, enabling more reasonable and efficient retrievals in practical combustion processes. The proposed CSTNet is both analytically and experimentally proven to be successful in high-fidelity imaging of H_2O concentration and temperature images using two H_2O transitions and a tomographic sensor with 32 laser beams.

The performance of CSTNet is first evaluated using simulated test sets. Given a measurement SNR of 35 dB, CSTNet can accurately reconstruct various distributions of H_2O concentration and temperature, outperforming state-of-the-art methods HuangNet and ELM by 13% (16%) and 32% (18%) for the reconstructions of H_2O concentration (temperature), respectively. For a wide range of SNR, numerical results indicate that CSTNet maintains excellent robustness against measurement noise. In the lab-scale experiments, image reconstruction using CSTNet achieves good agreement with the known locations of the original flames. The artifacts in the tomographic images are significantly eliminated, denoting strong resistance to the measurement noise in practical applications. Benefiting from GPU acceleration, the proposed CSTNet can simultaneously reconstruct images of H_2O concentration and temperature distributions with 3134 frames/s, exhibiting great potential for online CST toward real-time process control.

To the best of our knowledge, this is the first experimental application of deep learning in CST using an optical sensor with a severely limited number of laser beams. In practice, the low-complexity optical sensor is overwhelmingly preferred in order to maintain the integrity of the industrial reactors and chambers. In our future work, we will also deploy CSTNet on embedded devices such as FPGAs for the sake of more cost-efficient computation in industrial applications.

REFERENCES

[1] N. Terzija, S. Karagiannopoulos, S. Begg, P. Wright, K. Ozanyan, and H. McCann, "Tomographic imaging of the liquid and vapour fuel distributions in a single-cylinder direct-injection gasoline engine," *Int. J. Engine Res.*, vol. 16, no. 4, pp. 565–579, 2015.

[2] L. Ma *et al.*, "50-kHz-rate 2D imaging of temperature and H_2O concentration at the exhaust plane of a J85 engine using hyperspectral tomography," *Opt. Exp.*, vol. 21, no. 1, pp. 1152–1162, 2013.

[3] C. Liu, Z. Cao, Y. Lin, L. Xu, and H. McCann, "Online cross-sectional monitoring of a swirling flame using TDLAS tomography," *IEEE Trans. Instrum. Meas.*, vol. 67, no. 6, pp. 1338–1348, Jun. 2018.

[4] Q. Qu, Z. Cao, L. Xu, C. Liu, L. Chang, and H. McCann, "Reconstruction of two-dimensional velocity distribution in scramjet by laser absorption spectroscopy tomography," *Appl. Opt.*, vol. 58, no. 1, pp. 205–212, 2019.

[5] H. Xia *et al.*, "Two-step tomographic reconstructions of temperature and species concentration in a flame based on laser absorption measurements with a rotation platform," *Opt. Lasers Eng.*, vol. 90, pp. 10–18, Mar. 2017.

[6] C. Liu, L. Xu, J. Chen, Z. Cao, Y. Lin, and W. Cai, "Development of a fan-beam TDLAS-based tomographic sensor for rapid imaging of temperature and gas concentration," *Opt. Exp.*, vol. 23, no. 17, pp. 22494–22511, 2015.

[7] K. J. Daun, "Infrared species limited data tomography through Tikhonov reconstruction," *J. Quant. Spectrosc. Radiat. Transf.*, vol. 111, no. 1, pp. 105–115, Jan. 2010.

[8] L. Ma *et al.*, "Tomographic imaging of temperature and chemical species based on hyperspectral absorption spectroscopy," *Opt. Exp.*, vol. 17, no. 10, pp. 8602–8613, 2009.

[9] S. J. Grauer, P. J. Hadwin, and K. J. Daun, "Improving chemical species tomography of turbulent flows using covariance estimation," *Appl. Opt.*, vol. 56, no. 13, pp. 3900–3912, 2017.

[10] N. Polydorides *et al.*, "Constrained models for optical absorption tomography," *Appl. Opt.*, vol. 57, no. 7, pp. B1–B9, 2018.

[11] C. Liu and L. Xu, "Laser absorption spectroscopy for combustion diagnosis in reactive flows: A review," *Appl. Spectrosc. Rev.*, vol. 54, no. 1, pp. 1–44, Apr. 2019.

[12] D. Nie, L. Wang, Y. Gao, J. Lian, and D. Shen, "STRAINet: Spatially varying sTochastic residual AdversarIal networks for MRI pelvic organ segmentation," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 5, pp. 1552–1564, May 2019.

[13] S. Sakhavi, C. Guan, and S. Yan, "Learning temporal information for brain-computer interface using convolutional neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 11, pp. 5619–5629, Nov. 2018.

[14] F. Xing, Y. Xie, H. Su, F. Liu, and L. Yang, "Deep learning in microscopy image analysis: A survey," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 10, pp. 4550–4568, Oct. 2018.

