3D-CNN and Autoencoder-Based Gas Detection in Hyperspectral Images

Okan Bilge Özdemir ⬤ and Alper Koz ⬤

Abstract—The detection of gas emission levels is a crucial problem for ecology and human health. Hyperspectral image analysis offers many advantages over traditional gas detection systems with its detection capability from safe distances. Observing that the existing hyperspectral gas detection methods in the thermal range neglect the fact that the captured radiance in the longwave infrared (LWIR) spectrum is better modeled as a mixture of the radiance of background and target gases, we propose a deep learning-based hyperspectral gas detection method in this article, which combines unmixing and classification. The proposed method first converts the radiance data to luminance-temperature data. Then, a 3-D convolutional neural network (CNN) and autoencoder-based network, which is specially designed for unmixing, is applied to the resulting data to acquire abundances and endmembers for each pixel. Finally, the detection is achieved by a three-layer fully connected network to detect the target gases at each pixel based on the extracted endmember spectra and abundance values. The superior performance of the proposed method with respect to the conventional hyperspectral gas detection methods using spectral angle mapper and adaptive cosine estimator is verified with LWIR hyperspectral images including methane and sulfur dioxide gases. In addition, the ablation study with respect to different combinations of the proposed structure including direct classification and unmixing methods has revealed the contribution of the proposed system.

Index Terms—Autoencoders, convolutional neural networks (CNNs), gas detection, hyperspectral unmixing.

I. INTRODUCTION

IMAGING spectroscopy has been used by physicists and chemists for more than three decades to identify materials and their compositions. The concept of hyperspectral remote sensing started in the mid-80s and has been widely used by geologists for mapping minerals to this day [1]. The detectability of the material is determined depending on the spectral range of the spectrometer, its spectral resolution, the abundance of the material, and the strength of the absorption properties in the measured wavelength region [2].

The gas leaks in particular in developed countries in the last decade were one of the crucial environmental problems. Some gases are harmful to the environment and contribute to global warming. They present both short-term risks such as explosions and long-term risks such as cancer to workers or people living close to the leaking facility. To minimize these effects, environmental authorities need to monitor chemical and industrial plants to control gas emission levels. Infrared remote sensing technology, which offers many advantages over traditional gas detection systems, is one of the proposed solutions for this aim as such solutions allow monitoring the scene from a safe distance [3].

To this end, forward-looking infrared hyperspectral cameras are placed in potentially dangerous areas for gas detection from safe distances. These cameras, which are designed to capture images at different wavelengths, can operate in two different regions, which involve medium-wave infrared (3–5 μm) and long-wave infrared (7–14 μm) bands. Until now, these cameras have been utilized for the detection of different gases such as carbon dioxide, propane, methane, sulfur, butane, freon, ammonia, difluoroethane, diethyl ether, sulfur hexafluoride, and phosgene [4], [5], [6], [7]. The detection of gases in such studies is mainly achieved by utilizing conventional statistical detection methods along with the basic signal processing operations such as data transformation, background suppression, dimension reduction, linear regression, and matched filtering [4], [6], [7], [8], [9].

As one of the pioneer studies for gas detection, Pogorzala [10] proposed a pixel-based method using linear regression in synthetic images for the detection of ammonia (NH3) and Freon-114. Later, Vallières et al. [4] presented a method that first converts the hyperspectral radiance data to luminance temperature data. After performing background removal on the temperature data, the resulting cube undergoes spectral matched filtering [11] to distinguish gas-containing pixels. Finally, the detection is carried out by applying thresholding to the resulting scores after matched filtering. In another study, Spisz et al. [12] first applied principal component analysis for background removal, and then utilized matched filter and spectral angle mapper to detect various chemical compounds. A different study using hyperspectral imaging [13] focused on the automatic detection of waste gases. The proposed method first filters the possible areas in the scene by means of detecting critical wavelengths and using the correlation coefficient metrics to select pixels with high concentration. The target gases are then detected using a spectral matched filter algorithm on the selected pixels.

While the presented studies mainly utilize background information, Hirsch and Agassi [14] presented an algorithm for gas detection without requiring background information. The method first applies K-Means segmentation and performs a

spectral analysis on each segment. The final decision is made by applying thresholding on the correlation between the calculated information and the target gas signature. Another method of performing thresholding in the ultimate stage of gas detection is proposed by Kastek and Piątkowski to reveal the gases in turbulent stack fumes [15]. In particular, a spectral angle mapper is utilized to reveal the correlation between the pixel spectra and the signature of the target gas. Later, Kuflik and Rotman [16] conducted a study aiming to find the minimum number of bands required for gas detection with synthetically generated data. The proposed methods by Sabbah et al. [17] and Safak et al. [18] follow similar strategies. First, they convert the raw data into luminance temperature data. Then, the correlation between the pixel spectra and the signature of the target gas is computed and thresholded for detection. In another study presented by Theiler and Love [19], satellite images are employed for the detection of NO2 plumes and SO2 plumes rather than the captured hyperspectral data from in-scene sensors.

Other than the correlation and thresholding-based conventional methods, the new trend in gas detection is to exploit the accumulated experience in deep learning-based detection. One of the few examples of these studies proposed by Kim et al. demonstrated the performances of classification-based deep neural networks and convolutional neural networks (CNNs) for different gases [20]. Among these studies, Zhang et al. [21] developed a classification-based method using a CNN for the detection of CO2. Kumar et al. [22] utilized an region-based CNN (RCNN) structure, rather than CNN, in their work to detect Methane gas plume emissions. The authors stated that their hyperspectral mask-RCNN (H-MCRNN) method is suitable for the rapid scanning of large areas. Finally, Gu [23] performed a detailed study on the hyperparameter optimization of the H-MCRNN method for the detection of methane gas.

In addition to these studies, there are also studies using hyperspectral unmixing for gas detection. In the study presented by Henrot et al. [24], they demonstrated the performance of the hyperspectral unmixing-based model using the linear mixture model by using synthetic and real-time series of hyperspectral images. In another study, Shi et al. [25] presented a method based on hyperspectral unmixing where low-degree mixed pixels in the hyperspectral image were also used by using the sparse greedy algorithm. In a similar study, Tochon et al. [26] performed gas plume detection and tracking from hyperspectral video sequences using hyperspectral unmixing. A linear mixture model was used in the study to extract information about the concentration of the gas as well as the location information. Finally, Fiscante et al. [27] proposed a method based on unsupervised sparse unmixing for the detection of sulfur dioxide.

