



Article

# Time Series Classification with Multiple Wavelength Scattering Signals for Nuisance Alarm Mitigation

Kyuwon Han \* Soocheol Kim, Hoesung Yang, Kwangsoo Cho and Kangbok Lee

Defense & Safety ICT Research Department, Electronics and Telecommunications Research Institute, Daejeon 34129, Republic of Korea; skim88@etri.re.kr (S.K.); hsyang@etri.re.kr (H.Y.); choks@etri.re.kr (K.C.); kblee@etri.re.kr (K.L.)

\* Correspondence: wally.han@etri.re.kr; Tel.: +82-10-2659-0896

**Abstract:** Smoke detectors are the most widely used fire detectors due to their high sensitivity. However, they have persistently faced issues with false alarms, known as nuisance alarms, as they cannot distinguish smoke particles, and their responsiveness varies depending on the particle size and concentration. Although technologies for distinguishing smoke particles have shown promising results, the hardware limitations of smoke detectors necessitate an intelligent approach to analyze scattering signals of various wavelengths and their temporal changes. In this paper, we propose a pipeline that can distinguish smoke particles based on scattering signals of various wavelengths as input. In the data extraction phase, we propose methods for extracting datasets from time series data. We propose a method that combines traditional approaches, early detection methods, and a Dynamic Time Warping technique that utilizes only the shape of the signal without preprocessing. In the learning model and classification phase, we present a method to select and compare various architectures and hyperparameters to create a model that achieves the best classification performance for time series data. We create datasets for six different targets in our presented sensor and smoke particle test environment and train classification models. Through performance comparisons, we identify architecture and parameter combinations that achieve up to 98.7% accuracy. Ablation studies under various conditions demonstrate the validity of the chosen architecture and the potential of other models.



**Citation:** Han, K.; Kim, S.; Yang, H.; Cho, K.; Lee, K. Time Series Classification with Multiple Wavelength Scattering Signals for Nuisance Alarm Mitigation. *Fire* **2024**, *7*, 14. <https://doi.org/10.3390/fire7010014>

Academic Editors: António Vieira, Marcos Francos and Washington Rocha

Received: 18 November 2023

Revised: 21 December 2023

Accepted: 26 December 2023

Published: 29 December 2023



**Copyright:** © 2023 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

Fire detectors are the most commonly used automated fire detection devices in everyday life. In particular, smoke detectors have a simple structure, where infrared light is scattered when smoke particles enter the detector, and this scattered light is detected by a photodiode, generating signals to trigger alarms in various forms. Due to their high sensitivity, straightforward operating principles, and cost effectiveness, optical devices have become popular detectors [1]. However, due to their high sensitivity, they suffer from frequent errors, as they may not distinguish smoke particles, and their responsiveness varies depending on the particle size and concentration. To address this issue, UL (Underwriters Laboratories) has introduced UL-268 standards for all fire detectors and defined false alarms referred to as nuisance alarms [2].

Numerous studies on aerosol particles have been conducted using optical systems, especially on four major issues: particle distribution, multiple wavelengths, polarization and scattering angles.

A correlation method of selecting an optimal set of wavelengths and scattering angles was presented in [3], in which Q. Zhang et al., extracted an indicator for discrimination between fire and non-fire aerosols using scattering matrix analysis in polarized light scattering [4]. T. Deng et al., designed a sensor to measure surface area concentration with dual

wavelength sources [5]. Methods for the testing of optical smoke detectors in dusty environments was presented in [6]. Recently, there has been research conducted on reducing false alarms by utilizing the scattering characteristics of various wavelengths. K. Li et al. [7] conducted research to reduce false alarms by measuring the scattering coefficient and extinction coefficient through multiple photodiodes, and W. Wegrzyriski et al., measured characteristics using densitometers of various wavelengths [8].

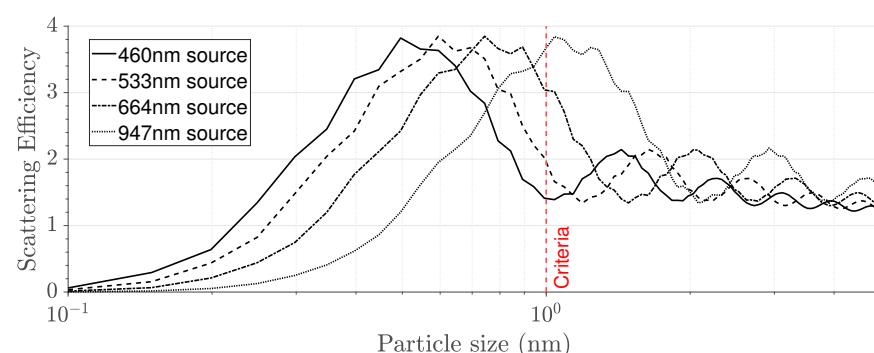
Precise system configurations and substantial data can lead to the creation of sophisticated detectors. However, considering the market for fire detectors, we should be able to successfully distinguish using only the simple scattering technology of LEDs and photodiodes. The limitations of hardware and the need for sensitive fire detection also call for efficient anomaly detection techniques. In this paper, we contribute to creating effective datasets that respond to continuous smoke particle data. Additionally, to the best of our knowledge, there has been no experimental study on the time series classification of scattering data from smoke particles; hence, our results can contribute to this issue.

This paper is composed of five sections. In Section 2, we calculate the scattering efficiency of smoke particles for various wavelengths and discuss techniques related to the multiple wavelengths we focus on. Section 3 proposes methods for efficiently extracting datasets from continuous scattering signals caused by smoke particles. Section 4 introduces the device designed for data collection and experiments, as well as introducing the experimental equipment, defines the learning models, and discusses the compared performance results. Section 5 presents the results of the performance comparisons and analyses for various hyperparameters. The entire paper concludes in Section 6.

## 2. Backgrounds

In an aerosol environment, optical absorption and scattering by particles can be analyzed based on various parameters. Specifically, according to Mie scattering theory, scattering efficiency is primarily influenced by particle size, light source intensity, light source wavelength, refractive index, and scattering angle [9]. To investigate the correlation between particle size and optical scattering intensity at different wavelengths, we performed MATLAB simulations for Mie scattering and absorption as outlined in [10]. In this paper, we focused on four specific wavelengths of interest (460 nm, 533 nm, 664 nm, and 947 nm), and scattering efficiency was computed based on a Lognormal distribution for particle size, following the research of [5]. Note that we specified a refractive index of  $n = 1.55 + 0.02i$ , and the particle size range was defined as  $0.1 \mu\text{m}$  to  $5 \mu\text{m}$ .