[15] T. Yu, W. Cai, and Y. Liu, "Rapid tomographic reconstruction based on machine learning for time-resolved combustion diagnostics," *Rev. Scientific Instrum.*, vol. 89, no. 4, Apr. 2018, Art. no. 043101.

[16] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 44–436, May 2015.

[17] J. Huang, H. Liu, J. Dai, and W. Cai, "Reconstruction for limited-data nonlinear tomographic absorption spectroscopy via deep learning," *J. Quant. Spectrosc. Radiat. Transf.*, vol. 218, pp. 187–193, Oct. 2018.

[18] J. Huang, J. Zhao, and W. Cai, "Compressing convolutional neural networks using POD for the reconstruction of nonlinear tomographic absorption spectroscopy," *Comput. Phys. Commun.*, vol. 241, pp. 33–39, Aug. 2019.

[19] C. Wei, K. K. Schwarm, D. I. Pineda, and R. M. Spearrin, "Deep neural network inversion for 3D laser absorption imaging of methane in reacting flows," *Opt. Lett.*, vol. 45, no. 8, pp. 2447–2450, 2020.

[20] T. Dreier, R. Chrystie, T. Endres, and C. Schulz, "Laser-based combustion diagnostics," in *Encyclopedia of Analytical Chemistry*. Hoboken, NJ, USA: Wiley, 2016, pp. 1–44.

[21] C. S. Goldenstein, R. M. Spearrin, J. B. Jeffries, and R. K. Hanson, "Infrared laser-absorption sensing for combustion gases," *Prog. Energy Combustion Sci.*, vol. 60, pp. 132–176, May 2017.

[22] H. McCann, P. Wright, and K. Daun, "5—Chemical species tomography," in *Industrial Tomography*, M. Wang, Ed. Cambridge, U.K.: Woodhead Publishing, 2015, pp. 135–174.

[23] W. Cai and C. F. Kaminski, "Tomographic absorption spectroscopy for the study of gas dynamics and reactive flows," *Prog. Energy Combustion Sci.*, vol. 59, pp. 1–31, Mar. 2017.

[24] R. Dian, S. Li, A. Guo, and L. Fang, "Deep hyperspectral image sharpening," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 11, pp. 5345–5355, Nov. 2018.

[25] H. Zhu, L. Jiao, W. Ma, F. Liu, and W. Zhao, "A novel neural network for remote sensing image matching," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 9, pp. 2853–2865, Sep. 2019.

[26] C. Liu, S.-A. Tsekenis, N. Polydorides, and H. McCann, "Toward customized spatial resolution in TDLAS tomography," *IEEE Sensors J.*, vol. 19, no. 5, pp. 1748–1755, Mar. 2019.

[27] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 1026–1034.

[28] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. Int. Conf. Mach. Learn.*, Lille, France, vol. 37, Jul. 2015, pp. 448–456.

[29] J. Liu, M. Gong, K. Qin, and P. Zhang, "A deep convolutional coupling network for change detection based on heterogeneous optical and radar images," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 3, pp. 545–559, Dec. 2016.

[30] K.-H. Shih, C.-T. Chiu, J.-A. Lin, and Y.-Y. Bu, "Real-time object detection with reduced region proposal network via multi-feature concatenation," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 6, pp. 2164–2173, Jun. 2020.

[31] Z. Gao *et al.*, "EEG-based Spatio–Temporal convolutional neural network for driver fatigue evaluation," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 9, pp. 2755–2763, Sep. 2019.
[32] P. Weigand, W. Meier, X. R. Duan, W. Stricker, and M. Aigner, "Investigations of swirl flames in a gas turbine model combustor: I. Flow field, structures, temperature, and species distributions," *Combust. Flame*, vol. 144, pp. 205–224, 2006.
[33] Y. Huang and V. Yang, "Dynamics and stability of lean-premixed swirl-stabilized combustion," *Prog. Energy Combustion Sci.*, vol. 35, no. 4, pp. 293–364, 2009.
[34] D. R. Richardson, N. Jiang, D. L. Blunck, J. R. Gord, and S. Roy, "Characterization of inverse diffusion flames in vitiated cross flows via two-photon planar laser-induced fluorescence of CO and 2-D thermometry," *Combustion Flame*, vol. 168, pp. 270–285, Jun. 2016.
[35] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Represent.*, 2014, pp. 1–15.
[36] L. N. Smith, "Cyclical learning rates for training neural networks," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Santa Rosa, CA, USA, Mar. 2017, pp. 464–472.
[37] T. S. Rögnvaldsson, "A simple trick for estimating the weight decay parameter," in *Neural Networks: Tricks of the Trade*. Berlin, Germany: Springer, 1998, pp. 71–92.
[38] J. Nickolls, I. Buck, M. Garland, and K. Skadron, "Scalable parallel programming with CUDA: Is CUDA the parallel programming model that application developers have been waiting for?" *Queue*, vol. 6, no. 2, pp. 40–53, Mar. 2008.
[39] Y. Bao *et al.*, "Relative entropy regularized TDLAS tomography for robust temperature imaging," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–9, 2021.
[40] S. I. Venieris and C.-S. Bouganis, "FpgaConvNet: Mapping regular and irregular convolutional neural networks on FPGAs," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 2, pp. 326–342, Feb. 2019.
[41] N. Shah, P. Chaudhari, and K. Varghese, "Runtime programmable and memory bandwidth optimized FPGA-based coprocessor for deep convolutional neural network," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 12, pp. 5922–5934, Dec. 2018.