When these studies are examined, one of the common processes for gas detection is the conversion of radiance data in the thermal domain to the luminance temperature values. The underlying reason for such a process is that the gases become visible due to the emission and/or absorption when there is a temperature contrast between the background and gas components in the scene. In addition, almost all of the presented methods utilize a thresholding operation in the final stage of the processing chain for the ultimate detection. This thresholding operation on the other hand is dependent on the statistical characteristics of the utilized data, which in turn makes the adaptation of these algorithms very challenging for a generic application. It is thus essential to develop methods which do not require changes in the threshold values with respect to the changes in the data for the ultimate decision.

Another important aspect of the previous hyperspectral gas detection methods is to work directly on the resulting data after the luminance temperature conversion. However, the acquired radiance data in the thermal LWIR range is always a mixture of the thermal radiation of the background and the gas molecules in the air. Therefore, the handling of the hyperspectral gas detection in the LWIR spectrum as a combined problem of unmixing and target detection is crucial for better detection performances. Finally, the existing studies on hyperspectral image analysis mostly handle the detection problem within a single stage. However, a two-stage detector, where the possible locations of the target are found in the first stage and then the target is classified in the second stage, is an alternative promising approach, as in the case of the latest deep learning-based object detection networks working on RGB images.

In this article, inspired by the latest two-stage deep learning-based object detectors, we have proposed a deep learning-based gas detection method combining unmixing with classification. The proposed method first converts the radiance data to luminance temperatures and then performs unmixing on the resulting data to detect endmembers and related abundance values for each pixel. Considering the success of autoencoders for the representation of complex scenes in deep learning, we have utilized 3D-CNN and autoencoder for the unmixing part to model the interactions in the thermal LWIR range, rather than using conventional geometry-based unmixing methods. Then, the proposed method further classifies the extracted endmembers by using a three-layer fully connected network, which examines the presence of the target gas signatures in the endmembers.

In this regard, the first contribution of the proposed 3D-CNN and autoencoder-based gas detection method for LWIR hyperspectral images is to eliminate the need for the selection of optimum threshold values in previous conventional methods. Second, the proposed deep learning-based detection method with hyperspectral unmixing significantly improves the detection performances for gases with respect to the existing gas detection methods using statistical detection methods such as spectral angle mapper (SAM) and adaptive cosine estimator (ACE) in LWIR thermal range. Finally, the ablation study for the possible different combinations of the proposed system reveals its better performances with respect to similar structures using conventional unmixing methods and also classification methods directly working on the data without unmixing.

The rest of this article is organized as follows. Section II gives the details of the proposed deep learning-based model for gas detection. Section III describes the test data, distance metrics, and performance metrics. Section IV presents the experimental results and the comparisons with the baseline gas detection methods. Finally, Section V concludes the article.

II. PROPOSED METHOD

The proposed gas detection method mainly consists of three stages. The first stage is preprocessing stage, where the raw data is converted to luminance temperature data. In the second stage, the pure materials in the data are determined by the proposed 3D convolution and autoencoder-based hyperspectral unmixing network along with the abundance ratios of these materials in each pixel. The proposed 3-D convolution and autoencoder-based model in this stage is a special design which addresses the essential constraints in unmixing problem, namely the positivity constraint and sum-to-one constraint for the weights of the abundances. Given the multidimensional characteristics of hyperspectral images and their specific features in LWIR range it is clear that CNN structures without autoencoder cannot sufficiently address these specific constraints for unmixing problem. In the last stage, it is determined whether there is gas in the endmembers obtained from the data with the proposed three-layer fully connected network for detection. Ultimately, the class assignment is performed for each pixel according to the abundance ratio. Every part in the proposed framework addresses one essential problem for the deep learning-based gas detection in hyperspectral images.

A. Radiance to Luminance Temperature Conversion

In the literature, most of the studies use luminance temperature of hyperspectral data for gas detection applications [4], [6], [7], [8], [9], [13], [14], [16], [18] as such information indicates a more steady characteristic than the radiance data for varying conditions. The conversion of hyperspectral radiance data to luminance temperatures is carried out in three stages. The data obtained in the first step is converted to the brightness temperature value [28]. Then, the blackbody curve is calculated in accordance with the Planck curve [29] with respect to the maximum temperature inherited in the observed data. This curve is then further processed to eliminate the atmospheric effects to obtain the ultimate data for the detection [30].

The first step of luminance-temperature conversion can be expressed as

$$T(\lambda, L) = \frac{c_2 \, \lambda}{\ln\left(\frac{c_1 \lambda^3}{L} + 1\right)} \tag{1}$$

where $c1$ and $c2$ are constants, λ is the wavenumber, and L represents the spectral radiance data [30].

In the second step, the Planck formula, which is used to obtain the blackbody curve, is calculated by using the luminance-temperature data, as

$$B(\lambda, L) = \frac{c_1 \lambda^3}{\exp\left(\frac{c_2 \lambda}{L} - 1\right)} \tag{2}$$

where $c1$ and $c2$ are fixed values, and λ is the wavenumber.

The last step is the elimination of atmospheric effects from the data with the resulting black body curve. To this aim, one of the common methods, namely, the blackbody radiation curve compensation algorithm [8] is adopted in this research. The final

TABLE I
RELATED PARAMETERS FOR THE PROPOSED MODEL

	Layer type	Input	Filter	Output	Activation function
Convolutional layer	3-D Convolution Layer 1	1xPx3x3	21x2x2		ReLU
	3-D Convolution Layer 2	3-D Convolution Layer 1	11x2x2		ReLU
	3-D Convolution Layer 3	3-D Convolution Layer 2	7x1x1		ReLU
	Flatten Layer	3-D Convolution Layer 3			ReLU
Encoder part	Linear 1	Flatten Layer		32	ReLU
	Linear 2	Linear 1		16	ReLU
	Linear 3	Linear 2		8	Linear
	Linear 4	Linear 3		2	ReLU
	Normalization Layer	Linear 4		2	-
Decoder part	Output	2		P	ReLU

corrected data is calculated as

$$C_\lambda = \frac{BB_{\max}}{B(\lambda, L)} S_\lambda, \tag{3}$$

where λ is wave number, C_λ is the corrected brightness value at λ, $BB_{\max}$ is the maximum brightness value in the Black-Body curve, B(λ, T) is the blackbody brightness value at λ and S_λ is the radiance value measured by the hyperspectral sensor at λ [30]. The corrected luminance temperature values are utilized for gas detection in the further stages.