Figure 1 shows the results of the total scattering efficiency as a function of particle size for each wavelength. The X-axis represents the particle size, while the Y-axis represents the scattering efficiency  $Q_{sca}$ . Non-fire sources (such as water vapor and dust) typically contain particles larger than  $1 \mu\text{m}$ , whereas smoke particles generated from actual fires are known to be smaller, typically below  $1 \mu\text{m}$  [11]. According to the simulations, ideally, smoke particles generated from actual fires exhibit higher total scattering efficiency at shorter wavelength ranges, while those from non-fire sources exhibit irregular scattering efficiency across all wavelength ranges.

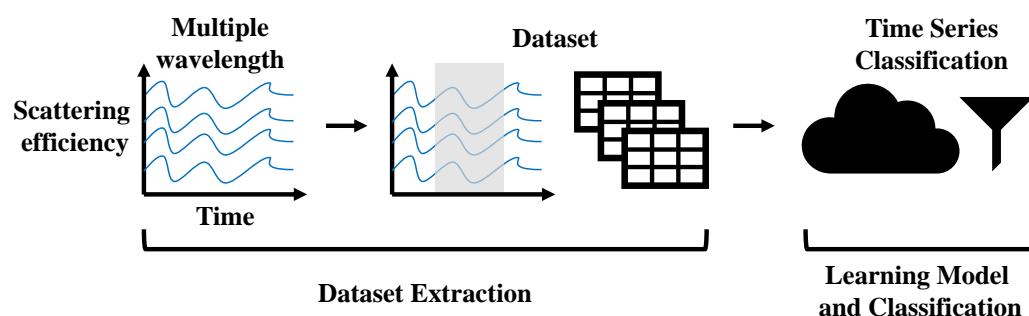


**Figure 1.** Total scattering efficiency based on particle size.

Focusing on the correlation between different wavelengths, it is possible to develop technologies for detecting non-fire sources or distinguishing smoke particles using light sources at various wavelengths. In [12], research focused on extracting characteristic values through the statistical analysis of correlations among multiple wavelengths, while [13] explored the discrimination of smoke particles using signals from multiple wavelengths. In the realm of products, both TI's TPS8802 [14] and Analog Devices' ADPD188BI [15] introduce modules that are based on the scattering efficiency between dual wavelengths, aiming not only for fire detection but also for distinguishing non-fire sources. However, unlike the ideal simulations that satisfy the scattering condition above, in the real environment, due to various variables, smoke particles do not guarantee consistent scattering. The distribution of smoke particles is strongly related to time, influenced by the diffusion speed, which varies due to spatial characteristics and fluid dynamics. Depending on the smoke material defined in the UL268 document [2], each has its own diffusion speed and concentration as detected by optical detectors. Additionally, in [12], the research aimed to extract an effective range due to the instability of smoke particles in time. Furthermore, as presented in Analog Devices' research [16], it was demonstrated that classification is possible based on the difference in diffusion speeds between fire smoke and cooking smoke. Intuitively, to address nuisance alarms and further classify smoke particles, the simultaneous analysis of wavelength-related features and time series data features is necessary.

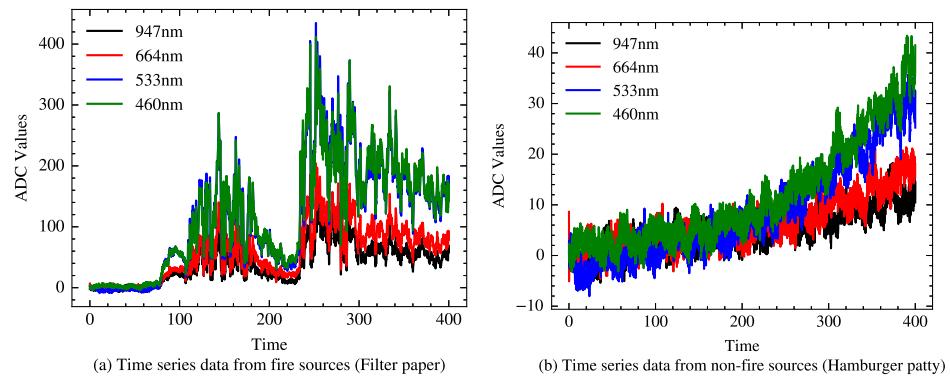
### 3. Methodology

In this paper, we focus on the analysis of scattering signals generated by multi-wavelength light sources to classify between fire smoke and non-fire sources. Figure 2 illustrates a simple overview of our pipeline. The pipeline consists primarily of two key components: the 'data extraction phase', which involves the detection of anomalous segments and the extraction of indices for inputs for the classification model, and the 'learning model and classification phase', used for generating the classification model and applying it to classify different types of smoke particles. Between each of these phases, there are additional steps involving signal preprocessing and tuning the hyperparameters for the training model, and further details will be addressed in the following sections.



**Figure 2.** A simple overview of our pipeline.

Features of time series data for smoke particles generated by fires or non-fire sources are illustrated in Figure 3. Note that Figure 3 is an example of the smoke particle signals collected through the device and experimental environment discussed in Section 4. These signals demonstrate a roughly linear progression over time, displaying either an increase or a decrease from their initial states. Subsequently, it can be observed that they reach a saturation point at a constant level and exhibit fluctuations. Note that this pattern is not limited to the experiments conducted in this paper but is observed in studies involving scattering signal collection through optical detectors, including similar fire tests [12,17]. Based on these shapes of signals, we extract data using three methods. Each method is applicable to every signal and allows for the extraction of indices corresponding to valid segments from the signal as output.



**Figure 3.** The time series data for smoke particles generated by (a) fires (filter paper) and (b) non-fire sources (cooking hamburger patty).

### 3.1. Dataset Extraction

#### 3.1.1. Threshold Detection

Generally, smoke detectors have been studied by calculating smoke concentrations within the detectors and setting thresholds to trigger alarms [18]. Intuitively, when a fire occurs, scattering caused by smoke particles entering the detector results in signal variations according to the internal design. As shown in Figure 3, the signal can increase or decrease depending on the reflection angle and chamber design, allowing for the empirical definition of upper and lower threshold values. In this paper, upper and lower threshold values equivalent to three times the standard deviation of signals in the idle state were set, and indices beyond this range were extracted. Note that these values were empirically determined to detect changes in multivariate data.

#### 3.1.2. Two-Sided Cumulative Sum

The CUSUM (cumulative sum) algorithm is a statistical technique designed to monitor sequential data and detect significant shifts or changes in those data over time. Its primary purpose is to identify deviations from a predefined baseline or expected behavior. This method is valuable in various applications, including quality control, process monitoring, and anomaly detection. It can be used for data extraction when the goal is to pinpoint the exact time or location of a change in the data. When this cumulative sum exceeds a user-defined upper or lower threshold, it detects the existence of a change point in the time series. In this paper, we introduce a two-sided CUSUM algorithm designed for the signal patterns illustrated in Figure 3. Furthermore, we extend this algorithm to handle multivariate data with multiple wavelengths.

