

## Article

# VOC-Net: A Deep Learning Model for the Automated Classification of Rotational THz Spectra of Volatile Organic Compounds

M. Arshad Zahangir Chowdhury \* , Timothy E. Rice  and Matthew A. Oehlschlaeger \* 

Department of Mechanical, Aerospace, and Nuclear Engineering, Rensselaer Polytechnic Institute, 110 8th Street, Troy, NY 12180, USA

\* Correspondence: arshad.zahangir.bd@gmail.com (M.A.Z.C.); oehlsm@rpi.edu (M.A.O.)

**Abstract:** Conventional black box machine learning (ML) algorithms for gas-phase species identification from THz frequency region absorption spectra have been reported in the literature. While the robust classification performance of such ML models is promising, the black box nature of these ML tools limits their interpretability and acceptance in application. Here, a one-dimensional convolutional neural network (CNN), VOC-Net, is developed and demonstrated for the classification of absorption spectra for volatile organic compounds (VOCs) in the THz frequency range, specifically from 220 to 330 GHz where prior experimental data is available. VOC-Net is trained and validated against simulated spectra, and also demonstrated and tested against experimental spectra. The performance of VOC-Net is examined by the consideration of confusion matrices and receiver-operator-characteristic (ROC) curves. The model is shown to be 99+% accurate for the classification of simulated spectra and 97% accurate for the classification of noisy experimental spectra. The model's internal logic is examined using the Gradient-weighted Class Activation Mapping (Grad-CAM) method, which provides a visual and interpretable explanation of the model's decision making process with respect to the important distinguishing spectral features.

**Keywords:** VOC; DNN; CNN; classification; THz; spectroscopy; rotational; microelectronic; spectrometer; gas sensing



**Citation:** Chowdhury, M.A.Z.; Rice, T.E.; Oehlschlaeger, M.A. VOC-Net: A Deep Learning Model for the Automated Classification of Rotational THz Spectra of Volatile Organic Compounds. *Appl. Sci.* **2022**, *12*, 8447. <https://doi.org/10.3390/app12178447>

Academic Editor: Luis Javier Garcia Villalba

Received: 21 July 2022

Accepted: 22 August 2022

Published: 24 August 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Absorption spectroscopy in the terahertz (THz) frequency region (0.1–10 THz) allows for the identification and quantitative detection of gas-phase polar molecules, including many volatile organic compounds (VOCs). Molecules of interest can be identified via their rich, complex, and unique THz region rotational absorption spectra, which act as fingerprints. Hence, THz absorption spectroscopy and gas sensing can be valuable in many scientific, industrial and environmental applications [1–5]. Moreover, recent advances in microelectronic THz sources and detectors has increased the viability of portable, miniature spectrometers capable of sensing a wide variety of gas molecules which are active in the THz region [6–16]. Similar to other spectroscopic techniques, rotational spectra in the THz frequency region require signal/data processing for spectral interpretation, gas species identification, and quantitative gas sensing. Spectral identification is often difficult because of overlapping transitions, among a single molecular spectra or for a complex gas mixture, which is further complicated by collisional broadening at realistic pressures found in applications.

Identification of rotational fingerprints with high accuracy, sensitivity, and selectivity is a challenge that can be addressed and automated with artificial intelligence (AI) and machine learning (ML) methods. Automated spectral identification using ML requires a model that can learn to identify unique spectral features that provide fingerprints for each molecule and separate those features from similar features of other molecules. Ideally, a

ML “artificial spectroscopist” would make spectral classification decisions based on identification of unique fingerprints within a spectra just as a human spectroscopist would, while ML methods, including shallow neural networks composed of multi-layer perceptrons, have been demonstrated to identify spectra of different molecules from THz absorption spectroscopy [17], IR and Raman spectroscopy [18,19], and excitation spectroscopy [20], their black box nature, and the difficulty of interpreting the logic behind their decision making, can lead to an erosion of acceptance and trust in such models [21], especially in industrial applications. Thus, there is a need for AI-based models to not only have high accuracy, sensitivity, and selectivity, but also to have good interpretability.

AI/ML based classification generally requires data with good descriptive and variable features. Often, the original raw data is reduced to derived values which are non-reducible, while preserving the most useful and distinctive information, in a process known as feature extraction. Since their resurgence in the late 2000s, neural networks have facilitated the training of complex models capable of wide variety of tasks such as image recognition, segmentation, classification, denoising, representation learning, and computer vision [22–27]. Deep networks are able to learn patterns in data by extracting good features. Particularly, convolutional neural networks (CNN) are widely used to extract features from data with the help of convolutional and subsampling or pooling layers [22]. These features are used to represent the data in a compact and useful form.

The most common CNNs are two dimensional (2D), since they are often used on image data [23] and numerous works have demonstrated 2D CNNs for learning useful representations such as edges, contours, contrast, saturation etc. However, 2D CNNs are difficult to use to interpret spectral data because spectral data is highly-dimensional along a single axis. On the other hand, 1D CNNs, widely used to learn time-series data for classification and regression purposes [22,28], are suited for data that is highly-dimensional along a single axis. A 1D CNN moves a sliding window filter kernel along a single axis to learn features from the data (spectrum). CNNs have an added advantage over fully-connected neural networks, in that the output from each of the applied filters can be viewed and interpreted, to better understand the limitation of the training data and how the model makes classification decisions. There are several visualization techniques available to interpret CNN classification decision processes, including Grad-CAM (Gradient-weighted Class Activation Mapping) [29], which can be used to evaluate the output from every layer of the CNN via a class activation map, to observe how the model is basing classification decisions on spectral features (peak locations, feature shapes, etc.) [30]. A class activation map relates localized patterns prioritized by the model to identify corresponding classes, offering an interpretable visualization of the neural network, and can be tailored to CNN-based classification models to better understand their inner decision making logic. Ultimately, Grad-CAM applied to spectral data produces a heat map illustrating the class-discriminative regions of the spectra. Interpretable deep learning models based on CNNs for classifying VOCs have been reported by Wang et al. [30] where the Grad-CAM method [29] was implemented to compute class activation maps to interpret the learning of a CNN model from optical emission spectra.

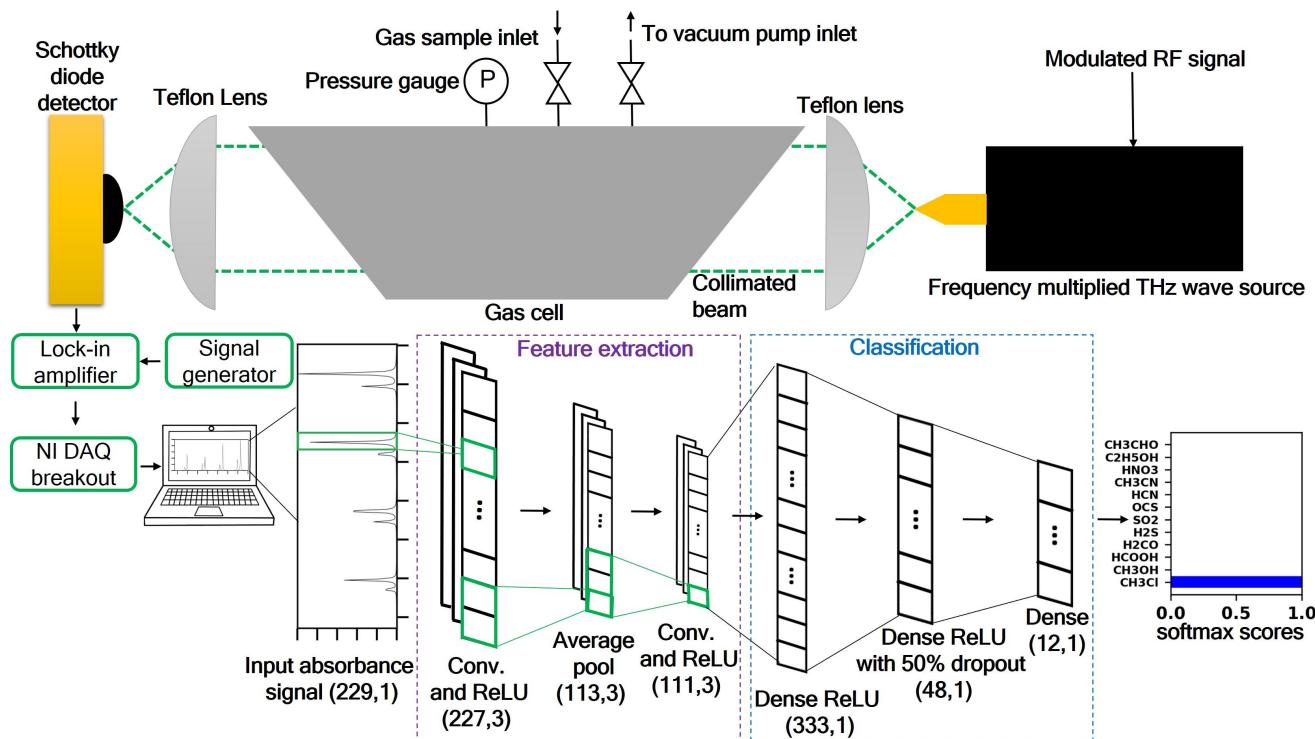
In prior work, we have demonstrated that conventional ML models, including random forests, fully connected neural networks, and support vector machines, can achieve high classification accuracy for the identification of pure gases from THz spectra [17], while the performance of these models based on specifically designed features or all available features within a frequency range is acceptable, they may fit specific spectral frequency locations instead of generalizing the spectral peaks, shapes, and widths. A deep learning classification network can offer advantages over such methods, both in terms of achieving better accuracy and other performance measures (e.g., area-under-curve (AUC) scores), but also by providing corresponding class activation maps allowing further insights into the network’s decision making process and its generalizability.

In this paper, we report a deep learning neural network classifier composed of convolutional layers for the identification of THz absorption spectra for twelve VOCs. The model

is trained and validated using simulated spectra and is demonstrated via the classification of experimental spectra. Its performance is evaluated based on confusion matrices and receiver-operating-characteristics curves (ROC). Finally, the model is analyzed using Grad-CAM activation maps to understand how the model learns from the fingerprint spectral features contained in each spectra and to determine the critical spectral features among the gases considered that should be prioritized for a cross-sensitive gas sensor design.

## 2. Methodology

The data sets, their development, and statistics, and the architecture of the VOC-Net and network training and validation approaches are described in the following subsections. A schematic of VOC-Net for automated classification of experimentally collected spectra is shown in Figure 1.