B. Proposed 3-D Convolution and Autoencoder-Based Hyperspectral Unmixing

After the luminance temperature conversion, the next step is unmixing to find the relevant endmembers and abundance ratios. The proposed hyperspectral unmixing model for gas detection consists of two main parts as 1) 3D-CNNs; and 2) autoencoder. The 3-D convolution and autoencoder-based hyperspectral unmixing network is used both to obtain the signatures of the pure pixels in the data and to calculate the abundance value in each pixel.

Fig. 1 illustrates the general scheme of the proposed deep learning-based network for hyperspectral unmixing. Table I illustrates the more detailed structure of the proposed model with the related parameters. In accordance with the table, the main stages of the proposed network layout are as follows.

The convolution layer, which is the first part illustrated on the left part of the figure, consists of a flattened layer followed by three different 3-D convolution layers. 3-D convolution filters allow to include both spectral and spatial information. In the figure, P is the input channel of the hyperspectral data corresponding to the number of spectral bands in the hyperspectral image. L1, L2, and L3 are the spectral dimensions of 3-D convolution filters. The output of this section is a flattened layer of 1-D vectors transformed from 3-D inputs. These 1-D vectors enter the next section as input. The x, y, and z parameters in the definition of the filters correspond to the terrestrial coordinates in the x and y
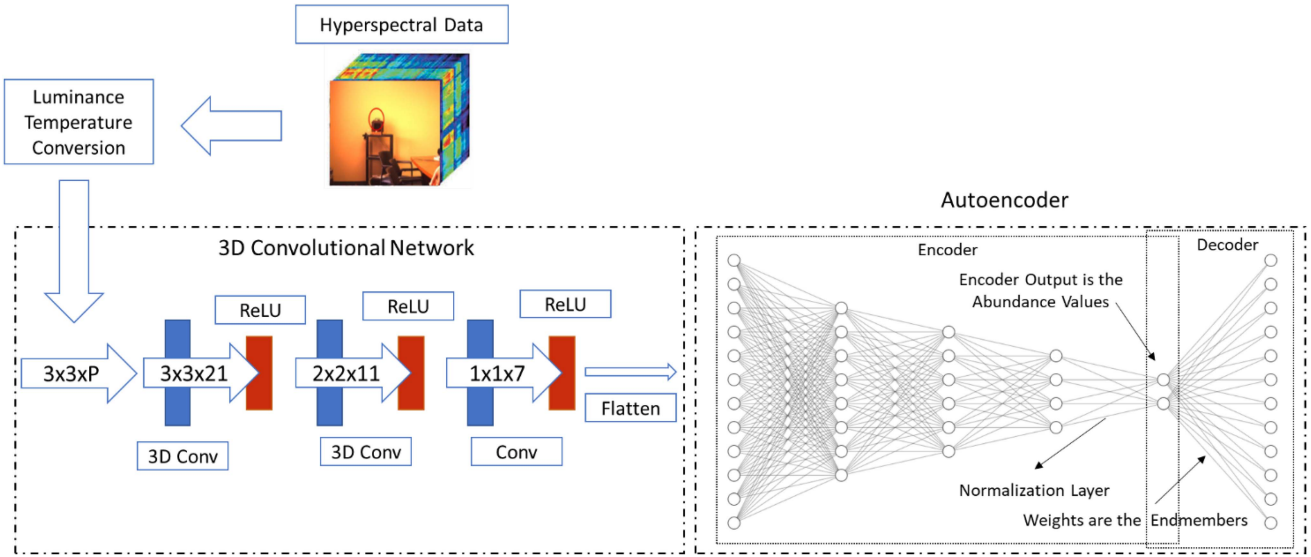
Fig. 1.   General scheme of the proposed model.

directions, and the third z dimension corresponds to the spectral coordinate.

The second part of the proposed model, namely the autoencoder part shown on the right side of Fig. 1, basically consists of two main stages: encoder and decoder. Autoencoders are often used to reduce the dimension of the input to a smaller size and learn data structures by restoring them to the input size. This structure primarily reduces the spectral signal received in the encoder part to the number of endmembers and performs the abundance detection process. Then, hyperspectral unmixing is performed by generating a signal again in the decoder layer. Constraints such as positivity and sum to one in the hyperspectral unmixing process can be applied in the last layer of the encoder part. In the proposed model, the output of the 3-D convolution section is given as input to the autoencoder section. Then, the encoder–decoder structure is established, which provides abundance estimation and endmember estimation operations. It should be noted that the weights between the encoder network and the decoder network correspond to the endmembers and the nodes in the last layer of the encoder network gives the abundance values for each endmember in the proposed design.

The most important factor in the proposed model, which affects the performance in the autoencoder part, is the normalization layer. This layer, which is applied before the encoder output, enforces the two most important constraints in hyperspectral unmixing. These are the constraint of positivity, which is the positive encoder output provided by rectified linear unit (ReLU), and the constraint of the sum to one. This layer is applied as follows:

$$w_{\text{output}, i} = \frac{w_{\text{input}, i}}{\sum_{j=1}^{T} w_{\text{input}, j}} \qquad (4)$$

where i is the index of the node, w is the weight matrix, and T is the length of the node. Finally, this layer provides both the detection of pure materials in the data and their abundance to be used in the ultimate decision.

TABLE II
DETAILS OF THE FULLY CONNECTED NETWORK

Layer	Input size	Output size	Filter
Linear layer	Band size	32	ReLU
Linear layer	32	8	ReLU
Linear layer	8	4	

The proposed 3D-CNN encoder-based hyperspectral unmixing algorithm is used to detect pure spectral signatures in data and the abundance values of these materials in each pixel. The detection of gas spectral signatures within these endmembers is performed in this ultimate layer. For this purpose, a 3-layer fully connected neural network is designed for gas detection within the endmembers. The parameters of the network with 3 layers are given in Table II. The loss function of this network is selected as the typical cross-entropy-loss metric. The proposed network is trained using the spectrum of background and spectral signatures of the gases which are determined as targets.