Initially, we define two cumulative sums, denoted as  $S_k^i[n]$  and  $S_k^d[n]$ , along with two corresponding decision functions  $G_k^i[n]$  and  $G_k^d[n]$  for both the upper and lower bounds, respectively. Note that  $k$  represents one of the multiple wavelengths  $K$ . We can express the instantaneous log-likelihood ratios  $s_k^i[n]$  and  $s_k^d[n]$  as follows:

$$s_k^i[n] = \frac{|\delta_k|}{\hat{\sigma}_k^2} \left( x_k[n] - \hat{\mu}_{0,k} - \frac{|\delta_k|}{2} \right), \quad (1)$$

$$s_k^d[n] = \frac{|\delta_k|}{\hat{\sigma}_k^2} \left( x_k[n] - \hat{\mu}_{0,k} + \frac{|\delta_k|}{2} \right), \quad (2)$$

where  $|\delta|$  denotes the absolute value of the change,  $x[n]$  denotes the value of  $N$  wavelengths at the current state  $n$ , and  $\mu$  and  $\sigma$  denote the initial mean and variance values for each wavelength [19]. Subsequently, for each sample, we iteratively compute  $S_k^i[n]$ ,  $S_k^d[n]$ ,  $G_k^i[n]$ , and  $G_k^d[n]$  as follows:

$$S_k^i[n] = S_k^i[n-1] + s_k^i[n], \quad S_k^d[n] = S_k^d[n-1] + s_k^d[n], \quad (3)$$

$$G_k^i[n] = \max\{G_k^i[n-1] + s_k^i[n], 0\}, \quad G_k^d[n] = \max\{G_k^d[n-1] + s_k^d[n], 0\}. \quad (4)$$

If  $G_k^i[n] > h$  is satisfied for the predefined threshold  $h$ , we set the change point  $n_c$  within the sample range of  $2t$  as follows:

$$n_c = \min \left\{ \forall k \in K : n-t \leq n < n+t \quad S_k^i[n] \right\}. \quad (5)$$

Similarly, if  $G_k^d[n] > h$  is satisfied, we set the change point  $n_c$  within the sample range of  $2t$  as follows:

$$n_c = \min \left\{ \forall k \in K : n-t \leq n < n+t \quad S_k^d[n] \right\}. \quad (6)$$

If neither of the above two conditions is satisfied, the sample  $n$  is incremented to  $n+1$ , and the process is repeated until the first instance of  $n_c$  detection, at which point it is extracted as an index.

### 3.1.3. Dynamic Time Warping with Reference Sequences

The DTW (Dynamic Time Warping) algorithm has been studied as a method for measuring the similarity between two time series data sequences. One of the notable strengths of the DTW algorithm is its capability to analyze aligned signals and measure time distortions between two time series data sequences, even when their overall lengths and magnitudes vary, achieved by locally stretching or compressing them. Considering that we can roughly estimate the shape of the signals as shown in Figure 3, in this paper, we define the expected reference signal as a state space model and adjust DTW to determine the dataset to be extracted.

Initially, as the state space model, it is defined as

$$\tilde{s}_k = cA^kx_0, \quad (7)$$

where  $A$  is the state transition matrix,  $c$  is the output vector,  $x_0$  is the state vector [20] and the  $k$  represents the sample index. We can split this model into three phases, and it also be expressed as:

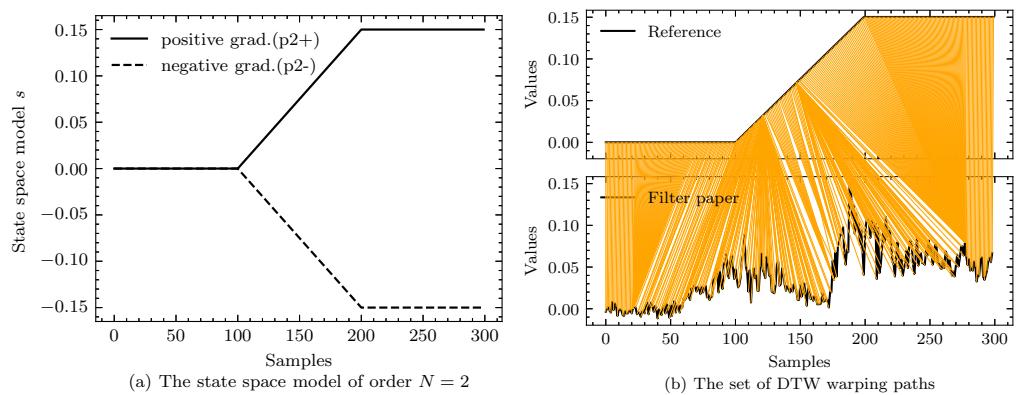
$$\tilde{s}_k = \begin{cases} cA_{p1}^{t_1-k}x_{p1}, & \text{for } k \leq t_1 \\ cA_{p2}^{k-t_1}x_{p2}, & \text{for } t_1 < k \leq t_2 \\ cA_{p2}^{t_2-t_1}x_{p2}, & \text{for } t_2 < k \leq t_3. \end{cases} \quad (8)$$

Assume that there is a one channel of scattering signals, vector  $c$  is a  $1 \times N$  matrix, vector  $A$  is an  $N \times N$  matrix, and vector  $x$  is an  $N \times 1$  matrix. Also, the parameters for the state space model can be set as:

$$c = (1 \quad 0), \quad A_{p1}^{-1} = A_{p2} = \begin{pmatrix} 1 & g \\ 0 & 1 \end{pmatrix}, \quad (9)$$

where  $g$  denotes the gradient for the state transition matrix. In the idle state, we set  $x_{p1} = [0, 0]^T$  except for the small change, but in the upper and lower states we set  $x_{p2}$  as  $x_{p2+} = [0, 1]^T$  and  $x_{p2-} = [0, -1]^T$ , respectively.

Figure 4 illustrates the state space model for  $N = 2$  and depicts the transitions for Phase 1, Phase 2, and Phase 3 with  $x_{p2+}$  and  $x_{p2-}$ . Note that the gradient  $g$  can be defined as the ratio of the maximum value of the signal within the range up to  $k$  to the difference  $(t_2 - t_1)$ , which represents the interval of Phase 2.