Yunfan Jiang received the B.Eng. degree (Hons.) in electronics and electrical engineering from the College of Science and Engineering, The University of Edinburgh, Edinburgh, U.K., in 2020.

During his undergraduate research, he focused on developing line-of-sight laser imaging systems empowered by deep learning.

Jingjing Si (Member, IEEE) received the B.S. degree in electronic engineering and the M.S. degree in communication and information system from Yanshan University, Hebei, China, in 2002 and 2005, respectively, and the Ph.D. degree in communication and information system from the Beijing University of Posts and Telecommunications, Beijing, China, in 2010.

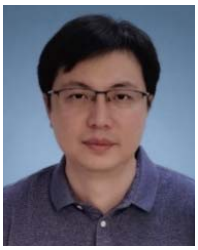
Since 2002, she has been with the School of Information Science and Engineering, Yanshan University, where she is currently a Professor. Her current research interests include multimedia signal processing, inverse problem, and optimization.

Rui Zhang received the B.Eng. degree (Hons.) in electronics and electrical engineering with management from The University of Edinburgh, Edinburgh, U.K., in 2019, and the B.Eng. degree (Hons.) in Internet of Things engineering from Tianjin University, Tianjin, China, in 2019. She is currently pursuing the Ph.D. degree with the Agile Tomography Group, School of Engineering, The University of Edinburgh.

Her current research interest is focused on laser absorption spectroscopic tomography with a customized spatial resolution for combustion diagnosis.

Godwin Enemali (Member, IEEE) received the B.Eng. degree (Hons.) in electrical and electronics engineering from the University of Agriculture, Makurdi, Nigeria, in 2010, and the M.Sc. and Ph.D. degrees in electronics engineering from The University of Edinburgh, Edinburgh, U.K., in 2014 and 2019, respectively.

Since 2019, he has been a Research Associate with the Agile Tomography Group, Institute for Digital Communication, The University of Edinburgh. His current research focuses on high-speed data acquisition systems, reliable field-programmable gate array (FPGA)-based system implementation, laser absorption spectroscopy (LAS), and chemical species tomography.



Bin Zhou received the B.Sc. and Ph.D. degrees from Southeast University, Nanjing, China, in 2002 and 2008, respectively.

He is currently a Professor at Southeast University, where he is also the Vice Dean of the Jiang Bei Innovation Center. His current research interests include NIR/MIR laser absorption spectroscopy (LAS), electrical capacitance tomography (ECT), and their applications to combustion diagnosis and industrial application.

Hugh McCann received the B.Sc. and Ph.D. degrees from the University of Glasgow, Glasgow, U.K., in 1976 and 1980, respectively.

He was a Professor of industrial tomography at The University of Manchester, Manchester, U.K., from 1996 to 2013, following ten years in research and development at the Royal Dutch/Shell Group. He was the Head of the School of Electrical and Electronic Engineering, The University of Manchester, from 1999 to 2002, and the Chair of the U.K. Professors and the Head of Electrical Engineering from 2003 to 2005. In 2018, he was appointed as an Honorary Professor at Beihang University, Beijing, China. He is currently a Professor of tomographic imaging at The University of Edinburgh, Edinburgh, U.K., where he was the Head of the School of Engineering from 2013 to 2018. He was involved in high-energy particle physics for ten years at Glasgow, Manchester, CERN, Geneva, Switzerland, and DESY, Hamburg, Germany. He has extended industrial tomography to provide specific chemical contrast using high-speed all-optoelectronic techniques and has developed electrical impedance tomography for medical applications, collaborating intensively with users in both academia and industry.

Prof. McCann was elected as a fellow of the Royal Academy of Engineering in 2009 and the Royal Society of Edinburgh in 2015.

Chang Liu (Senior Member, IEEE) received the B.Sc. degree in automation from Tianjin University, Tianjin, China, in 2010, and the Ph.D. degree in testing, measurement technology and instrument from Beihang University, Beijing, China, in 2016.

From April 2016 to January 2018, he was a Post-Doctoral Researcher with the Department of Air Pollution and Environmental Technology, Empa-Swiss Federal Laboratories for Materials Science and Technology, Dübendorf, Switzerland. Since February 2018, he has been a Lecturer at the School of Engineering, The University of Edinburgh, Edinburgh, U.K. His current research interests include NIR/MIR laser absorption spectroscopy (LAS), active/passive optical tomography techniques and system design, and their applications to combustion diagnosis and environmental monitoring.