C. Gas Detection Over Endmembers and Abundances

The hyperspectral images in a natural scene can be represented with two components, which are: 1) the background thermal radiation; and 2) the radiation due to the targeted gas. Given the input spectrum to the trained network, the network classifies the test data as background or one of the gas classes. In the proposed model, the endmembers are given as input to this network to determine the class for each endmember. If a gas signature is detected in one of the tested endmembers, the pixels with more than 50% abundance values in the hyperspectral data are marked as the detected gas type. Considering that the background signature is very dominant in LWIR images, we experimentally observed that the given threshold is the logical
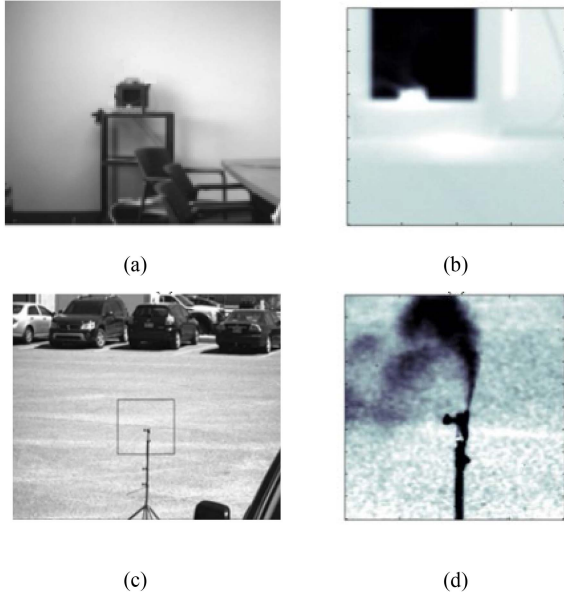
Fig. 2. (a) Grayscale image of the scene for the capturing of methane, (b) a sample band image from the captured hyperspectral LWIR image for methane, (c) the grayscale image of the scene for the capturing of sulfur dioxide, and (d) a sample band image from the captured hyperspectral LWIR image.

Fig. 3. Brightness temperature signals for methane, sulfur dioxide, and butane gases.

threshold to decide whether the tested pixel is closer to the background or the target gas signature.

III. DATASETS AND DISTANCE METRICS

Two different real hyperspectral LWIR gas images are used in this study. These images are acquired with the Long Wave InfraRed camera manufactured by TELOPS. Images of the scenes are given in Fig. 2. Since gas diffusion is not fully predictable, ground truth data are roughly generated.

The Methane images in Fig. 2(a) are taken at 877 cm⁻¹ and 1285 cm⁻¹ spectral ranges. The methane image has 124 bands and a spatial size of 200×200 pixels. The ambient temperature was determined as 300 K. There is only Methane gas in this scene shot from 2–3 m away. The sulfur dioxide images in Fig. 2 are taken at 851 cm⁻¹ and 1288 cm⁻¹ spectral ranges. This image has 171 bands, and the ambient temperature is determined as 302 K. This image also has 200×200 pixels. There is only sulfur dioxide gas in this scene shot from a distance of 5 m. Ground truth maps for methane and sulfur dioxide gases are given in Fig. 2. Method performance comparisons are performed using the ground truth information given in the figure.

Fig. 3 also shows the spectral signatures of methane, butane, and sulfur dioxide gases with their absorbance characteristics between 851 cm⁻¹ and 1290 cm⁻¹ to cover both images. These signatures are taken from the NIST database [31]. The gas detection process which is carried out in the last stage of the proposed system utilizes these signatures with the background information.

Different distance metrics are used for abundance estimation accuracy, endmember estimation accuracy, and the cost function of the optimization algorithm. These metrics are selected as mean squared error (MSE) and SAM. The distance metric SAM
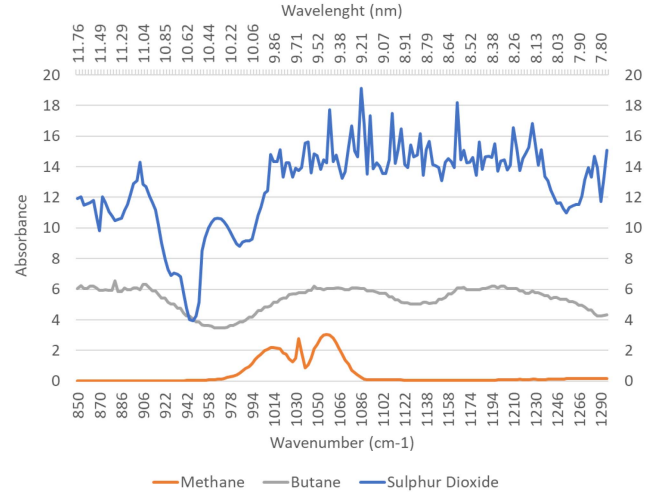
[32] has been used both as a cost function and for endmember estimation accuracy. The formulation of SAM is given as

$$\text{SAM} = \arccos\left(\frac{\langle \boldsymbol{x}_i, \hat{\boldsymbol{x}}_i \rangle}{||\boldsymbol{x}_i||_2 ||\hat{\boldsymbol{x}}_i||_2}\right) \quad (5)$$

where $\boldsymbol{x}_i$ is the reference spectral signature, $\hat{\boldsymbol{x}}_i$ is the estimated spectral signature.

IV. EXPERIMENTAL RESULTS

The experimental results and comparisons for the proposed 3-D convolution and autoencoder-based method and the conventional gas detection methods are presented in this section. The first part of the experiments illustrates the performance of conventional gas detection methods, based on SAM and ACE methods. The performances are given for both methane and sulfur dioxide gases. The second part reveals the performance of the proposed method in comparison with the conventional methods. The performance of the endmember estimation and abundance estimation processes are both discussed for different distance metrics such as SAM and MSE for the optimization of the proposed 3-D convolution and autoencoder-based network.

Due to the technical limitations to exactly determine the position of the gas molecules in gas detection studies, it is not possible to have a regular receiver operating characteristics analysis as in the other detection studies for solid targets. Therefore, the overall experimental evaluation goes over the score images or approximate ground truth data extracted with visual inspection. Rather than using uncertain ground truth data, we perform the experimental evaluation over score images in the performed study.