**Figure 4.** (a) The state space model of order  $N = 2$  with three phases and (b) The set of DTW warping paths.

Subsequently, the DTW algorithm allows us to calculate the warping path between a reference sequence  $Q = (q_1, q_2, \dots, q_{t_3})$  and a scattered signal  $S = (s_1, s_2, \dots, s_N)$ . These two discrete time series data are aligned to form a  $t_3 \times N$  matrix. The  $(i, j)$  element of this matrix represents the Euclidean distance  $d(q_i, s_j) = (q_i - s_j)^2$  between two points  $q_i$  and  $s_j$ , and we can search for the set  $W$  that satisfies three conditions, the boundary condition, continuity, and monotonicity [21]:

$$W = w_1, w_2, \dots, w_K, \max(t_3, N) \leq K < t_3 + N - 1. \quad (10)$$

Our goal is to find the optimal path that minimizes the warping cost, and we define the cumulative distance up to the current cell  $\gamma$  as follows [22]:

$$\gamma(i, j) = d(q_i, s_j) + \min\{\gamma(i - 1, j - 1), \gamma(i - 1, j), \gamma(i, j - 1)\}. \quad (11)$$

Finally, we can compute the cumulative distances, denoted as  $\gamma^+$ , between the state space models  $\tilde{s}$  with  $x_{p2+}$  and the scattered signal  $S$ . Similarly, we can calculate  $\gamma^-$  for the cumulative distances between  $\tilde{s}$  with  $x_{p2-}$  and  $S$ . By selecting the reference sequence with the smaller warping distance, we can extract the set of optimal warping path elements. In this paper, we describe the phases for the dataset as Phase 2, which features periods of increasing or decreasing signal magnitude, and Phase 3, where the magnitude is maintained at a constant level. Due to the varying waveform of the signal depending on each source, we crop and extract the data to the length of the dataset based on the transition point between Phase 2 and Phase 3.

### 3.2. Learning Model and Classification

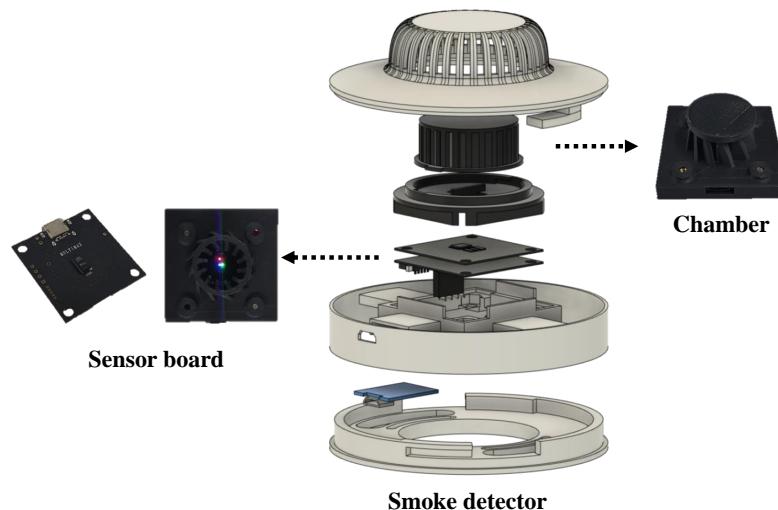
Considerable research has been conducted on learning models for the classification of time series data. The time series data in this paper are characterized by patterns that increase or decrease due to the smoke particles. To compare the performance of learning models for this type of data, three representative architecture categories were selected. Recurrent neural networks, typically used for voice and sentence processing, have already shown effective results for time series data. To compare the number of parameters and computational efficiency, we selected the LSTM (Long Short-Term Memory) [23] and GRU (Gated Recurrent Unit) [24], which can be trained using time series data as input. Among the methods for extracting features from time series data, the mWDN (multilevel Wavelet Decomposition Network), which employs a structure utilizing multi-stage discrete wavelet decomposition and LSTM [25], can also be trained using time series data as input. Secondly, we included the Transformer architecture, which has shown significant performance in various fields [26]. Through the Transformer, the influence of inductive bias can be observed, and it uses tokens of time series data applied with multi-head attention as input. Lastly, the Rocket-based [27] MiniRocket [28], which has achieved state-of-the-art

performance on the UCR univariate time series classification datasets, was also included in the performance comparison. Minirocket generates a wide range of features for time series data as input, allowing for arbitrary lengths, expansion, padding, weights, and biases.

For the selected architectures, we can train the models using the data collected in this paper and choose the model that achieves the highest classification accuracy for the pipeline. For all models, we use the CrossEntropyLossFlat as the loss function, which applies a flattened loss function to the inputs and targets. The selected model, which includes the architecture and parameters that satisfy the optimal performance, can be extracted for extension into a real-time classifier through the same pipeline.

#### 4. Results

From fire and non-fire sources, scattering signals were collected using the following device in Figure 5. We developed a device to enable data acquisition for scattering at approximately 460 nm, 533 nm, 664 nm, and 947 nm wavelengths by selecting the analog front end of the MAXIM Integrated Optical Sensor Module (San Jose, CA, USA) [29]. The device was designed with a chamber to allow the flow of smoke as shown in Figure 5. Note that the chamber in this paper is not designed solely to observe scattering as in [15], but rather it is a simple design intended to centralize the flow of smoke. The intensity of the LEDs was set based on the efficiency of the photodiodes in ambient conditions at the respective wavelengths to ensure the reproducibility of every device. Ultimately, this structured device was developed to operate at a 50 Hz sampling rate for data across the four wavelengths.