**Figure 1.** VOC-Net model for the automated classification of absorption spectra measured with a THz spectrometer.

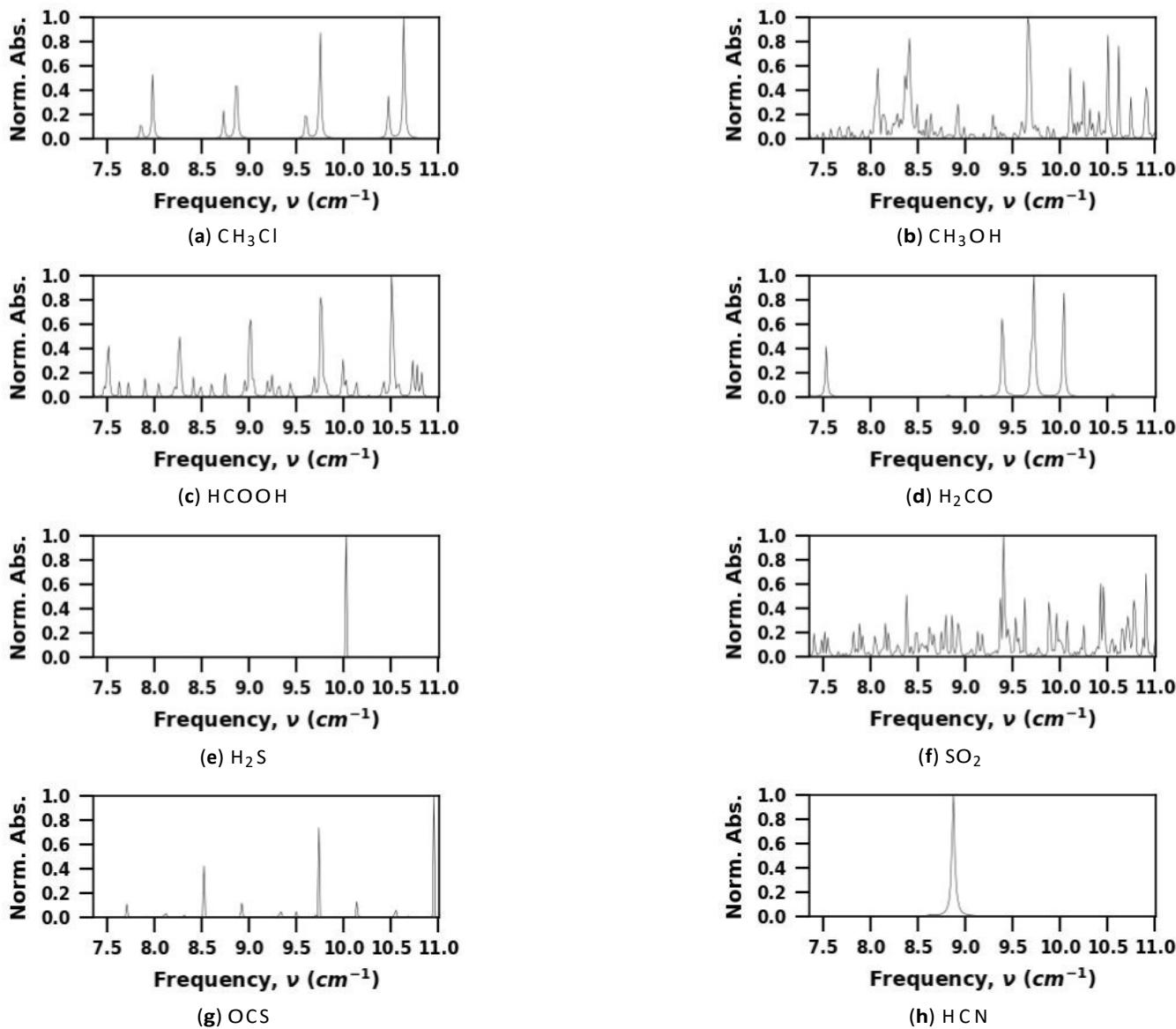
In the above schematic, an overview of how VOC-Net can be integrated with a spectrometer is illustrated. A spectrum, measured using a THz spectrometer (described in the experimental data section), is fed to VOC-Net which consists of convolutional, pooling, and dense layers. VOC-Net broadly performs two tasks, extracting relevant features from the spectra and classification. Feature extraction is a process where a raw spectrum is reduced to a set of derived values which captures the spectrum's most relevant aspects, by preserving its distinctive and unique information. Classification is the process of categorizing the spectrum as belonging to a specific class or molecule. VOC-Net classifies by producing a set of raw score values that can be converted to softmax scores, which indicate the probabilities that the spectrum belongs to each class. VOC-Net outputs the class with the highest softmax score/probability; hence, classifying the input spectrum. The softmax scores,  $s$ , are given by,

$$s(z)_i = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}} \quad (1)$$

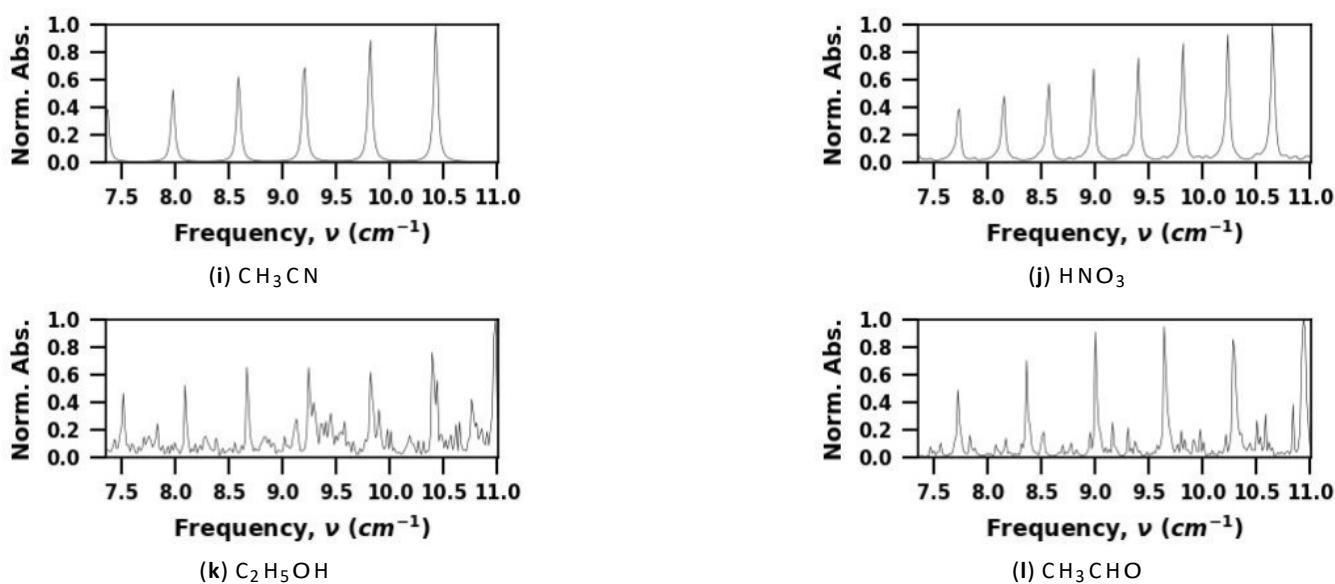
where,  $z$  is the input to the softmax layer and  $K$  is the number of molecule classes.

### 2.1. Training and Validation Data

The CNN model developed in this study was trained and validated using spectra for twelve pure VOCs that were simulated using spectroscopic parameters found in the HITRAN [31] and JPL [32] databases. Spectral simulations were carried out using the HAPI tool [33]. Representative simulated spectra for each molecule are shown in Figure 2. The frequency range considered in this study is based on the range available in our THz spectrometer, 220–330 GHz (7.33–11 cm<sup>−1</sup>). The dataset consists of spectra for the twelve VOC molecules at pressures from 0.1 to 16.5 Torr (13.3 to 2200 Pa). Each spectra is for a single pure molecule and consists of 229 absorbance values at a frequency resolution of 0.016 cm<sup>−1</sup>.



**Figure 2.** Cont.



**Figure 2.** Representative simulated spectra; conditions: 1 Torr, 297 K, 21.6 cm pathlength.

## 2.2. Experimental Data

Data used for demonstration of the trained VOC-Net model consists of experimental spectra measured in our laboratory for six VOCs: ethanol, methanol, formic acid, acetonitrile, and acetaldehyde. Absorption spectra for these molecules have been previously measured in the 220–330 GHz range using a THz microelectronics spectrometer (Figure 1), as reported by Rice et al. [34–37]. Measurements were carried out at room temperature (297 K) in a gas cell with an absorption pathlength of 21.6 cm. The spectrometer provides frequency resolution of 0.5–15 MHz, depending on operating parameters.

For the measurement of absorption spectra, THz radiation is generated using microelectronics based frequency multiplication of a radio frequency source. That radiation is coupled to free space with a diagonal horn antenna that produces a diverging beam of THz radiation. The diverging output is collimated with a Teflon lens and passed through the gas cell. Following passage through the gas cell, the radiation is focused using a second Teflon lens onto a Schottky diode detector, where the signal is captured and sent to a data acquisition system. Absorbance is determined by measuring a baseline reference intensity,  $I_0$ , with the gas cell under vacuum and then measuring the absorbed intensity,  $I$ , when the gas cell is pressurized with a gas of interest. The spectral absorbance  $A$  is given by,

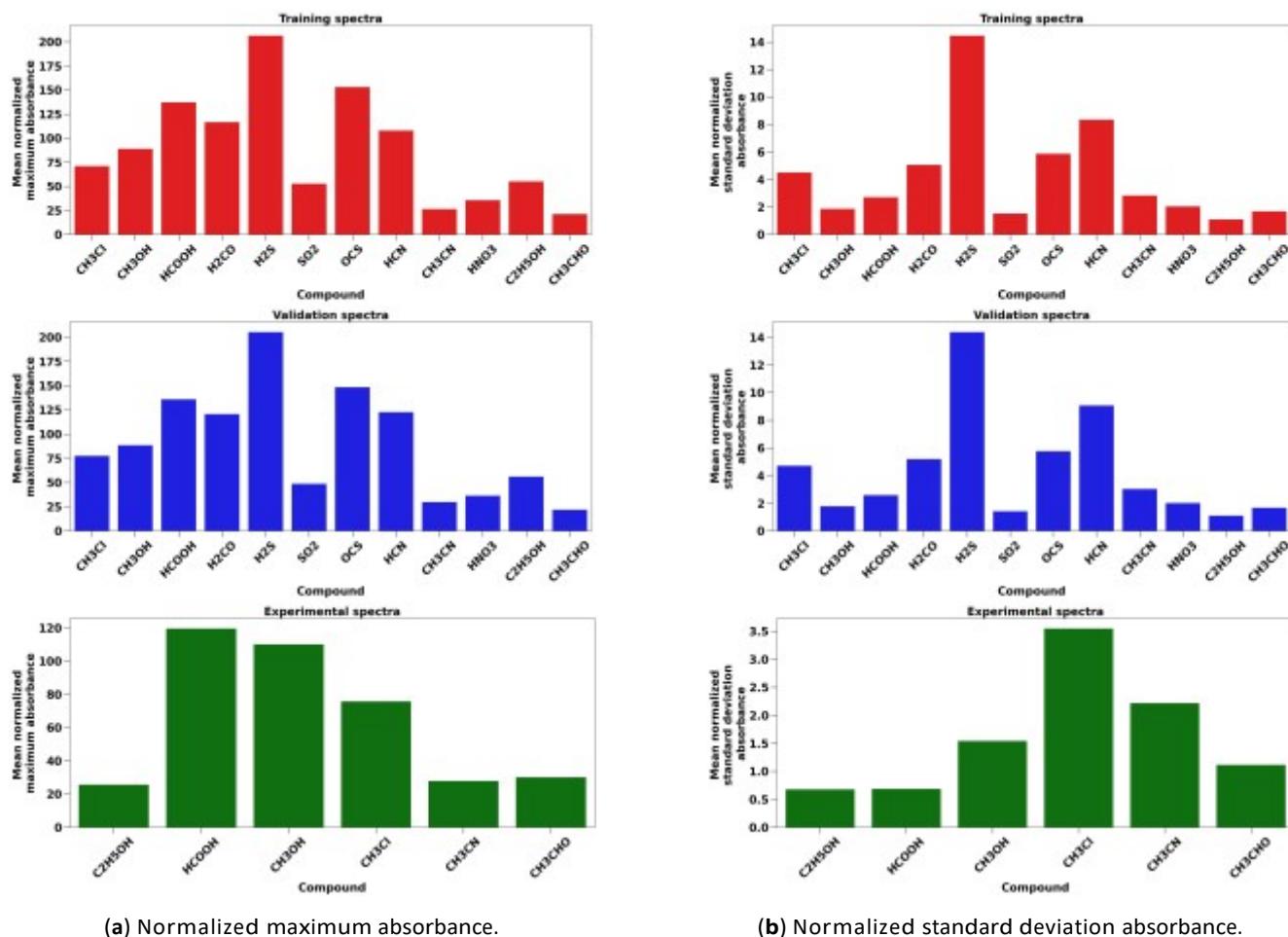
$$A = \ln\left(\frac{I}{I_0}\right) = ecl \quad (2)$$