A. Experimental Results for Conventional Methods

The SAM [32] and ACE [33] algorithms are used for methane and sulfur dioxide gas detection. The SAM method measures the angle between two spectral signals in radians. The similarity between the two vectors is high when the angle is small in the
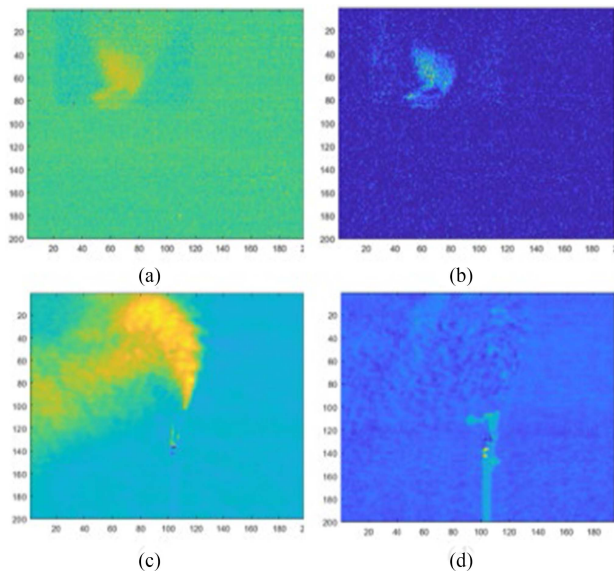
Fig. 4. Results for methane gas, (a) SAM and (b) ACE, and results for sulfur dioxide gas, (c) SAM and (d) ACE.
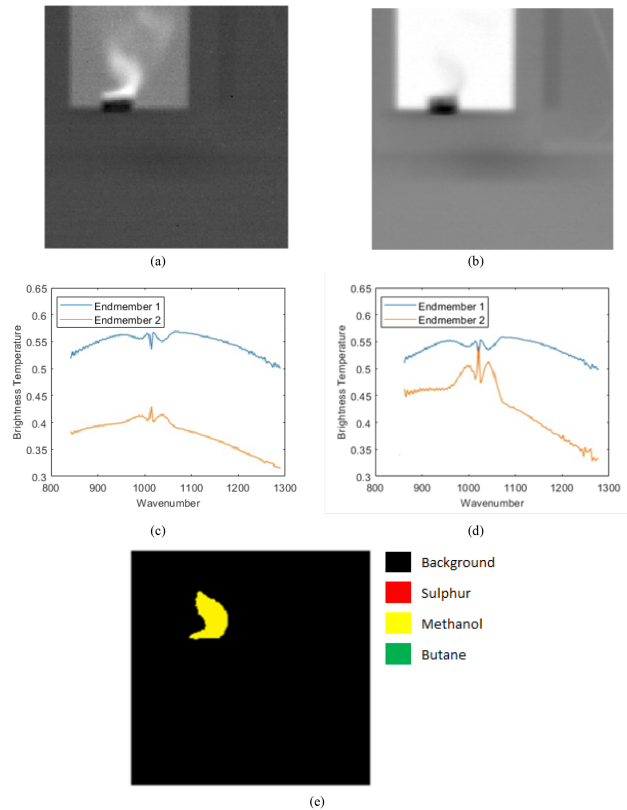

Fig. 5. Results for the methane gas. (a) After the unmixing stage in the proposed system (distance metric is SAM), (b) after the unmixing stage in the proposed system (distance metric is MSE), (c) estimated endmembers with the proposed method. (distance metric is SAM), (d) estimated endmembers with the proposed method. (distance metric is MSE.) (e) The obtained result with the proposed framework.

SAM method. SAM metric ensures a more robust evaluation due to its invariance to scaling compared to the MSE metric. Different than the SAM algorithm, ACE suppresses the background information by utilizing the covariance matrix of the data during the detection. The conventional gas detection methods based on SAM and ACE generally perform the detection process by comparing the reference target signature with the spectral signature of each pixel. These methods are applied to the resulting data cubes which are obtained after transforming the hyperspectral cubes into luminance temperature data. In this regard, in this study, the SAM-based method for gas detection was proposed by Öztürk et al. [18] and the ACE-based gas detection method proposed by Omruuzun and Cetin [8] are selected as the baseline methods for the comparisons.

Fig. 4 shows the output scores for detection of methane and sulfur dioxide gases with the SAM and ACE-based gas detection methods for each pixel before the thresholding. The regions where the sulfur dioxide gas is found seem to be more clearly selected by the SAM method compared to ACE. In addition, the background suppression is more apparent for sulfur dioxide in Fig. 4(a) and (b) compared to the methane in Fig. 4(c) and (d). The main problem for both of the output images obtained with the statistical detection methods is to regularly determine the threshold value for the decision. As such a selection is dependent on the statistical characteristics of the utilized images, which makes the generic application more challenging, there is a need for a different method to determine the desired targets in the data. One of the main motivations of this work was to eliminate such a thresholding process by means of utilizing a deep neural network-based approach.

B. Experimental Results for the Proposed Method

The experiments for the proposed method involve determining the parameters of the deep learning structure used in the hyperspectral unmixing part. For this purpose, the performance of different cost functions with the proposed system has been examined. The SAM and MSE are selected as cost functions.

1) Experimental Results for Methane Gas: The results obtained with the proposed method are given in Fig. 5(a) and (b) for the distance metrics SAM and MSE. The pure pixel signatures obtained for these results are given in Fig. 5(c) and (d), respectively. The spectral signature of one of the estimated endmembers, illustrated with orange color in Fig. 5(c) and (b) are similar to the spectral characteristics of methane gas taken from the NIST database [31] which is illustrated in Fig. 3. Although there is such similarity in both methods, the methane gas characteristics are more apparent in the results obtained with SAM. The SAM distances between the estimated endmembers and the spectral signature of methane are 0.05 and 0.11 radians when the SAM metric is utilized as the distance metric in the proposed method. Accordingly, the SAM distances between the estimated endmembers and the spectral signature of methane are 0.07 and 0.12 radians for the case of the MSE metric. The proposed method successfully extracts the endmembers for the background and inscene gas.

The extracted endmembers are classified as described in Section II-C. The results obtained after the classification process g are given in Fig. 5(e). While the presented method can
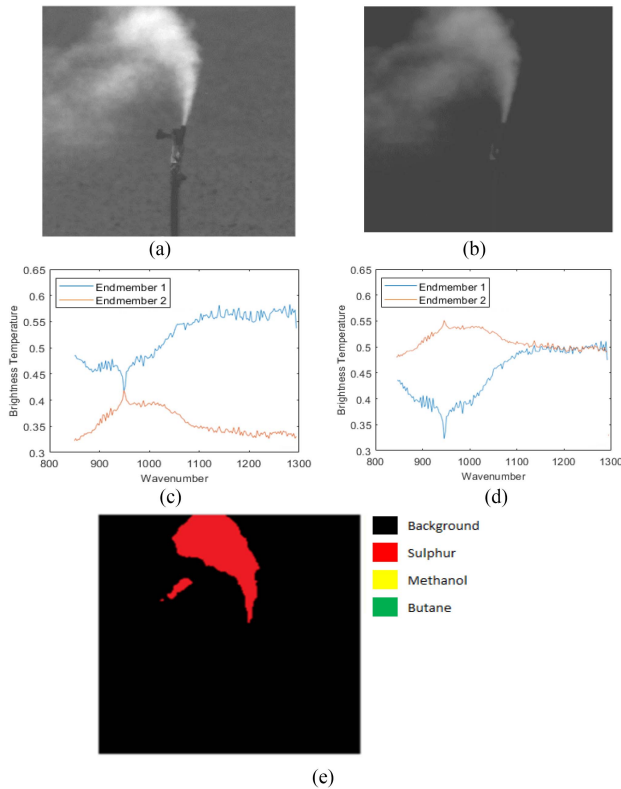
Fig. 6. Results for the sulfur dioxide gas. (a) After the unmixing stage in the proposed system (distance metric is SAM), (b) after the unmixing stage in the proposed system (distance metric is MSE), (c) estimated endmembers with the proposed method. (distance metric is SAM), (d) estimated endmembers with the proposed method. (distance metric is MSE.) (e) The obtained result with the proposed framework.

successfully extract the gas regions more apparently, its performance with respect to the SAM and ACE-based methods are also superior as can be revealed from the comparison of Figs. 4 and 5.