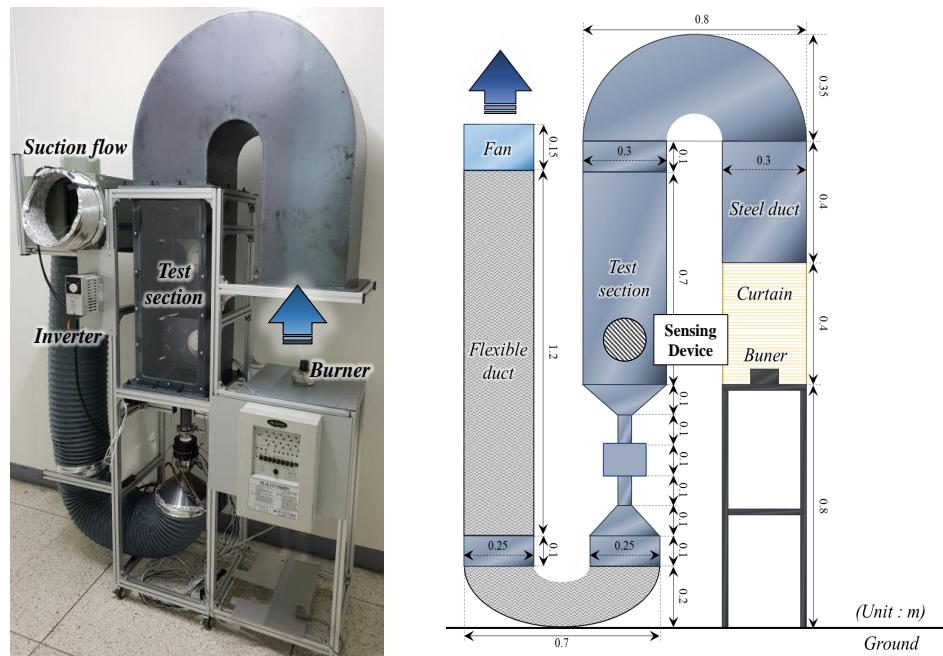


**Figure 5.** Design of a smoke detector for measuring scattering signals at multiple wavelengths.

##### 4.1. Design of the Sensing Device and the Experimental Equipment

Figure 6 illustrates the structure and photographs of the experimental setup used for generating smoke particles. The experimental equipment includes a square duct with a cross-sectional area of  $0.09 \text{ m}^2$  ( $0.3 \text{ m} \times 0.3 \text{ m}$ ) made of carbon steel. Materials for fire experiments were placed in a burner for combustion, and a fan was installed at the end of the smoke path to create a uniform flow. The generated particles first entered the duct and then the test section. In the test section, a total of 10 sensor devices were installed perpendicular to the direction of particle flow, according to the design mentioned above. A total of six types of fire sources and non-fire sources were selected for the smoke particle generation experiment. As shown in Table 1, the fire sources included paper, kerosene, and polyethylene, while the non-fire sources were dust, vapor, and hamburger patties. The experiments were conducted 20 times for each fire and non-fire source, and identical samples of the same quantity were used for each smoke particle generation experiment as much as possible. Tests #01 to #20 used filter paper from ADVANTEC (Tokyo, Japan).

Tests #21 to #40 generated smoke using a wick with the liquid fuel kerosene. Tests #41 to #60 utilized polymer pellets of polyethylene. For Tests #61 to #80, dust of  $\mu\text{m}$  size was dispersed through a blower. Tests #81 to #100 generated smoke particles by cooking hamburger patties on a heated pan, and Tests #101 to #120 produced vapor using an ultrasonic humidifier. For a more detailed description of the experimental environment, refer to [30].



**Figure 6.** The structure and photographs of experimental equipment for generating smoke particles.

**Table 1.** Smoke particle generation experiments: fire and non-fire sources.

Label	Source
Test #01~Test #20	Filter paper
Test #21~Test #40	Kerosene
Test #41~Test #60	Polyethylene
Test #61~Test #80	Dust
Test #81~Test #100	Hamburger patty
Test #101~Test #120	Vapor

#### 4.2. Dataset and Learning Models

In Section 3, we addressed methods for extracting datasets from time series data. Specifically, we identified three methods that enable us to extract indices representing the starting point of a dataset. The first method, based on a threshold, sets the threshold at  $\pm \text{std} * 3$  of the idle state for each wavelength and selects the smallest index that exceeds this threshold. The second method, based on CUSUM, selects the smallest index detected by the algorithm for each wavelength. Lastly, the DTW-based method chooses indices corresponding to Phases 2 and 3 as defined in the reference sequence. From these indices, we use continuous data of various lengths following each index as our dataset. For performance comparison according to the length of the time series dataset in this paper, we utilize datasets of 500 samples, 1500 samples, and 3000 samples in length. Recall that, with a 50 Hz sampling rate, these correspond to periods of 10 s, 30 s, and 60 s, respectively.

The dataset is divided into training, testing, and validation sets as follows. According to Table 1, a total of 1200 sequence data are generated from 10 sensors, with 200 sequences per source. While it is possible to separate these into training and testing sets using a random train-test split function, the sensor data from a single test are not entirely

independent, as they coexist within the same chamber. To reduce the issue of leakage appearing due to non-independence when randomly splitting the train and test sets, we allocated the datasets as follows: 17 out of 20 trials (1020 sets) were used for the training set, 1 trial (60 sets) for the testing set, and 2 trials (120 sets) for the validation set.

The dataset includes eight multivariate features, each of which represents a different aspect of the scattering signal analysis. These features are derived from the scattering signals obtained at various wavelengths and are processed to capture distinct characteristics related to smoke particle detection. Features of the dataset can be expressed as given in Table 2.

**Table 2.** Features of the dataset.

Feature 1	The difference between the ADC values and the mean of the initial idle ADC values at 947 nm wavelength
Feature 2	The difference between the ADC values and the mean of the initial idle ADC values at 664 nm wavelength
Feature 3	The difference between the ADC values and the mean of the initial idle ADC values at 533 nm wavelength
Feature 4	The difference between the ADC values and the mean of the initial idle ADC values at 460 nm wavelength
Feature 5	The ratio of the ADC values to the mean of the initial idle ADC values at 947 nm wavelength
Feature 6	The ratio of the ADC values to the mean of the initial idle ADC values at 664 nm wavelength
Feature 7	The ratio of the ADC values to the mean of the initial idle ADC values at 533 nm wavelength
Feature 8	The ratio of the ADC values to the mean of the initial idle ADC values at 460 nm wavelength

Each of these features is obtained by processing the scattering signals, subtracting the idle state signal, and dividing by the mean of the idle state signal. This transformation allows us to capture both the absolute and relative differences in magnitude between the sensor's idle state and the scattered signal from smoke particles. The resulting dataset is organized into two-dimensional tensors with a shape of  $(8, length)$ , where  $length$  represents the number of data points in the dataset.

For training with varying parameters, the LSTM and GRU models were expanded by adjusting the hyperparameters, such as the number of layers, hidden size, and bidirectional settings. mWDN utilized the results of a 4-level decomposition. Training and evaluation were carried out using the open-source deep learning package “tsai” with modifications, and default values were used for unspecified specific parameters [31]. Table 3 represents the architecture combinations compared in this paper. Training was conducted in an NVIDIA A6000 environment with a batch size of 16, and early stopping was applied with a maximum of 200 epochs.

**Table 3.** The combinations of model architectures.

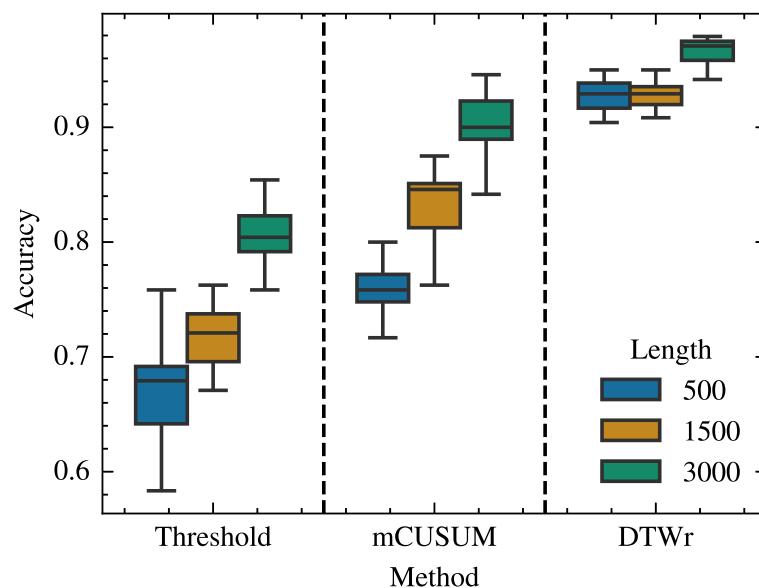
Methodology	Threshold-based method (Threshold) Multivariate CUSUM-based method (mCUSUM) DTW-based method with reference sequences (DTWr)
Length	500 / 1500 / 3000
Models	TransformerModel
	mWDN
	MiniRocket
	GRU
	LSTM

#### 4.3. Performance Comparison

Figure 7 is a boxplot of the entire accuracy results, plotted according to the dataset generation method and length. Table 4 presents the top five models ranked by the highest accuracy and F1 score to compare classification performance according to the dataset extraction methods and length. It also includes additional details regarding performance differences due to hyperparameters or the number of parameters, as well as additional ablation study results.