where  $e$  is spectral absorption coefficient, which is a function of the spectroscopic parameters for the probed transitions (i.e., line position, strengths, and shapes), thermodynamic conditions, and gaseous composition;  $c$  is concentration of the absorbing gas; and  $l$  is the pathlength. The acquired experimental spectra are resampled to match the  $0.016 \text{ cm}^{-1}$  resolution of the simulated spectra. The initial frequency resolution of the experimental spectra was  $0.000016 \text{ cm}^{-1}$ , except for chloromethane which was  $0.00036 \text{ cm}^{-1}$ .

## 2.3. Data Analytics

In total, 1968 simulated spectra for twelve VOCs were generated, 70% were used for training (1377 spectra) and 30% were used for validation (591 spectra). The training and validation spectra were chosen using random stratification to ensure all molecules are equally represented in both the training and validation datasets as listed in Table 1. To demonstrate and test the model, 36 experimental spectra for the six VOC molecules were considered. In Figure 3, the number of the spectra in each dataset and their normalized maximum absorbance and normalized standard deviation in absorbance are illustrated to

show that there is no bias among molecules in training and validation, and that there is significant variation among absorbance from molecule-to-molecule and from spectrum-to-spectrum for each molecule.



**Figure 3.** Data analytics for simulated training and validation spectra and experimental spectra.

**Table 1.** Number of spectra in training, validation, and experimental datasets.

Molecule	Molecular Formula	Training Counts	Validation Counts	Experiment Counts
Chloromethane	CH <sub>3</sub> Cl	115	49	6
Methanol	CH <sub>3</sub> OH	115	49	6
Formic acid	HCOOH	114	50	6
Formaldehyde	H <sub>2</sub> CO	115	49	-
Hydrogen sulfide	H <sub>2</sub> S	115	49	-
Sulfur dioxide	SO <sub>2</sub>	115	49	-
Carbonyl sulfide	OCS	114	50	-
Hydrogen cyanide	HCN	115	49	-
Acetonitrile	CH <sub>3</sub> CN	115	49	6
Nitric acid	HNO <sub>3</sub>	115	49	-
Ethanol	C <sub>2</sub> H <sub>5</sub> OH	115	49	6
Acetaldehyde	CH <sub>3</sub> CHO	114	50	6
<b>Totals</b>		<b>1377</b>	<b>591</b>	<b>36</b>

#### 2.4. Model Development

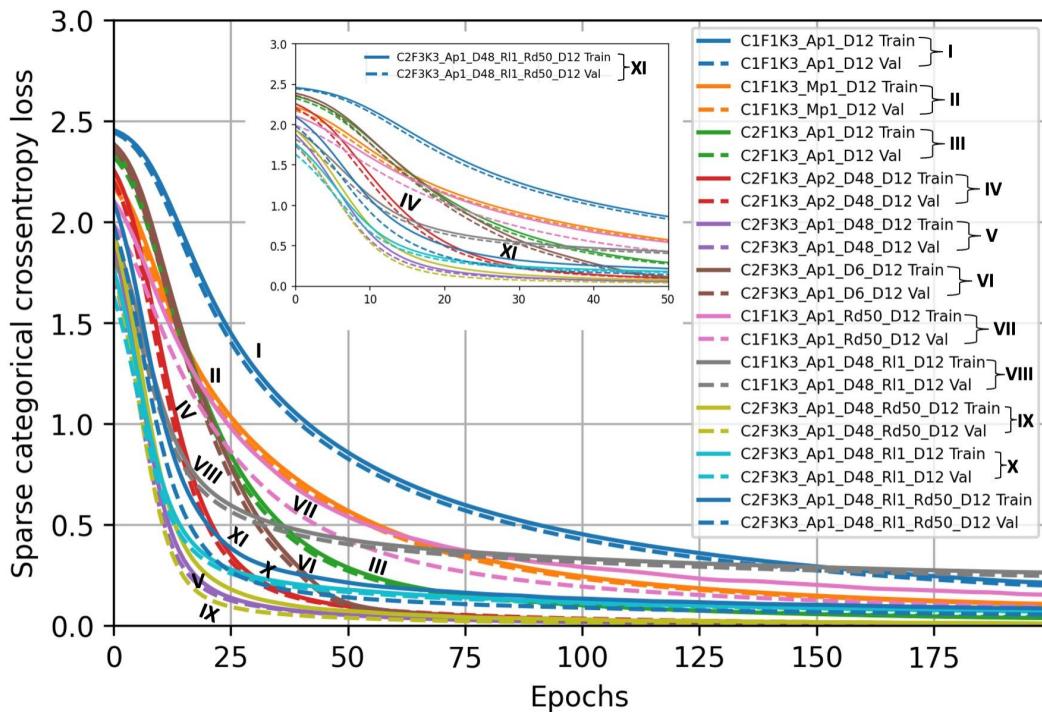
To construct a 1D CNN based classifier, we examined the performance of a number of CNN models in 70–30% training-validation studies using the simulated spectra. These

models were implemented using Tensorflow [38] in Python [39]. The training time per epoch was on average two seconds and computations were performed on a Dell 5820 Precision Tower Workstation computer with 64 GB RAM, Intel Xeon 3.6 GHz processor with NVIDIA QUADRO RTX4000 GPU.

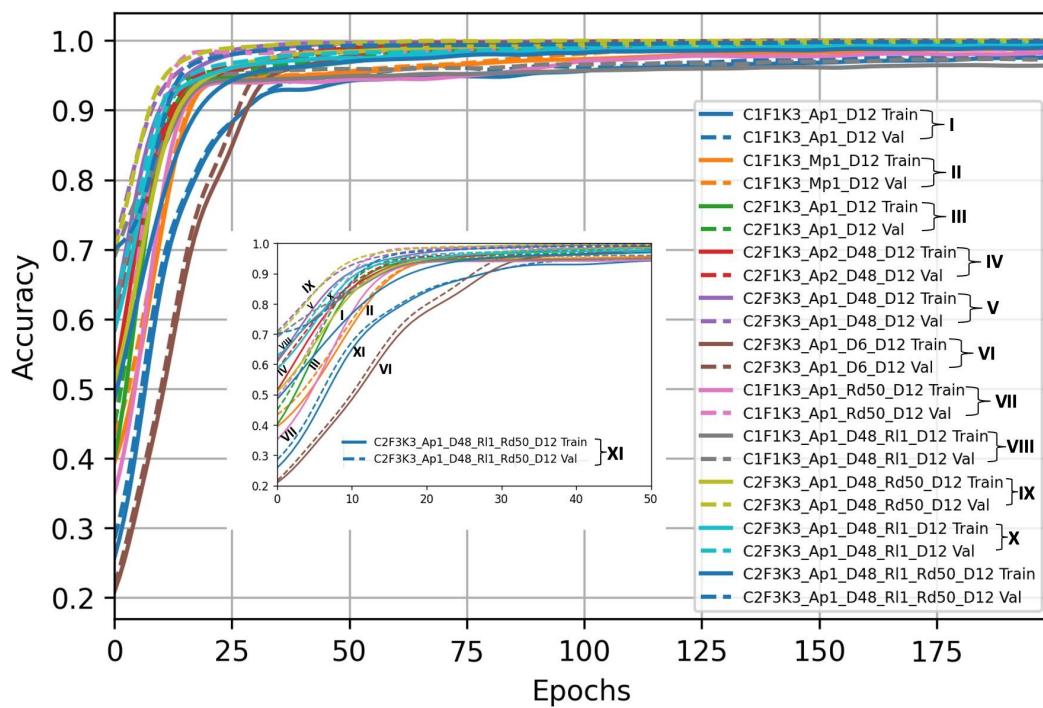
#### 2.4.1. Model Architecture

CNN models primarily consist of a feature extraction section, comprising of convolution and subsampling or pooling blocks, followed by a classifier section of dense layers. Here, twelve CNN architectures have been considered and a standardized naming convention to describe these architectures has been used as shown in Table 2. For example, C2f3k3\_AP1\_D48\_L1\_R50\_D12 refers to a deep neural network consisting of two convolutional layers each with three filters of kernel size three, one subsampling or pooling layer between the convolutional layers, followed by flattened dense layer, a hidden layer with 48 neurons and the output layer with 12 neurons corresponding to each molecule. The prefix L and R refers to L2 and dropout regularized models. Models are also identified by roman numerals in Figures 4 and 5.

The loss and accuracy on training and validation data for the twelve CNN architectures considered are shown in Figures 4 and 5. The simplest model consisting of one convolutional layer with a single filter and kernel size of three, followed by pooling, flattening, and a 12 neuron dense layer minimizes the sparse categorical crossentropy loss and achieves accuracy above 90%. The convolution and the pooling layer extracts relevant features from the input spectrum consisting of 229 absorbance values corresponding to the  $0.016\text{ cm}^{-1}$  resolution. All other model architectures are constructed incrementally and regularization is added for certain models. In constructing models, simplicity was preferred and a balance of network width and depth was maintained. The penultimate layer in each model outputs raw scores that are then converted to a probability distribution using a softmax layer.