2) Experimental Results for Sulfur Dioxide Gas: The experimental results for sulfur dioxide are given in Fig. 6 for the SAM and MSE metrics in the proposed method Similar to methane gas results, the cost function SAM reveals more successful results for sulfur dioxide gas. This can be explained by the fact that when background and gas signatures are selected, the angle difference when using SAM is greater than the one when using MSE.

The examples marked with red color among the results obtained using the SAM and MSE distance metrics are given in Fig. 6(a) and (b). The pure pixel signatures obtained for the marked results are given in Fig. 6(c) and (d). Although the spectral characteristics of the sulfur dioxide gas appearing in the orange signatures are included in both methods, the sulfur dioxide gas characteristic is more apparent in the results obtained with SAM. 0.9 and 0.12 SAM were taken for 0.9 and 0.21 MSE. The result of the automatic detection of sulfur dioxide gas is given in Fig. 6(e).

Similar to methane gas, the proposed method also reveals a successful detection performance for sulfur dioxide gas. The proposed deep learning-based gas detection method eliminates the need for thresholding. While such a thresholding operation requires a delicate selection for different data in traditional

methods, the proposed method only utilizes the abundance values for such a selection with its deep learning-based structure without being affected by data changes. This threshold value, which is not affected by data changes, appears as a parameter change in the proposed deep learning-based method. The proposed deep learning-based method gives much better results compared to the conventional gas detection methods. Without loss of generality, the proposed method can also be easily adapted to different gases. In addition, the experiments reveal that the proposed method can successfully work for different gases with the same system parameters such as learning rate, batch size, and cost function.

V. DISCUSSION

The proposed deep learning-based gas detection method combining unmixing with classification gives better performances than the other gas detection methods proposed for hyperspectral LWIR images as illustrated in the previous section. Another important discussion for the performed study is however to understand the contributions of the proposed system with respect to its possible different combinations. To this aim, we have performed an ablation study in this section by considering the main parts of the proposed system.

The proposed system is composed of three main stages namely, 1) luminance temperature conversion; 2) 3D-CNN based autoencoder; and 3) a 3-layer fully connected network for gas type detection. The first stage, luminance temperature conversion, is one of the essential processes in gas detection to suppress the variations in hyperspectral gas radiation in LWIR thermal range. Therefore, the ablation study has skipped this part and focused on the following two main aspects after temperature conversion.

1) The proposed system combining unmixing with classification is compared with direct classification without performing unmixing on the data to reveal the contribution of unmixing to gas detection.
2) The proposed system using 3-D convolution and autoencoder-based unmixing is compared with a similar structure using conventional unmixing methods in order to reveal the contribution of autoencoders for gas detection.

A. Comparison of the Proposed System With Direct Classification Without Unmixing

In order to compare the proposed gas detection using hyperspectral unmixing with the direct classification, we have utilized support vector machines (SVM) directly on the data after luminance temperature conversion. For the experiments, the SVM algorithm [34] is trained with the average background signals extracted from the data and predictions are made on the data with the spectral signatures taken from NIST database [31]. The radial basis function kernel is utilized for the SVM. Implementation is carried out with the Python Sklearn library [35].

Fig. 7(a) and (b) illustrates the detection results for SVM for the sulfur dioxide and methane, respectively. Although the results for sulfur dioxide contain a gas region, the gas pipe in the scene is also detected as the target, which significantly decreases the performance. In the case of the LWIR image for
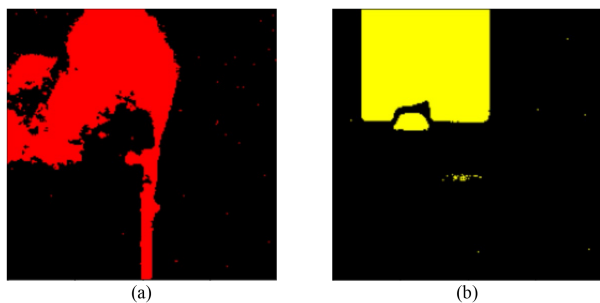
(a)   (b)

Fig. 7.   Classification results with SVM for (a) sulfur dioxide (b) methane.

methane, both the part, where the gas outlet is located and the entire closed area are marked as methane. In both of the LWIR images, SVM detects different small areas as gases in different regions. The detection performances are lower compared to the results in Figs. 5(e) and 6(e) for the proposed method, and the results in Fig. 4(a)–(d) given for the state-of-the-art gas detection methods.

B. Comparison of the Proposed System With the Similar Structure Using Conventional Unmixing

The experimental results in the case of using a traditional end-member estimation and abundance estimation method instead of the utilized 3-D convolution and autoencoder-based unmixing are presented in this section. Vertex component analysis [36] is used for endmember estimation as one of the baseline methods for endmember estimation in the literature. Fully constrained least square [37] method is utilized for the abundance estimation for the extracted endmembers.

Fig. 8(a) and (b) show the abundance estimation results for Sulfur Dioxide and Methane data for one of the endmembers. The conclusions for the abundance results for the other endmember are also similar. The threshold value used for the abundance estimation is determined as 0.5 as in the proposed method while presenting the gas dominant regions. The results indicate that the abundances for the gas regions are not found correctly with vertex component analysis (VCA). The pure pixel signatures corresponding to the extracted endmembers are given in Fig. 8(c) and (d). As observed in the figures, the two endmembers indicate some similar characteristics with respect to each other. The background signal in LWIR thermal range is selected as an endmember by VCA. However, this background signal also exists with different abundances in the other pixels as well. As VCA selects endmembers from pure pixels in the data, the other endmember also resembles the background.