First, by considering the dataset extraction methods in Figure 7, along the rows of the table, we can compare the overall performance. In almost all length cases, DTWr achieved superior accuracy and F1 scores. Even when comparing the top five models, the overall performance improvement was evident in the order of DTWr, mCUSUM, and Threshold. This addressed that DTWr efficiently selected features for time series data. Unlike the other two methods that select the starting point of signal changes, DTWr implied effective dataset extraction through phase division using the reference sequence.

Next, we compared the overall performance of the dataset based on its length in Figure 7, along the columns of the table. We found that as the dataset length increases, which is equivalent to taking a longer time, the classification performance improves. Particularly noteworthy is the significant improvement in the classification performance as the dataset length increases for Threshold and mCUSUM. While the performance improvement for DTWr may not be as pronounced, an overall improvement can be observed in the top five performance results. Intuitively, this indicates that as time progresses in the case of scattering signals from smoke particles, the concentration of smoke particles increases, leading to improved classification performance when using longer time data as the dataset. This is also related to trade-offs in real-world applications, considering a sampling rate of 50 Hz, where obtaining classification results requires 10 s, 30 s, and 60 s, respectively, for each length. (Note that the classification execution time for each model is not considered.) Depending on the domain, collecting data over a long time can become a bottleneck for applications.



**Figure 7.** Classification performance comparison by dataset generation method and length.

Lastly, we compared the overall performance based on the learning models. Table 4 shows that the top-performing model is MiniRocket, achieving a very high classification accuracy of 0.9875 at a length of 1500. MiniRocket generates a wide range of features through arbitrary length, expansion, padding, weights, and biases, and captures patterns at various frequencies and scales through kernel expansion for classification [28]. We found that in datasets of shorter length, higher performance is achieved through kernel expansion, and as the length of the dataset increases, there is an improvement in the performance of the wavelet-based mWDN, which utilizes both time- and frequency-domain information, as well as the RNN-based models that leverage long sequences. However, considering the overall performance results, it appears that MiniRocket is the most suitable for the characteristics of the scattering signals.

**Table 4.** Top five classification performance comparison by dataset generation method and length.

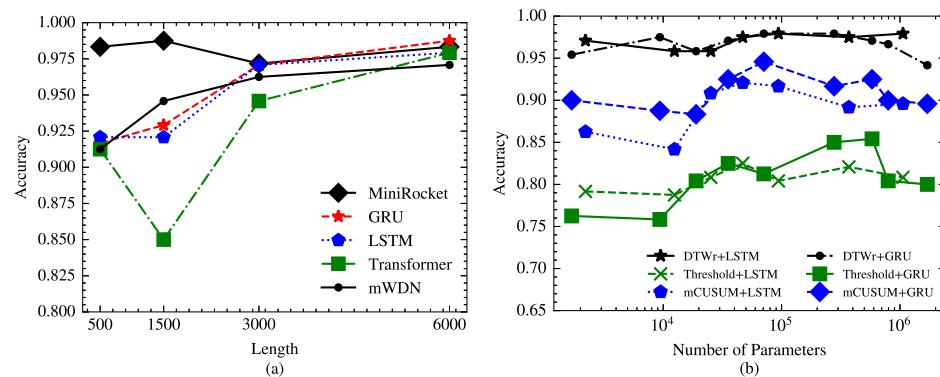
Length	No.	Method									
		Threshold			mCUSUM			DTWr			
		Arch. Hyperparams.	Accuracy	F1 Score	Arch. Hyperparams.	Accuracy	F1 Score	Arch. Hyperparams.	Accuracy	F1 Score	
500	1	MiniRocket {}	0.7583	0.7576	MiniRocket {}	0.9167	0.9164	MiniRocket {}	0.9833	0.9833	GRU
	2	GRU {'n_layers': 5, 'bidirectional': True, 'hidden_size': 100}}	0.7125	0.7120	TransformerModel {}	0.8000	0.8014	GRU {'n_layers': 5, 'bidirectional': True, 'hidden_size': 100}}	0.9450	0.9495	LSTM
	3	GRU {'n_layers': 10, 'bidirectional': True, 'hidden_size': 100}}	0.7042	0.7043	GRU {'n_layers': 1, 'bidirectional': True, 'hidden_size': 10}}	0.7958	0.7910	GRU {'n_layers': 5, 'bidirectional': False, 'hidden_size': 100}}	0.9458	0.9451	LSTM
	4	GRU {'n_layers': 10, 'bidirectional': False, 'hidden_size': 100}}	0.7041	0.7042	GRU {'n_layers': 5, 'bidirectional': True, 'hidden_size': 10}}	0.7792	0.7778	GRU {'n_layers': 5, 'bidirectional': True, 'hidden_size': 100}}	0.9417	0.9410	LSTM
	5	GRU {'n_layers': 5, 'bidirectional': False, 'hidden_size': 100}}	0.6958	0.6965	GRU {'n_layers': 10, 'bidirectional': True, 'hidden_size': 100}}	0.7750	0.7703	GRU {'n_layers': 1, 'bidirectional': True, 'hidden_size': 100}}	0.9417	0.9409	LSTM
1500	1	MiniRocket {}	0.8458	0.8455	MiniRocket {}	0.9208	0.9201	MiniRocket {}	0.9875	0.9875	GRU
	2	GRU {'n_layers': 1, 'bidirectional': False, 'hidden_size': 100}}	0.7625	0.7629	GRU {'n_layers': 5, 'bidirectional': True, 'hidden_size': 10}}	0.8750	0.8715	GRU {'n_layers': 1, 'bidirectional': False, 'hidden_size': 100}}	0.9500	0.9492	LSTM
	3	TransformerModel {}	0.7500	0.7464	GRU {'n_layers': 1, 'bidirectional': True, 'hidden_size': 100}}	0.8708	0.8656	mWDN {'levels': 4}	0.9458	0.9450	LSTM
	4	GRU {'n_layers': 1, 'bidirectional': True, 'hidden_size': 100}}	0.7458	0.7461	GRU {'n_layers': 1, 'bidirectional': True, 'hidden_size': 100}}	0.8625	0.8582	GRU {'n_layers': 1, 'bidirectional': True, 'hidden_size': 100}}	0.9417	0.9410	LSTM
	5	GRU {'n_layers': 10, 'bidirectional': True, 'hidden_size': 100}}	0.7417	0.7427	TransformerModel {}	0.8542	0.8532	GRU {'n_layers': 1, 'bidirectional': True, 'hidden_size': 100}}	0.9417	0.9410	LSTM
3000	1	MiniRocket {}	0.9125	0.9090	GRU {'n_layers': 1, 'bidirectional': True, 'hidden_size': 100}}	0.9458	0.9455	GRU {'n_layers': 5, 'bidirectional': False, 'hidden_size': 100}}	0.9792	0.9791	LSTM
	2	GRU {'n_layers': 10, 'bidirectional': False, 'hidden_size': 100}}	0.8542	0.8531	MiniRocket {}	0.9375	0.9334	GRU {'n_layers': 1, 'bidirectional': True, 'hidden_size': 100}}	0.9717	0.9791	LSTM
	3	GRU {'n_layers': 5, 'bidirectional': False, 'hidden_size': 100}}	0.9500	0.8489	mWDN {'levels': 4}	0.9333	0.9330	GRU {'n_layers': 5, 'bidirectional': True, 'hidden_size': 100}}	0.9792	0.9790	LSTM
	4	GRU {'n_layers': 1, 'bidirectional': False, 'hidden_size': 100}}	0.8250	0.8229	GRU {'n_layers': 10, 'bidirectional': False, 'hidden_size': 100}}	0.9250	0.9239	GRU {'n_layers': 1, 'bidirectional': True, 'hidden_size': 100}}	0.9792	0.9790	LSTM
	5	GRU {'n_layers': 1, 'bidirectional': False, 'hidden_size': 100}}	0.8250	0.8225	GRU {'n_layers': 1, 'bidirectional': False, 'hidden_size': 100}}	0.9250	0.9239	GRU {'n_layers': 5, 'bidirectional': False, 'hidden_size': 100}}	0.9750	0.9750	LSTM