**Figure 4.** Loss on training and validation data for the considered CNN model architectures.



**Figure 5.** Accuracy on training and validation data for the considered CNN model architectures.

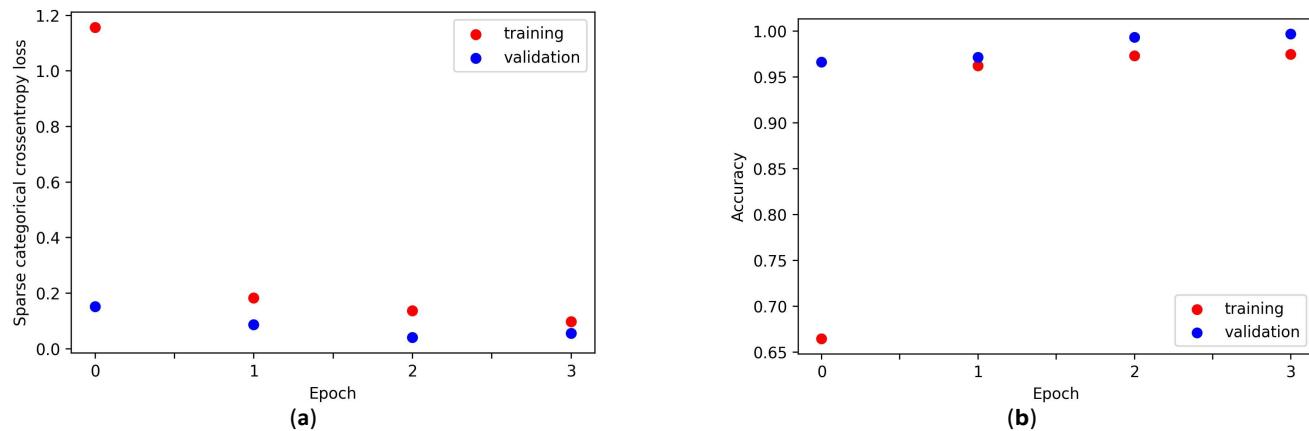
**Table 2.** Overview of CNN model architectures. The accuracy is calculated at the 200th epoch. The filter kernel size is three, the pool size is two, and padding is valid with stride of two.

Model Name	Conv. Layers	Filters	Pooling	Dense Layers	Regularization	Remarks
C1f1k3_AP1_D12 (I)	1	1	1, Average	(113,12)	-	initial model
C1f1k3_MP1_D12 (II)	1	1	1, Max	(113,12)	-	accuracy improves
C2f1k3_AP1_D12 (III)	2	1	1, Average	(111,12)	-	accuracy improves
C2f1k3_AP1_D48_D12 (not plotted)	2	1	1, Average	(111,48,12)	-	negligible improvement
C2f1k3_AP2_D48_D12 (IV)	2	1	2, Average	(55,48,12)	-	accuracy improves
C2f3k3_AP1_D48_D12 (V)	2	3	1, Average	(333,48,12)	-	accuracy improves
C2f3k3_AP1_D6_D12 (VI)	2	3	1, Average	(333,6,12)	-	accuracy worsens
C1f1k3_AP1_RD50_D12 (VII)	1	1	1, Average	(113,12)	dropout	accuracy improves
C1f1k3_AP1_D48_RL1_D12 (VIII)	1	1	1, Average	(113,48,12)	L2	accuracy worsens
C2f3k3_AP1_D48_RD50_D12 (IX)	2	3	1, Average	(333,48,12)	dropout	best accuracy, VOC-Net
C2f3k3_AP1_D48_RL1_D12 (X)	2	3	1, Average	(333,48,12)	L2	accuracy worsens
C2f3k3_AP1_D48_RL1_R50_D12 (XI)	2	3	1, Average	(333,48,12)	L2 + dropout	accuracy worsens

In Table 2, a brief description of the performance of the considered CNN models is noted in the remarks column, while max pooling improves performance, it increases the emphasis of model decision making on very strong transitions within each spectrum and, thus, weak but important transitions for defining spectral fingerprints may be ignored. Hence, we kept average pooling to preserve as much information from the raw spectrum for feature extraction. Adding additional filters, convolutional blocks, pooling layers generally improves accuracy, but is likely to result in a model that overfits the data. We also found that dropout regularization of model weights work best for this classification task compared to L2 regularized weights and combined L2 with dropout regularization. Having a second convolutional layer and/or increasing the number of filters from one to three improves the performance on the training and the validation set. After consideration of these observations regarding performance of the various CNN model architectures, the C2f3k3\_AP1\_D48\_RD50\_D12 model was selected, given its training and validation (generalization) losses are closely matched and this model does not overfit the simulated dataset, as shown in Figures 4 and 5.

#### 2.4.2. Hyperparameter Tuning and Model Training

Some of the hyperparameters in the selected CNN architecture were tuned using KerasTuner [40], including the number of filters in the convolutional layers, number of neurons in the dense layer after flattening, and the learning rate. The best epoch for training was found to be four. The final classification model, trained using the optimized hyperparameters, is shown in Figure 1. The sparse categorical crossentropy losses and classification accuracy for training and validation spectra are shown in Figure 6. Using a small epoch helps to prevent overfitting.



**Figure 6.** (a) Sparse categorical crossentropy loss and (b) accuracy for simulated training and validation spectra.

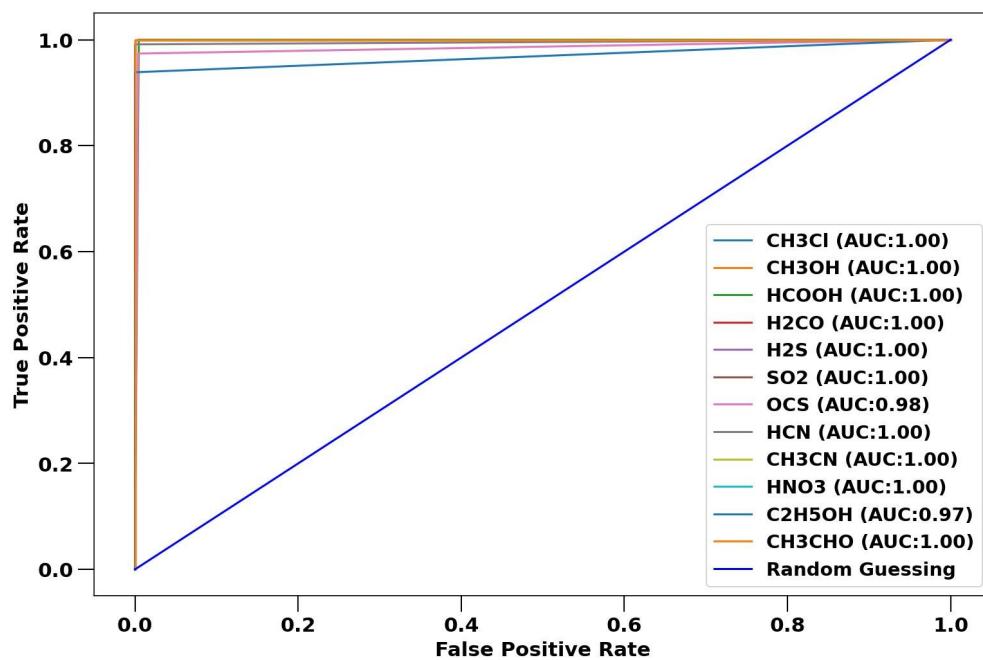
### 3. Results

The performance of the VOC-Net classifier was evaluated by consideration of confusion matrices and ROC curves. The confusion matrix and ROC curves for performance against the training spectra are shown in Figures 7 and 8. There are eleven misclassifications for the 1377 training spectra, resulting in a classification accuracy for the training dataset of 99.2%. By minimizing the sparse categorical cross-entropy loss, we have limited training to some extent, which reduces the accuracy against the training dataset. However, this approach results in better generalization for the CNN model and more robust performance in validation and in application against experimental spectra, as illustrated in Figure 6b, where the accuracy on validation data is greater than accuracy on training data, indicating that the generalization error is small compared to the training error. The misclassifications for the training spectra are for carbonyl sulfide, hydrogen cyanide, and ethanol spectra. Interestingly, the ethanol spectra is misclassified as carbonyl sulfide and formic acid. As a consequence, in the ROC curve shown in Figure 8, the ethanol and carbonyl sulfide have reduced area-under-curve (AUC) scores.

For the classification of simulated validation spectra, the VOC-Net model produces two misclassifications, resulting in a classification accuracy of 99.7% for the 591 validation spectra. The confusion matrix and ROC curves for the validation spectra are shown in Figures 9 and 10. The confusion matrix illustrates that two ethanol spectra were misclassified out of the 591 validation spectra. These misclassifications cause the reduction in the true positive rate of ethanol shown in the ROC curve in Figure 10.

Actual Molecule	CH <sub>3</sub> Cl	0 8.4%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
CH <sub>3</sub> OH	0 0.0%	115 8.4%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
HCOOH	0 0.0%	0 0.0%	115 8.4%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
H <sub>2</sub> CO	0 0.0%	0 0.0%	0 0.0%	115 8.4%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
H <sub>2</sub> S	0 0.0%	0 0.0%	0 0.0%	0 0.0%	115 8.4%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
SO <sub>2</sub>	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	115 8.4%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
OCS	0 0.0%	0 0.0%	3 0.2%	0 0.0%	0 0.0%	0 0.0%	112 8.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
HCN	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 0.1%	113 8.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
CH <sub>3</sub> CN	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	115 8.4%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
HNO <sub>3</sub>	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	115 8.4%	0 0.0%	0 0.0%	0 0.0%
C <sub>2</sub> H <sub>5</sub> OH	0 0.0%	0 0.0%	3 0.2%	0 0.0%	0 0.0%	0 0.0%	4 0.3%	0 0.0%	0 0.0%	0 0.0%	107 7.8%	0 0.0%	0 0.0%
CH <sub>3</sub> CHO	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	114 8.3%	0 0.0%
	CH <sub>3</sub> Cl	CH <sub>3</sub> OH	HCOOH	H <sub>2</sub> CO	H <sub>2</sub> S	SO <sub>2</sub>	OCS	HCN	CH <sub>3</sub> CN	HNO <sub>3</sub>	C <sub>2</sub> H <sub>5</sub> OH	CH <sub>3</sub> CHO	
	Predicted Molecule												