Fig. 8(e) and (f) show the final decision images for sulfur dioxide and methane, respectively. For both images, the gas exit regions are detected instead of the gas regions. In the visual examinations made for this experiment, it has been observed that when different threshold values are used, very similar results to the ones in the SVM experiments can be obtained as well. However, as in other methods involving manual threshold selection, threshold selection is considered as a separate problem. For this reason, the selection of the most abundant material in the pixel was applied here as in the proposed method.
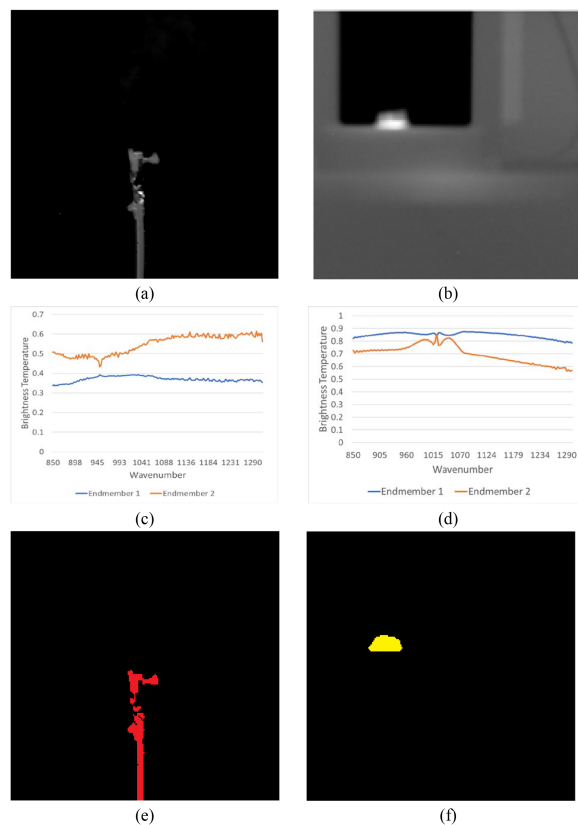

(a)   (b)

(c)   (d)

(e)   (f)

Fig. 8.   Estimated abundances with conventional hyperspectral unmixing for (a) sulfur dioxide (b) methanol (c) Estimated endmembers with VCA for sulfur dioxide (d) Estimated endmembers with VCA for methane (e) Final result after thresholding for sulfur dioxide (f) Final result after thresholding for methane.

VI. Conclusion

We have proposed a deep learning-based gas detection method which combines 3D-CNN and autoencoder-based hyperspectral unmixing with neural network-based classification. The experiments reveal that a detection approach combining deep learning based unmixing with classification is better than the existing methods to handle the gas detection problem in LWIR range. An ablation study with respect to the possible different combinations for such a system, such as using direct classification methods or using the same structure with other unmixing methods are also performed. The proposed system combining unmixing with classification has given better performances than direct classification without performing unmixing. In addition, the 3D-CNN and autoencoder-based unmixing has indicated better results than the conventional unmixing for the proposed gas detection framework. The experiments have revealed that using SAM as a cost function in the proposed method yields more successful results than the MSE metric. The performed study does not require thresholding, unlike the conventional gas detection methods. Finally, the proposed gas detection method achieves better results than state of the art gas detection methods in LWIR range due to its high learning capacity with 3-D convolutional layers. Without loss of generality, the proposed system can be adapted to different gases by integrating the target gas signature into the classification module of the proposed system in the last stage.

REFERENCES

[1] R. O. Green et al., "Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (AVIRIS)," *Remote Sens. Environ.*, vol. 65, no. 3, pp. 227–248, 1998.

[2] M. Govender, K. Chetty, and H. Bulcock, "A review of hyperspectral remote sensing and its application in vegetation and water resource studies," *Water Sa*, vol. 33, no. 2, pp. 145–151, 2007.

[3] P. Y. Foucher and S. Doz, "Real time gas quantification using thermal hyperspectral imaging: Ground and airborne applications," Accessed: Jan. 18, 2023. [Online]. Available: https://www.sto.nato.int/publications/STO%20Meeting%20Proceedings/STO-MPSET-277/MP-SET-277-18.pdf

[4] A. Vallières et al., "Algorithms for chemical detection, identification and quantification for thermal hyperspectral imagers," in *Proc. Chem. Biol. Standoff Detection III*, vol. 5995, 2005, Art. no. 59950G.

[5] J.-M. Thériault, G. Fortin, F. Bouffard, H. Lavoie, P. Lacasse, and J. Lévesque, "Hyperspectral gas and polarization sensing in the LWIR: Recent results with MoDDIFS," in *Proc. 5th Workshop Hyperspectral Image Signal Process.: Evol. Remote Sens.*, 2013, pp. 1–4.

[6] D. W. Messinger, "Gaseous plume detection in hyperspectral images: A comparison of methods," in *Proc. Algorithms Technol. for Multispectral, Hyperspectral, Ultraspectral Imagery X*, vol. 5425, 2004, pp. 592–603.

[7] M. Kastek, T. Piatkowski, R. Dulski, M. Chamberland, P. Lagueux, and V. Farley, "Method of gas detection applied to infrared hyperspectral sensor," *Photon. Lett. Poland*, vol. 4, no. 4, pp. 146–148, 2012.

[8] F. Omruuzun and Y. Y. Cetin, "Endmember signature based detection of flammable gases in LWIR hyperspectral images," in *Proc. Adv. Environ., Chem., Biol. Sens. Technol. XII*, vol. 9486, 2015, pp. 168–176.

[9] C. C. Funk, J. Theiler, D. A. Roberts, and C. C. Borel, "Clustering to improve matched filter detection of weak gas plumes in hyperspectral thermal imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 7, pp. 1410–1420, Jul. 2001.

[10] D. R. Pogorzala, D. W. Messinger, C. Salvaggio, and J. R. Schott, "Gas plume species identification by regression analyses," in *Proc. Algorithms Technol. for Multispectral, Hyperspectral, Ultraspectral Imagery X*, vol. 5425, 2004, pp. 583–591.

[11] F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 1, pp. 208–216, Jan. 1992.

[12] T. S. Spisz, P. K. Murphy, C. C. Carter, A. K. Carr, A. Vallières, and M. Chamberland, "Field test results of standoff chemical detection using the FIRST," in *Proc. Chem. Biol. Sens. VIII*, vol. 6554, 2007.