All results can be stored with the model that satisfies the optimal performance, and through the connection with sensors, outcomes can be derived using this model. Returning to the pipeline in Figure 2, we can either develop a standalone device with GPU acceleration capable of running the classifier using the generated trained model, or create a real-time classifier by establishing a communication interface with a server that performs classification. Particularly, in the case of MiniRocket, which is deemed most suitable, it

is available for operation on edge neural network accelerators, such as Edge TPU and NVIDIA Jetson series, provided via tsai and implemented in PyTorch v0.3.8 [31].

## 5. Ablation Study

In Figure 8, we present the results of our ablation study. First, Figure 8a illustrates the performance comparison according to the length of the dataset for each model, with the dataset extraction method fixed to the DTWr proposed in this paper. For MiniRocket, there is no significant change in accuracy with respect to the dataset length, whereas other models achieved higher accuracy as the data length increased. Additionally, we found that the Transformer performed lower than RNN-based models such as LSTM and GRU. RNNs are designed to process sequential information, and the dataset generated in this paper is strongly related to time series data. This indicates that RNN-based models are more fitted to scattered data than Transformers, which have a relatively weak inductive bias [32]. Next, Figure 8b fixes the data length at 1500 and shows the performance comparison according to the number of parameters of the training model, categorized by the dataset extraction method. The order of the performance was addressed as DTWr, mCUSUM, and Threshold, discussed in Section 4. Both LSTM and GRU have an increasing number of parameters with settings related to the number of recurrent layers, the dimensions of the hidden state, and the use of bidirectional hyperparameters. As shown in Figure 8b, there is a slight tendency for performance to increase with the number of parameters, but this is not distinctly evident for all three methods. Note that the choice of bi-LSTM also did not result in significant performance differences. This suggests that a higher-complexity model does not guarantee performance improvement and that it can vary depending on the model design.



**Figure 8.** Performance comparison results for ablation study: (a) accuracy with fixed method and varying length, and (b) accuracy with fixed length and varying the number of parameters.

## 6. Conclusions

We present a classification method for time series scattered signals of smoke particles aimed at mitigating nuisance or false alarms. The optical scattering characteristics of aerosol smoke particles are analyzed, and a methodology for extracting datasets from time series data is proposed. Various learning models for time series data are selected, and their performance is compared based on a variety of hyperparameters, dataset lengths, and dataset generation methods. We find that the DTW method, used in conjunction with reference signals defined through the state space model, generally shows higher performance. An ablation study is conducted to compare the performance results while varying different variables.

From a practical perspective, there are still limitations. In certain scenarios, such as rapidly spreading fires, there may not be enough time to collect data of the length discussed in this paper. Performance-wise, the DTW-based method, although effective in data extraction, necessitates analyzing either the entire dataset or a sufficiently large subset, which can be a drawback. In the case of CUSUM, while early detection is feasible, setting

the initial parameters can be challenging, varying with environmental conditions. This highlights the need for technology adept at detecting variations specific to different stages of fire combustion. Following this, we hope that our approach will lead to the application of time series classification in the stages of fire combustion.

**Author Contributions:** Conceptualization, K.H.; methodology, K.H.; software, K.H. and K.C.; hardware, S.K. and H.Y.; validation, K.H., S.K., H.Y. and K.L.; formal analysis, K.H.; investigation, S.K. and H.Y.; resources, H.Y. and K.C.; data curation, K.H.; writing—original draft preparation, K.H.; writing—review and editing, K.H., H.Y. and K.L.; visualization, K.H.; supervision, K.L.; project administration, H.Y.; funding acquisition, K.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was supported by the Institute for Information & Communications Technology Promotion (IITP) grant funded by the Korean government (MSIT) (no. 2020-0-00012, Development of Intelligent Fire Detection Equipment Based on Smoke Particle Spectrum Analysis).