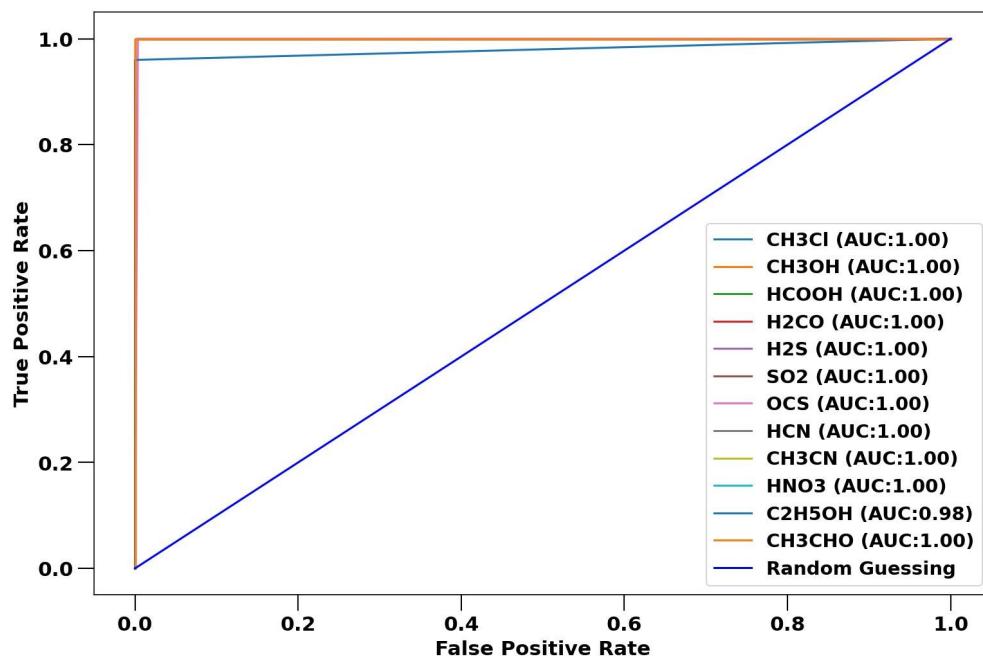
**Figure 7.** Confusion matrix for classification of 1377 simulated training spectra. Red boxes indicate misclassifications.



**Figure 8.** ROC curve for the classification of 1377 simulated training spectra.

Actual Molecule	CH <sub>3</sub> Cl	CH <sub>3</sub> OH	HCOOH	H <sub>2</sub> CO	H <sub>2</sub> S	SO <sub>2</sub>	OCS	HCN	CH <sub>3</sub> CN	HNO <sub>3</sub>	C <sub>2</sub> H <sub>5</sub> OH	CH <sub>3</sub> CHO
CH <sub>3</sub> Cl	49 8.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
CH <sub>3</sub> OH	0 0.0%	49 8.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
HCOOH	0 0.0%	0 0.0%	49 8.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
H <sub>2</sub> CO	0 0.0%	0 0.0%	0 0.0%	49 8.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
H <sub>2</sub> S	0 0.0%	0 0.0%	0 0.0%	0 0.0%	49 8.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
SO <sub>2</sub>	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	49 8.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
OCS	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	49 8.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
HCN	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	50 8.5%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
CH <sub>3</sub> CN	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	49 8.3%	0 0.0%	0 0.0%	0 0.0%
HNO <sub>3</sub>	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	49 8.3%	0 0.0%	0 0.0%
C <sub>2</sub> H <sub>5</sub> OH	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	2 0.3%	0 0.0%	0 0.0%	0 0.0%	48 8.1%	0 0.0%
CH <sub>3</sub> CHO	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	50 8.5%

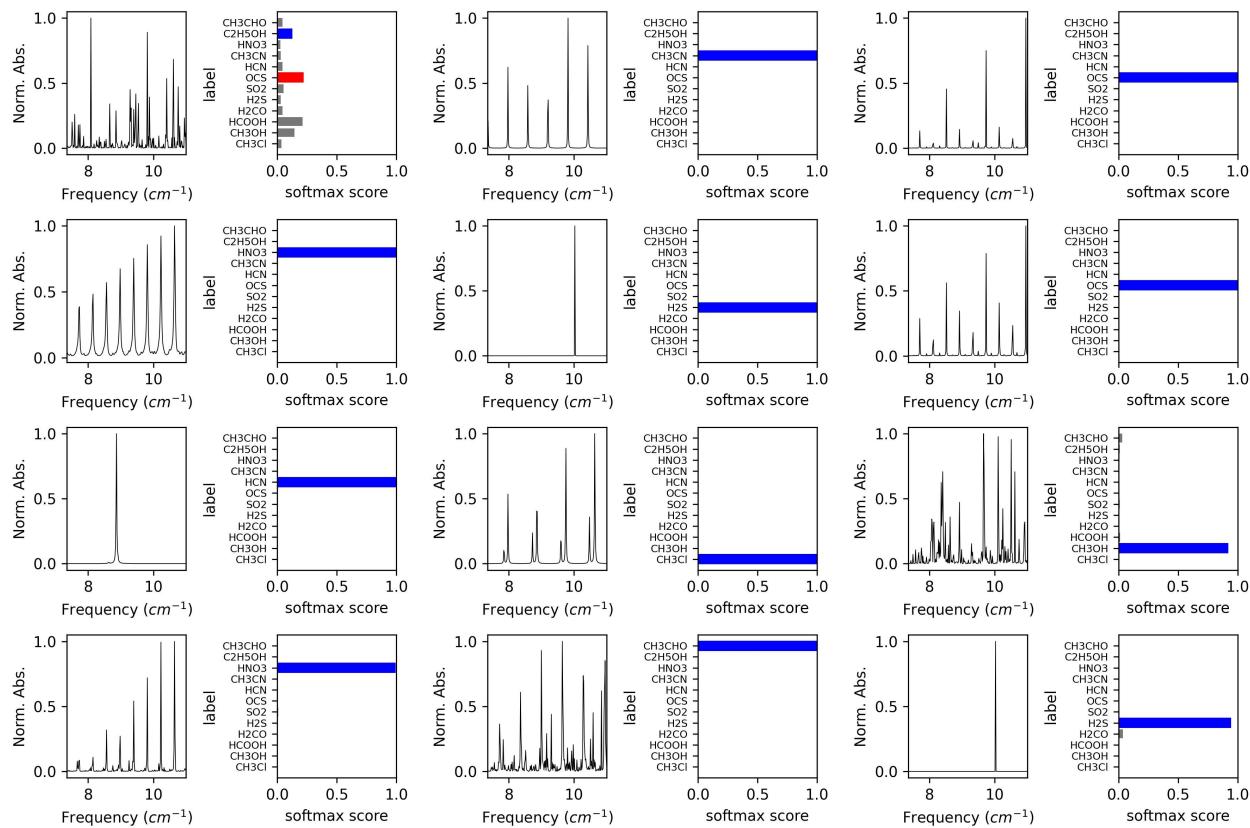
**Figure 9.** Confusion matrix for the classification of 591 simulated validation spectra. Red boxes indicate misclassifications.



**Figure 10.** ROC curve for the classification of 591 simulated validation spectra.

In Figure 11, we observe twelve spectra sampled from the validation dataset. One of the two misclassifications is illustrated in Figure 11, a misclassification for an ethanol spectra at 4.4 Torr. The rest of the sampled Figure 11 spectra are correctly classified. Based on the softmax scores, the model is performing very well, with softmax scores for individual classifications of approximately one. The softmax scores represent the probability that a

spectra belongs to one of the considered molecules; hence, the softmax scores near unity indicate that the model has a high degree of confidence in these classifications.



**Figure 11.** Sampled validation spectra and their corresponding VOC-Net softmax scores.

The misclassified ethanol spectra shown in Figure 11 (top left) is very complex. Based on the softmax scores for this case, it is evident that the model is confused by this complexity and is uncertain as to whether the spectra belongs to methanol, carbonyl sulfide, ethanol, or formic acid. All four of these molecules have complex spectra with hundreds of significant transitions in the present spectral region. Of course, the number of epochs can be increased, additional filters can be added, and/or additional convolutional blocks can be added to increase the classification accuracy and avoid model confusion for these complex spectra, but that runs the risk of building a model that overfits the spectral dataset and will not perform well for the classification of noisy experimental spectra. The spectra for methanol in the third row and the fifth column is similarly complex but it is correctly classified but with the model assigning a small probability that this spectra could belong to acetaldehyde. Thus, classification decisions with softmax scores, as illustrated, offer insight into the certainty of model predictions and encourages acceptance of classification models like VOC-Net. Since the training and validation spectra are simulated and do not contain noise, we can conclude that the misclassifications for the validation spectra arise from similar transitions present in very complex spectra for more than one molecule, and are not from overfitting. We conclude that VOC-Net is not overfitting, since similar training and validation accuracies were achieved and the model was trained with only four epochs. The smaller number of epochs essentially implements an early stopping strategy that prevents overfitting. Ensuring the model is not overfitting does not preclude misclassifications; however, it makes the model capable of generalization.

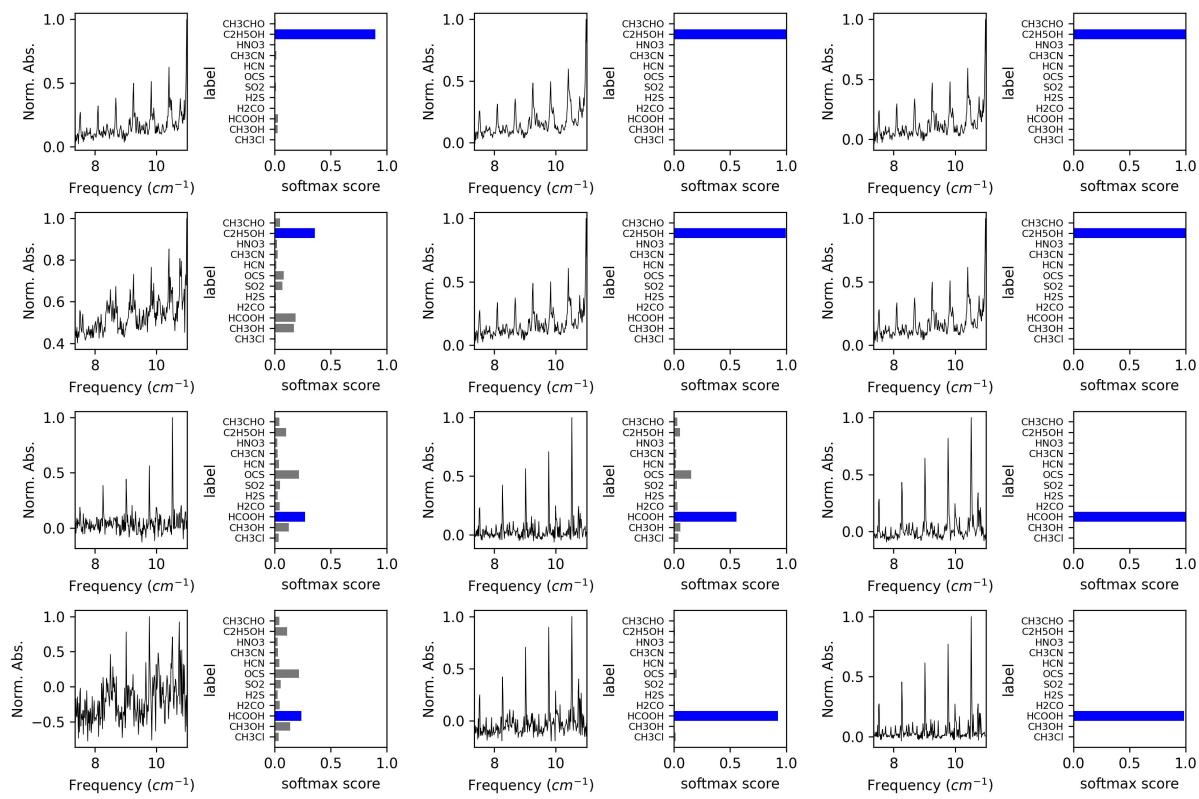
Thirty-six experimental spectra measured as described above were used to demonstrate and test the model. The softmax scores produced by the VOC-Net for these experimental spectra and their subsequent classifications are given in Figures 12–14. Against the experimental dataset, the model yields one misclassification. The misclassified spectra

is for methanol at 1 Torr. The details of the measurement conditions of the experimental spectra are given in Table 3. We observe that all of the spectra contain some amount of noise (noise floor is on average 0.001 in absorbance space) and other experiment-to-experiment variations and anomalies. For example, some measurements exhibit a subtle upward drift in measured absorbance at higher frequencies that is a purely experimental artifact resulting from intensity baseline drift in the radiation source.