[13] L. Sagiv, S. R. Rotman, and D. G. Blumberg, "Detection and identification of effluent gases by long wave infrared (LWIR) hyperspectral images," in *Proc. IEEE 25th Conv. Elect. Electron. Engineers Isr.*, 2008, pp. 413–417.

[14] E. Hirsch and E. Agassi, "Detection of gaseous plumes in IR hyperspectral images using hierarchical clustering," *Appl. Opt.*, vol. 46, no. 25, pp. 6368–6374, 2007.

[15] M. Kastek, T. Piatkowski, and P. Trzaskawka, "Infrared imaging fourier transform spectrometer as the stand-off gas detection system," *Metrol. Meas. Syst.*, vol. 18, no. 4, pp. 607–620, 2011.

[16] P. Kuflik and S. R. Rotman, "Band selection for gas detection in hyperspectral images," in *Proc. IEEE 27th Conv. Elect. Electron. Engineers Isr.*, 2012, pp. 1–4.

[17] S. Sabbah, R. Harig, P. Rusch, J. Eichmann, A. Keens, and J.-H. Gerhard, "Remote sensing of gases by hyperspectral imaging: System performance and measurements," *Opt. Eng.*, vol. 51, no. 11, 2012, Art. no. 111717.

[18] Ş. Öztürk, Y. Artan, and Y. E. Esin, "Ethene and CO2 gas detection in hyperspectral imagery," in *Proc. 24th Signal Process. Commun. Application Conf. (SIU)*, 2016, pp. 357–360.

[19] J. Theiler and S. P. Love, "Algorithm development with on-board and ground-based components for hyperspectral gas detection from small satellites," in *Proc. Algorithms, Technol., Appl. for Multispectral Hyperspectral Imagery XXV*, vol. 10986, 2019.

[20] Y. C. Kim, H.-G. Yu, J.-H. Lee, D.-J. Park, and H.-W. Nam, "Hazardous gas detection for FTIR-based hyperspectral imaging system using DNN and CNN," in *Proc. Electro-Opt. Infrared Syst.: Technol. Appl. XIV*, vol. 10433, 2017.

[21] L. Zhang, J. Wang, and Z. An, "Classification method of CO2 hyperspectral remote sensing data based on neural network," *Comput. Commun.*, vol. 156, pp. 124–130, 2020.

[22] S. Kumar, C. Torres, O. Ulutan, A. Ayasse, D. Roberts, and B. S. Manjunath, "Deep remote sensing methods for methane detection in overhead hyperspectral imagery," in *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis.*, 2020, pp. 1776–1785.

[23] R. Gu, "Methane gas emission detection using deep learning and hyperspectral imagery," in *Proc. IEEE 3rd Int. Conf. Front. Technol. Inf. Comput.*, 2021, pp. 36–44.

[24] S. Henrot, J. Chanussot, and C. Jutten, "Dynamical spectral unmixing of multitemporal hyperspectral images," *IEEE Trans. Image Process.*, vol. 25, no. 7, pp. 3219–3232, Jul. 2016.

[25] Z. Shi, W. Tang, Z. Duren, and Z. Jiang, "Subspace matching pursuit for sparse unmixing of hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 6, pp. 3256–3274, Jun. 2014.

[26] G. Tochon, D. Pauwels, M. D. Mura, and J. Chanussot, "Unmixing-based gas plume tracking in LWIR hyperspectral video sequences," in *Proc. 8th Workshop Hyperspectral Image Signal Process.: Evol. Remote Sens.*, 2016, pp. 1–5.

[27] N. Fiscante, P. Addabbo, F. Biondi, G. Giunta, and D. Orlando, "Unsupervised sparse unmixing of atmospheric trace gases from hyperspectral satellite data," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, 2022, Art. no. 6006405.

[28] R. Harig and G. Matz, "Toxic cloud imaging by infrared spectrometry: A scanning FTIR system for identification and visualization," *Field Anal. Chem. Technol.*, vol. 5, no. 1–2, pp. 75–90, 2001.

[29] M. Planck, *The Theory of Heat Radiation*. Philadelphia, PA, USA: P. Blakiston's Son & Co., 1914.

[30] Y. Gür, "Detection and identification of greenhouse gases using infrared hyperspectral imagery," M.Sc. Dissertation, Inf. Syst. Dept. Middle East Tech. Univ., Ankara, Turkey, 2017.

[31] C. I. Watson and C. L. Wilson, "NIST special database 4," *Fingerprint Database, Nat. Inst. Standards Technol.*, vol. 17, no. 77, p. 5, 1992.

[32] F. A. Kruse et al., "The spectral image processing system (SIPS)—interactive visualization and analysis of imaging spectrometer data," *Remote Sens. Environ.*, 1993.

[33] S. Kraut, L. L. Scharf, and R. W. Butler, "The adaptive coherence estimator: A uniformly most-powerful-invariant adaptive detection statistic," *IEEE Trans. Signal Process.*, vol. 53, no. 2, pp. 427–438, Feb. 2005.

[34] L. Lin and Z. Xiaolong, "Optimization of SVM with RBF Kernel," *Comput. Eng. Appl.*, vol. 29, 2006, pp. 190–193.

[35] F. Pedregosa et al., "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011.

[36] J. M. P. Nascimento and J. M. B. Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 898–910, Apr. 2005.

[37] D. C. Heinz and C.-I. Chang, "Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 3, pp. 529–545, Mar. 2001.

Okan Bilge Özdemir received the M.S. and Ph.D. degrees in information systems from the Middle East Technical University (METU), Ankara, Turkey, in 2013 and 2021.

He is currently an Assistant Professor with the Department of Computer Engineering, Artvin Çoruh University, Artvin, Turkey. His research interests include deep learning, image processing, hyperspectral image analysis, and medical image processing.

Alper Koz was born in Kahramanmaras, Turkey, in 1978. He received the B.S., M.S., and Ph.D. degrees in electrical and electronics engineering from Middle East Technical University (METU), Ankara, Turkey, in 2000, 2002, and 2007, respectively.

He performed postdoctoral research studies with Multimedia Signal Processing Laboratory, Delft University of Technology, Delft, The Netherlands, from 2008 to 2010; with Laboratory Signals and Systems, SUPELEC, Paris, France, from 2010 to 2011; and with Multimedia Group of Telecom ParisTech, Paris, France, from 2011 to 2013. He joined the Center for Image Analysis of METU in 2014. His research interests include hyperspectral image analysis, high dynamic range image and video processing, and multimedia watermarking and fingerprinting.

Dr. Koz is an Associate Editor for *Signal Processing: Image Communication* (Elsevier).