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

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

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

## Abbreviations

The following abbreviations are used in this manuscript:

UL	Underwriters Laboratories
MATLAB	Matrix Laboratory
CUSUM	Cumulative Sum
DTW	Dynamic Time Warping
LSTM	Long Short-Term Memory
GRU	Gated Recurrent Unit

## References

1. Lee, A.; Pineda, D. *Smoke Alarms—Pilot Study of Nuisance Alarms Associated with Cooking*; U.S. Consumer Product Safety Commission: Bethesda, MD, USA, 2010; pp. 1–3.
2. UL 268-2016; UL268: UL Standard for Safety Smoke Detectors for Fire Alarm Systems, 7th ed.; UL: Northbrook, IL, USA, 2023; pp. 192–208.
3. Yan, S.; Deng, T.; Xu, W.; Wang, S. Selecting an Optimal Set of Scattering Angles and Wavelengths for Practical Photoelectric Smoke Detector. In Proceedings of the 2017 Suppression, Detection, and Signaling Research and Applications Conference, College Park, MD, USA, 12–14 September 2017.
4. Zhang, Q.; Liu, Z.; Luo, J.; Wang, F.; Wang, J.; Zhang, Y. Characterization of Typical Fire and Non-fire Aerosols by Polarized Light Scattering for Reliable Optical Smoke Detection. In Proceedings of the 11th Asia-Oceania Symposium on Fire Science and Technology, Taipei, Taiwan, 22–24 October 2018 ; pp. 791–801.
5. Deng, T.; Wang, S.; Zhu, M. Dual-wavelength optical sensor for measuring the surface area concentration and the volume concentration of aerosols. *Sens. Actuators B Chem.* **2016**, *236*, 334–342. [[CrossRef](#)]
6. Krüll, W.; Schultze, T.; Tobera, R.; Willms, I. Characterization of Dust and Water Steam Aerosols in False Alarm Scenarios—Design of a Test Method for Fire Detectors in Dusty and in Highly Foggy Environments. In Proceedings of the 2015 Suppression, Detection, and Signaling Research and Applications Conference, Orlando, FL, USA, 3 March 2015.
7. Li, K.; Liu, G.; Yuan, H.; Chen, Y.; Dai, Y.; Meng, X.; Kang, Y.; Huang, L. Dual-Wavelength Smoke Detector Measuring Both Light Scattering and Extinction to Reduce False Alarms. *Fire* **2023**, *6*, 140. [[CrossRef](#)]
8. Wegrzyriski, W.; Antosiewicz, P.; Fangrat, J. Multi-Wavelength Densitometer for Experimental Research on the Optical Characteristics of Smoke Layers. *Fire Technol.* **2021**, *57*, 2683–2706. [[CrossRef](#)]
9. van de Hulst, H.C. Light Scattering by Small Particles. *Q. J. R. Meteorol. Soc.* **1957**, *84*, 198–199. [[CrossRef](#)]
10. Christian, M. *MATLAB Functions for Mie Scattering and Absorption*; Institut für Angewandte Physik: Bern, Switzerland, 2002.
11. Cole, M. Aerosol characterisation for reliable ASD operation. In Proceedings of the 14th International Conference on Automatic Fire Detection, Duisburg, Germany, 8–10 September 2009.
12. Han, K.; Kim, S.; Yang, H.; Cho, K.S.; Lee, K.; Han, H. Statistical Characteristics of Scattering Ratio Based on Three Optical Wavelengths for Smoke Particles. *Int. J. Fire Sci. Eng.* **2022**, *36*, 40–49. [[CrossRef](#)]

13. Ahn, Y.; Han, K.; Yang, H.; Kim, S.; Ryu, J.; Lee, K. Smoke Particle-Source Prediction Model Based on Multiple Optical Wavelengths Using Deep Learning. *Int. J. Fire Sci. Eng.* **2023**, *37*, 20–29. [[CrossRef](#)]
14. TPS8802 Smoke Alarm AFE. Available online: <https://www.ti.com/document-viewer/tps8802/datasheet> (accessed on 13 November 2023).
15. Integrated Optical Module for Smoke Detection. Available online: <https://www.analog.com/media/en/technical-documentation/data-sheets/adpd188bi.pdf> (accessed on 13 November 2023).
16. Smoke Testing with the ADPD188BI Optical Smoke and Aerosol Detection Module. Available online: <https://www.analog.com/media/en/technical-documentation/app-notes/an-1567.pdf> (accessed on 13 November 2023).
17. Ren, E.; Zalmai, N.; Loeliger H. Multi-channel Information Processing for Fire Detection. In Proceedings of the 2017 Suppression, Detection, and Signaling Research and Applications Conference, Hyattsville, MD, USA, 12 September 2017.
18. Geiman, J.; Gottuk, D.T. Alarm Thresholds for Smoke Detector Modeling. In Fire Safety Science, Proceedings of the Seventh International Symposium, Worcester, MA, USA, 16–21 June 2002; pp. 197–208.
19. Granjon, P. The CuSum Algorithm—A Small Review. 2013. ID: hal-00914697f. Available online: <https://hal.science/hal-00914697> (accessed on 26 December 2023).
20. Wildhaber, R.A.; Zalmai, N.; Jacomet, M.; Loeliger, H. Windowed State-Space Filters for Signal Detection and Separation. *IEEE Trans. Signal Process.* **2018**, *66*, 3768–3783. [[CrossRef](#)]
21. Keogh, E.; Ratanamahatana, C.A. Exact indexing of dynamic time warping. *Knowl. Inf. Syst.* **2005**, *7*, 357–386. [[CrossRef](#)]
22. Giorgino, T. Computing and Visualizing Dynamic Time Warping Alignments in R: The dtw Package. *J. Stat. Softw.* **2009**, *31*, 1–24. [[CrossRef](#)]
23. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [[CrossRef](#)] [[PubMed](#)]
24. Chung, J.; Gulcehre, C.; Cho, K.; Bengio Y. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. In Proceedings of the NIPS 2014 Deep Learning and Representation Learning Workshop, Montréal, QC, Canada, 8 December 2014.
25. Wang, J.; Wang, Z.; Li, J.; Wu, J. Multilevel Wavelet Decomposition Network for Interpretable Time Series Analysis. In Proceedings of the KDD '18: 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, London, UK, 19 August 2018; pp. 2437–2446.
26. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin I. Transformer: Attention is All You Need. In Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA, 4–9 December 2017.
27. Dempster, A.; Petitjean, F.; Webb, G.I. ROCKET: Exceptionally fast and accurate time series classification using random convolutional kernels. *Data Min. Knowl. Discov.* **2020**, *34*, 1454–1495. [[CrossRef](#)]
28. Dempster, A.; Schmidt, D.F.; Webb, G.I. MiniRocket: A Very Fast (Almost) Deterministic Transform for Time Series Classification. In Proceedings of the KDD '21: 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining, New York, NY, USA, 14 August 2021.
29. Integrated Optical Sensor Module for Mobile Health. Available online: <https://www.analog.com/media/en/technical-documentation/data-sheets/max86916.pdf> (accessed on 13 November 2023).
30. Jang H.; Hwang, C. Preliminary Study for Smoke Color Classification of Combustibles Using the Distribution of Light Scattering by Smoke Particles. *Appl. Sci.* **2023**, *13*, 669. [[CrossRef](#)]
31. tsai—A State-of-the-Art Deep Learning Library for Time Series and Sequential Data. Available online: <https://github.com/timeseriesAI/tsai> (accessed on 13 November 2023).
32. Kharitonov, E.; Chaabouni, R. What they do when in doubt: A study of inductive biases in seq2seq learners. In Proceedings of the ICLR 2021, Vienna, Austria, 4 May 2021.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.