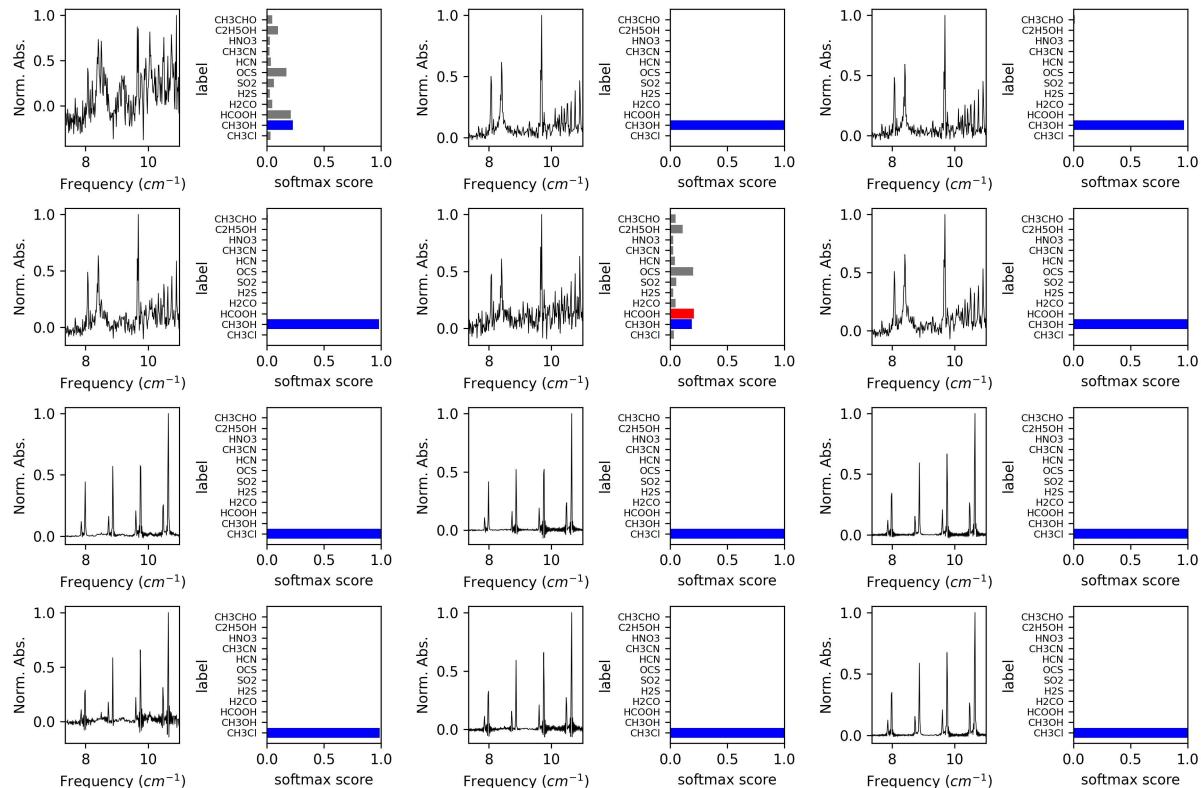
Despite the noise and other aberrations not found in the training datasets, VOC-Net performs very well against the experimental dataset, yielding a 97.2% classification accuracy. The strong performance against experiments illustrates that the VOC-Net has managed to generalize and learn to distinguish the salient spectral features belonging to each molecule. However, when the noise floor is large relative to the signal, the softmax scores illustrate that the model has lower confidence and offers probabilities that the spectra under consideration may belong to multiple molecules. For the single misclassification (spectra no. 17), we observe that the model marginally misses the correct classification (methanol), confusing the methanol spectra with that of formic acid and carbonyl sulfide. Note, similar uncertainty was shown by the model with respect to methanol, carbonyl sulfide, and formic acid for simulated validation spectra. Due to this particular misclassification, in the ROC curves shown in Figure 15, the AUC scores for formic acid and methanol are reduced. The overall AUC score for the six molecule classes that make up the experimental spectra is 0.983, which is very similar to the training and validation AUC scores of 0.996 and 0.998, respectively. The overall classification accuracy for the VOC-Net against experimental data (97.2%) is somewhat better than the performance of a support vector machine (SVM) classifier we have previously reported, that under similar conditions produced an accuracy of 93.5% against experiments [17].

**Table 3.** Experiment conditions. The experimental spectra are sequentially plotted from left to right in Figures 12–14.

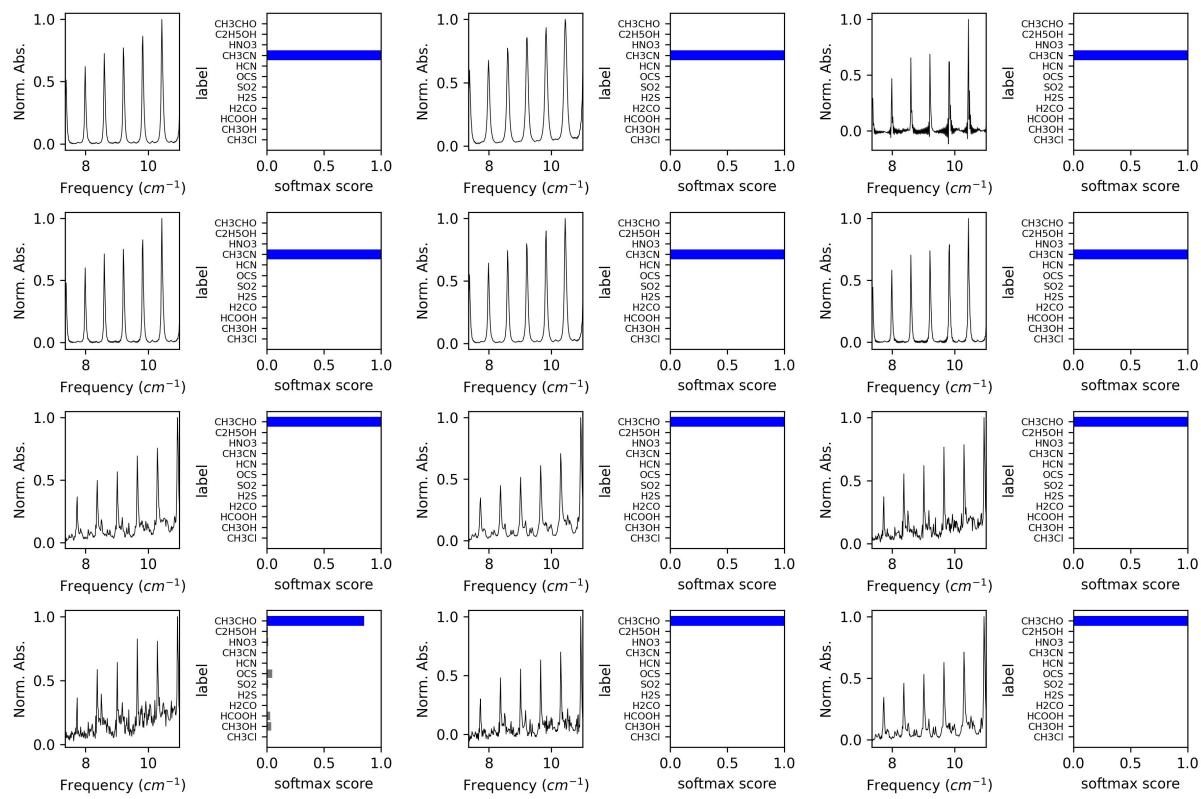
Exp. Spectrum No.	Molecule	Pressure (Torr)	Exp. Spectrum No.	Molecule	Pressure (Torr)
1	Ethanol	2	19	Chloromethane	8
2	Ethanol	16	20	Chloromethane	1
3	Ethanol	8	21	Chloromethane	5
4	Ethanol	1	22	Chloromethane	0.5
5	Ethanol	4	23	Chloromethane	1
6	Ethanol	8	24	Chloromethane	10
7	Formic acid	1	25	Acetonitrile	4
8	Formic acid	2	26	Acetonitrile	16
9	Formic acid	16	27	Acetonitrile	0.5
10	Formic acid	1	28	Acetonitrile	2
11	Formic acid	4	29	Acetonitrile	8
12	Formic acid	4	30	Acetonitrile	1
13	Methanol	1	31	Acetaldehyde	2
14	Methanol	4	32	Acetaldehyde	8
15	Methanol	2	33	Acetaldehyde	1
16	Methanol	2	34	Acetaldehyde	0.5
17	Methanol	1	35	Acetaldehyde	1
18	Methanol	8	36	Acetaldehyde	2



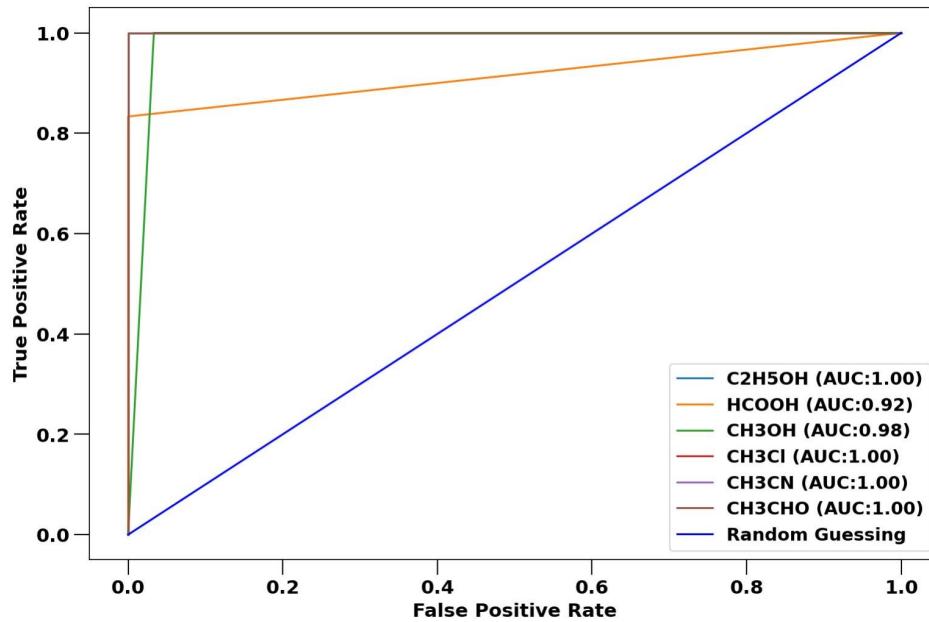
**Figure 12.** VOC-Net classification of experimental spectra (1–12). The corresponding softmax scores for each classification are given with each spectra.



**Figure 13.** VOC-Net classification of experimental spectra (13–24). The corresponding softmax scores for each classification are given with each spectra.



**Figure 14.** VOC-Net classification of experimental spectra (25–36). The corresponding softmax scores for each classification are given with each spectra.

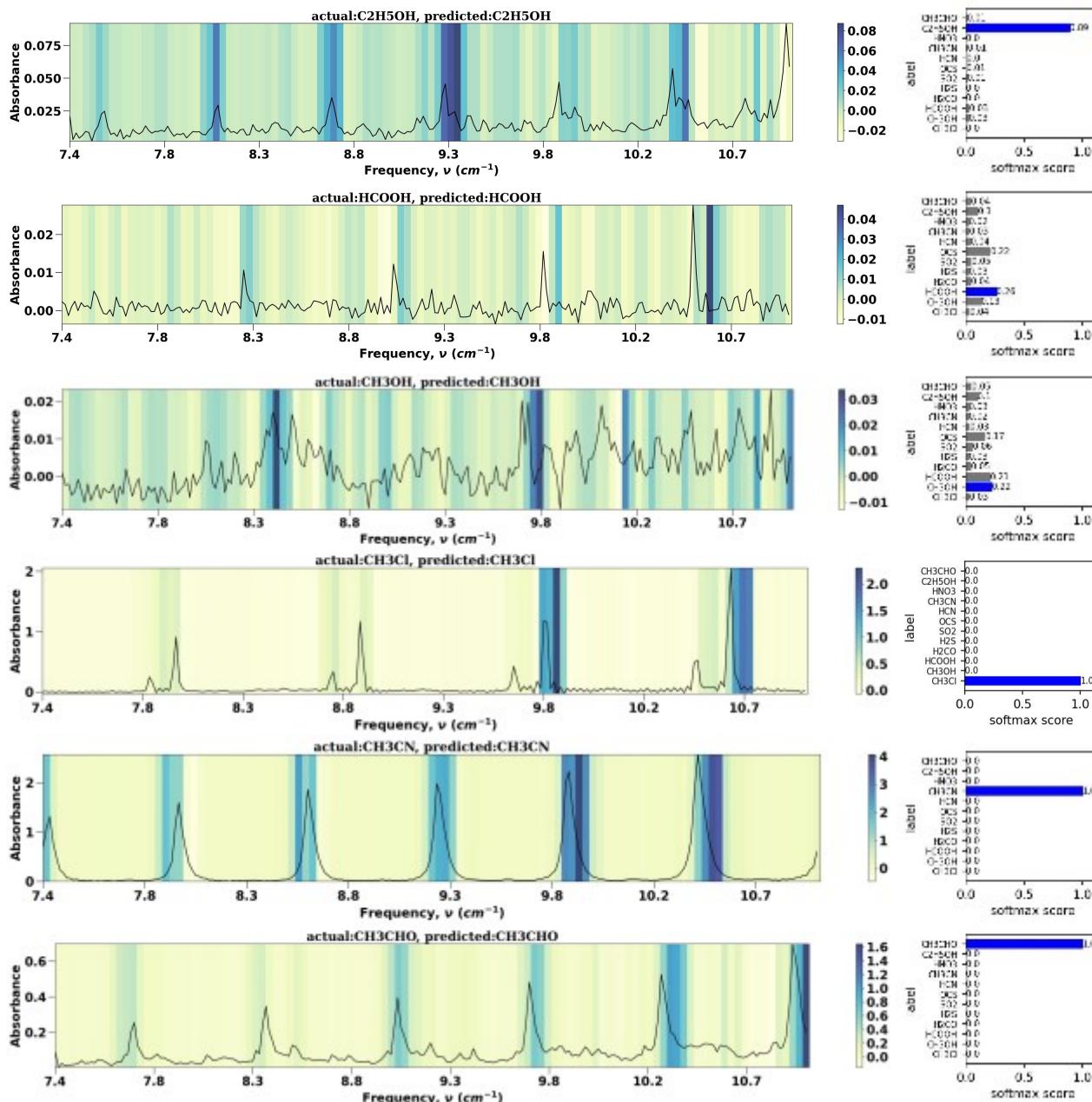


**Figure 15.** ROC curve for the classification of 36 experimental spectra.

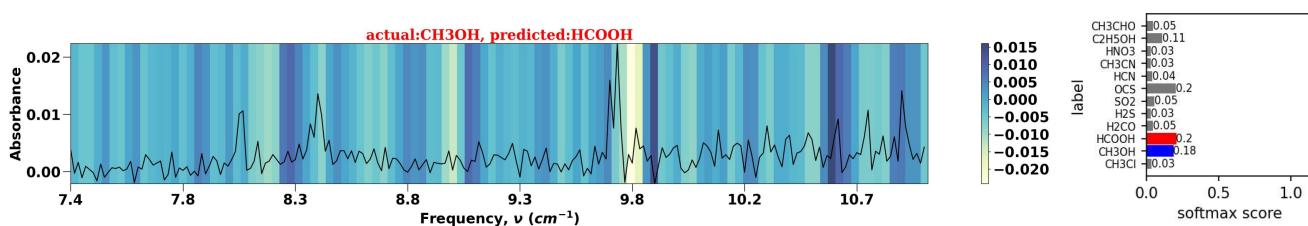
#### 4. Discussion

The VOC-Net is further examined using the Grad-CAM method [29], which produces a heatmap superimposed on the raw spectral input to visualize and elucidate the model's decision making process. The gradients of each target molecule class from the inputs to the last convolutional layer (in our case the second convolution layer) are averaged. Then the gradient multipliers are summed with the corresponding feature maps to generate the

heatmap. Localized spectral peaks prioritized by the model for classification are easily visualized and illustrate the molecule-discriminating regions of the spectra. The class activation maps for six experimental spectra correctly classified by VOC-Net are shown in Figure 16. The class activation map for the single misclassification of the methanol spectra at 1 Torr, which VOC-Net incorrectly predicted to be formic acid, is shown in Figure 17. Figures 16 and 17 depicts the softmax scores corresponding to each spectrum next to the class activation maps.



**Figure 16.** Class activation maps for the classification of six experimental spectra (No. 1, 7, 13, 19, 25, and 31). All of these spectra were correctly classified. The corresponding softmax scores are shown next to each class activation map.



**Figure 17.** Class activation map for the classification of experimental spectra number 17. The spectra is for methanol at 1 Torr but is misclassified as formic acid. The corresponding softmax scores are shown next to the class activation map.

For accurate and confident classification (high softmax scores), the Grad-CAM results illustrate that distinguishing transitions, characterized by strong absorption peaks, should be activated, where activation refers to the features that the model prioritizes in its classification decisions. Furthermore, weak transitions and noise fluctuations should have low activation. The model misclassifies when large numbers of weak absorption features within the noise floor causes activation. Both methanol and formic acid have a large number of absorption peaks in the considered frequency range. In the activation map for the misclassified methanol spectrum (Figure 17, misclassified as formic acid), we observe that the major transitions near  $8.4\text{ cm}^{-1}$  and  $9.8\text{ cm}^{-1}$  are not well activated. The class activation maps also provide valuable insight regarding the identification of the necessary or preferred features for spectral classification. Thus, in terms of selecting spectroscopic transitions or frequency ranges for sensor design, the Grad-CAM method may act as a guiding tool.

## 5. Conclusions

A deep learning neural network, VOC-Net, has been developed and demonstrated for the automated identification of VOCs based on their fingerprint rotational absorption spectra in the 220–330 GHz frequency range. VOC-Net performed at an greater than 99% against simulated spectra and 97% against experimental spectra. The architecture of the CNN-based VOC-Net and model and data parameters affecting its performance have been evaluated. Grad-CAM class activation maps for the CNN provide valuable insight into the decision making process of the network, enabling acceptance and trust from potential users. The model demonstrates a high degree of accuracy, selectivity, and sensitivity for simulated spectra for twelve VOCs of interest in remote gas sensing in industrial applications and for six VOCs for which prior experimental data is available. The model further demonstrated the importance of specific spectral features which it prioritizes for making classification decision. VOC-Net maybe used for classification of spectra and sensor design (transition or frequency range selection), enabling future selective and cross-sensitive automated THz remote gas sensing. The VOC-Net model reported here may be extended to different frequency regions, types of spectroscopy, or other spectral classification problems, provided suitable training datasets exist or can be generated.

**Author Contributions:** Conceptualization, M.A.Z.C. and M.A.O.; methodology, M.A.Z.C. and T.E.R.; software, M.A.Z.C.; validation, M.A.Z.C. and T.E.R.; formal analysis, M.A.Z.C.; investigation, M.A.Z.C.; data curation, M.A.Z.C. and T.E.R.; writing—original draft preparation, M.A.O.; writing—review and editing, M.A.Z.C. and M.A.O.; supervision, M.A.O.; project administration, M.A.O.; funding acquisition, M.A.O. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Science Foundation under Grant CBET-1851291.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The code and data used in this paper can be found at <https://github.com/arshadzahangirchowdhury/VOC-Net> (accessed on 20 July 2022).

**Acknowledgments:** The authors acknowledges the support of the National Science Foundation under Grant CBET-1851291.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
ML	Machine Learning
CNN	Convolutional Neural Network
DNN	Deep Neural Network
SVM	Support Vector Machines
VOC	Volatile Organic Compounds
Grad-CAM	Gradient-weighted Class Activation Mapping
CAM	Class Activation Map

## References

1. Zulkifli, M.F.H.; Hawari, N.S.S.L.; Latif, M.T.; Abd Hamid, H.H.; Mohtar, A.A.A.; Idris, W.M.R.W.; Mustaffa, N.I.H.; Juneng, L. Volatile organic compounds and their contribution to ground-level ozone formation in a tropical urban environment. *Chemosphere* **2022**, *302*, 134852. [[CrossRef](#)] [[PubMed](#)]
2. Zheng, H.; Kong, S.; Chen, N.; Niu, Z.; Zhang, Y.; Jiang, S.; Yan, Y.; Qi, S. Source apportionment of volatile organic compounds: Implications to reactivity, ozone formation, and secondary organic aerosol potential. *Atmos. Res.* **2021**, *249*, 105344. [[CrossRef](#)]
3. Hewitt, C.N. *Reactive Hydrocarbons in the Atmosphere*; Elsevier: Amsterdam, The Netherlands, 1998.
4. Paul, S.; Bari, M.A. Elucidating sources of VOCs in the Capital Region of New York State: Implications to secondary transformation and public health exposure. *Chemosphere* **2022**, *299*, 134407. [[CrossRef](#)] [[PubMed](#)]
5. McDonald, B.C.; De Gouw, J.A.; Gilman, J.B.; Jathar, S.H.; Akherati, A.; Cappa, C.D.; Jimenez, J.L.; Lee-Taylor, J.; Hayes, P.L.; McKeen, S.A.; et al. Volatile chemical products emerging as largest petrochemical source of urban organic emissions. *Science* **2018**, *359*, 760–764. [[CrossRef](#)] [[PubMed](#)]
6. Schmalz, K.; Rothbart, N.; Neumaier, P.F.X.; Borngräber, J.; Hübers, H.W.; Kissinger, D. Gas spectroscopy system for breath analysis at mm-wave/THz using SiGe BiCMOS circuits. *IEEE Trans. Microw. Theory Tech.* **2017**, *65*, 1807–1818. [[CrossRef](#)]
7. Rothbart, N.; Stanley, V.; Koczulla, R.; Jarosch, I.; Holz, O.; Schmalz, K.; Huebers, H.W. Millimeter-wave gas spectroscopy for breath analysis of COPD patients in comparison to GC-MS. *J. Breath Res.* **2022**, *16*, 046001. [[CrossRef](#)]
8. Chevalier, P.; Meister, T.; Heinemann, B.; Van Huylenbroeck, S.; Liebl, W.; Fox, A.; Sibaja-Hernandez, A.; Chantre, A. Towards thz sige hbts. In Proceedings of the 2011 IEEE Bipolar/BiCMOS Circuits and Technology Meeting, Atlanta, GA, USA, 9–11 October 2011; IEEE: Piscataway, NJ, USA, 2011; pp. 57–65.
9. Mansha, M.W.; Wu, K.; Rice, T.E.; Oehlschlaeger, M.A.; Hella, M.M.; Wilke, I. Detection of volatile organic compounds using a single transistor terahertz detector implemented in standard BiCMOS technology. In Proceedings of the 2019 IEEE SENSORS, Montreal, QC, Canada, 27–30 October 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 1–4.
10. Wang, C.; Han, R. Dual-terahertz-comb spectrometer on CMOS for rapid, wide-range gas detection with absolute specificity. *IEEE J. Solid-State Circuits* **2017**, *52*, 3361–3372. [[CrossRef](#)]
11. Wang, D.; Schmalz, K.; Eissa, M.H.; Borngräber, J.; Kucharski, M.; Elkhouly, M.; Ko, M.; Ng, H.J.; Kissinger, D. Integrated 240-GHz dielectric sensor with DC readout circuit in a 130-nm SiGe BiCMOS technology. *IEEE Trans. Microw. Theory Tech.* **2018**, *66*, 4232–4241. [[CrossRef](#)]
12. Schmalz, K.; Rothbart, N.; Eissa, M.; Borngräber, J.; Kissinger, D.; Hübers, H.W. Transmitters and receivers in SiGe BiCMOS technology for sensitive gas spectroscopy at 222–270 GHz. *AIP Adv.* **2019**, *9*, 015213. [[CrossRef](#)]
13. Neese, C.F.; Medvedev, I.R.; Plummer, G.M.; Frank, A.J.; Ball, C.D.; De Lucia, F.C. Compact submillimeter/terahertz gas sensor with efficient gas collection, preconcentration, and ppt sensitivity. *IEEE Sens. J.* **2012**, *12*, 2565–2574. [[CrossRef](#)]
14. Naftaly, M.; Vieweg, N.; Deninger, A. Industrial applications of terahertz sensing: State of play. *Sensors* **2019**, *19*, 4203. [[CrossRef](#)] [[PubMed](#)]
15. Galstyan, V.; D’Arco, A.; Di Fabrizio, M.; Poli, N.; Lupi, S.; Comini, E. Detection of volatile organic compounds: From chemical gas sensors to terahertz spectroscopy. *Rev. Anal. Chem.* **2021**, *40*, 33–57. [[CrossRef](#)]
16. Medvedev, I.R.; Schueler, R.; Thomas, J.; Kenneth, O.; Nam, H.J.; Sharma, N.; Zhong, Q.; Lary, D.J.; Raskin, P. Analysis of exhaled human breath via terahertz molecular spectroscopy. In Proceedings of the 2016 41st International Conference on Infrared, Millimeter, and Terahertz waves (IRMMW-THz), Copenhagen, Denmark, 25–30 September 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 1–2.
17. Chowdhury, M.; Rice, T.E.; Oehlschlaeger, M.A. Evaluation of machine learning methods for classification of rotational absorption spectra for gases in the 220–330 GHz range. *Appl. Phys. B* **2021**, *127*, 34. [[CrossRef](#)]

18. Liu, Y.; Upadhyaya, B.R.; Naghedi Feizi, M. Chemometric data analysis using artificial neural networks. *Appl. Spectrosc.* **1993**, *47*, 12–23. [\[CrossRef\]](#)

19. Lussier, F.; Thibault, V.; Charron, B.; Wallace, G.Q.; Masson, J.F. Deep learning and artificial intelligence methods for Raman and surface-enhanced Raman scattering. *TrAC Trends Anal. Chem.* **2020**, *124*, 115796. [\[CrossRef\]](#)

20. Ghosh, K.; Stuke, A.; Todorović, M.; Jørgensen, P.B.; Schmidt, M.N.; Vehtari, A.; Rinke, P. Deep learning spectroscopy: Neural networks for molecular excitation spectra. *Adv. Sci.* **2019**, *6*, 1801367. [\[CrossRef\]](#)

21. Guidotti, R.; Monreale, A.; Ruggieri, S.; Turini, F.; Giannotti, F.; Pedreschi, D. A survey of methods for explaining black box models. *ACM Comput. Surv. (CSUR)* **2018**, *51*, 1–42. [\[CrossRef\]](#)

22. Kiranyaz, S.; Avci, O.; Abdeljaber, O.; Ince, T.; Gabbouj, M.; Inman, D.J. 1D convolutional neural networks and applications: A survey. *Mech. Syst. Signal Process.* **2021**, *151*, 107398. [\[CrossRef\]](#)

23. Rawat, W.; Wang, Z. Deep convolutional neural networks for image classification: A comprehensive review. *Neural Comput.* **2017**, *29*, 2352–2449. [\[CrossRef\]](#)

24. LeCun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **1998**, *86*, 2278–2324. [\[CrossRef\]](#)

25. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.

26. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.

27. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015; Springer: Berlin/Heidelberg, Germany, 2015; pp. 234–241.

28. Khan, A.; Sohail, A.; Zahoor, U.; Qureshi, A.S. A survey of the recent architectures of deep convolutional neural networks. *Artif. Intell. Rev.* **2020**, *53*, 5455–5516. [\[CrossRef\]](#)

29. Selvaraju, R.R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. Grad-cam: Visual explanations from deep networks via gradient-based localization. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 618–626.

30. Wang, C.Y.; Ko, T.S.; Hsu, C.C. Interpreting convolutional neural network for real-time volatile organic compounds detection and classification using optical emission spectroscopy of plasma. *Anal. Chim. Acta* **2021**, *1179*, 338822. [\[CrossRef\]](#) [\[PubMed\]](#)

31. Gordon, I.E.; Rothman, L.S.; Hill, C.; Kochanov, R.V.; Tan, Y.; Bernath, P.F.; Birk, M.; Boudon, V.; Campargue, A.; Chance, K.; et al. The HITRAN2016 molecular spectroscopic database. *J. Quant. Spectrosc. Radiat. Transf.* **2017**, *203*, 3–69. [\[CrossRef\]](#)

32. Pickett, H.; Poynter, R.; Cohen, E.; Delitsky, M.; Pearson, J.; Müller, H. Submillimeter, millimeter, and microwave spectral line catalog. *J. Quant. Spectrosc. Radiat. Transf.* **1998**, *60*, 883–890. [\[CrossRef\]](#)

33. Kochanov, R.V.; Gordon, I.; Rothman, L.; Wcisło, P.; Hill, C.; Wilzewski, J. HITRAN Application Programming Interface (HAPI): A comprehensive approach to working with spectroscopic data. *J. Quant. Spectrosc. Radiat. Transf.* **2016**, *177*, 15–30. [\[CrossRef\]](#)

34. Rice, T.E.; Chowdhury, M.; Mansha, M.W.; Hella, M.M.; Wilke, I.; Oehlschlaeger, M.A. VOC gas sensing via microelectronics-based absorption spectroscopy at 220–330 GHz. *Appl. Phys. B* **2020**, *126*, 152. [\[CrossRef\]](#)

35. Rice, T.E.; Mansha, M.W.; Chowdhury, A.; Hella, M.M.; Wilke, I.; Oehlschlaeger, M.A. All Electronic THz Wave Absorption Spectroscopy of Volatile Organic Compounds Between 220–330 GHz. In Proceedings of the 2020 45th International Conference on Infrared, Millimeter, and Terahertz Waves (IRMMW-THz), Buffalo, NY, USA, 8–13 November 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 1–2.

36. Rice, T.E.; Chowdhury, M.A.Z.; Mansha, M.W.; Hella, M.M.; Wilke, I.; Oehlschlaeger, M.A. Halogenated hydrocarbon gas sensing by rotational absorption spectroscopy in the 220–330 GHz frequency range. *Appl. Phys. B* **2021**, *127*, 123. [\[CrossRef\]](#)

37. Rice, T.E.; Chowdhury, M.A.Z.; Powers, M.N.; Mansha, M.W.; Wilke, I.; Hella, M.M.; Oehlschlaeger, M.A. Gas Sensing for Industrial Relevant Nitrogen-Containing Compounds Using a Microelectronics-Based Absorption Spectrometer in the 220 to 330 GHz Frequency Range. *Sens. Actuators B Chem.* **2022**, *367*, 132030. [\[CrossRef\]](#)

38. Abadi, M.; Agarwal, A.; Barham, P.; Brevdo, E.; Chen, Z.; Citro, C.; Corrado, G.S.; Davis, A.; Dean, J.; et al. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. *arXiv* **2016**, arXiv:1603.04467.

39. Van Rossum, G.; Drake F.L., Jr. Python Reference Manual; Centrum voor Wiskunde en Informatica: Amsterdam, The Netherlands, 1995.

40. O’Malley, T.; Bursztein, E.; Long, J.; Chollet, F.; Jin, H.; Invernizzi, L. KerasTuner. 2019. Available online: <https://github.com/keras-team/keras-tuner> (accessed on 20 July 2022).