

## Article

# Fire Detection in Urban Areas Using Multimodal Data and Federated Learning

Ashutosh Sharma <sup>1,2</sup>, Rajeev Kumar <sup>3</sup> , Isha Kansal <sup>3</sup>, Renu Popli <sup>3</sup>, Vikas Khullar <sup>3,\*</sup> , Jyoti Verma <sup>4</sup>   
and Sunil Kumar <sup>5</sup> 

- <sup>1</sup> Business School, Henan University of Science and Technology, Luoyang 471300, China; sharmaashutosh1326@gmail.com
- <sup>2</sup> Department of Informatics, School of Computer Science, University of Petroleum and Energy Studies, Dehradun 248007, Uttarakhand, India
- <sup>3</sup> Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura 140401, Punjab, India; rajeev.kumar@chitkara.edu.in (R.K.); singla.isha1@gmail.com (I.K.); renu.popli@chitkara.edu.in (R.P.)
- <sup>4</sup> Department of Computer Science and Engineering, Punjabi University, Patiala 147002, Punjab, India; jyoti.snehiverma@gmail.com
- <sup>5</sup> Department of Computer Science, Graphic Era Hill University, Dehradun 248001, Uttarakhand, India; drskumar.cs@gmail.com
- \* Correspondence: vikas.khullar@gmail.com

**Abstract:** Fire chemical sensing for indoor detection of fire plays an essential role because it can detect chemical volatiles before smoke particles, providing a faster and more reliable method for early fire detection. A thermal imaging camera and seven distinct fire-detecting sensors were used simultaneously to acquire the multimodal fire data that is the subject of this paper. The low-cost sensors typically have lower sensitivity and reliability, making it impossible for them to detect fire at greater distances. To go beyond the limitation of using solely sensors for identifying fire, the multimodal dataset is collected using a thermal camera that can detect temperature changes. The proposed pipeline uses image data from thermal cameras to train convolutional neural networks (CNNs) and their many versions. The training of sensors data (from fire sensors) uses bidirectional long-short memory (BiLSTM-Dense) and dense and long-short memory (LSTM-DenseDenseNet201), and the merging of both datasets demonstrates the performance of multimodal data. Researchers and system developers can use the dataset to create and hone cutting-edge artificial intelligence models and systems. Initial evaluation of the image dataset has shown densenet201 as the best approach with the highest validation parameters (0.99, 0.99, 0.99, and 0.08), i.e., Accuracy, Precision, Recall, and Loss, respectively. However, the sensors dataset has also shown the highest parameters with the BiLSTM-Dense approach (0.95, 0.95, 0.95, 0.14). In a multimodal data approach, image and sensors deployed with a multimodal algorithm (densenet201 for image data and Bi LSTM-Dense for Sensors Data) has shown other parameters (1.0, 1.0, 1.0, 0.06). This work demonstrates that, in comparison to the conventional deep learning approach, the federated learning (FL) approach performs privacy-protected fire leakage classification without significantly sacrificing accuracy and other validation parameters.

**Keywords:** fire detection; multimodal; sensors; thermal camera; image enhancement



**Citation:** Sharma, A.; Kumar, R.; Kansal, I.; Popli, R.; Khullar, V.; Verma, J.; Kumar, S. Fire Detection in Urban Areas Using Multimodal Data and Federated Learning. *Fire* **2024**, *7*, 104. <https://doi.org/10.3390/fire7040104>

Academic Editors: Yongzheng Yao, Jinlong Zhao, Qiang Wang and Ziheng Gao

Received: 25 January 2024  
Revised: 27 February 2024  
Accepted: 28 February 2024  
Published: 22 March 2024



**Copyright:** © 2024 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

Indoor fire identification is of utmost importance as it enables early detection and timely response to potential fire hazards within buildings. Early fire monitoring is essential. In conventional fire detection techniques, sensors are utilized to identify several fire parameters, such as smoke, fire scale, initial flame area, and air temperature. Due to their affordability and convenience, these sensors have been applied extensively. However,

there are a lot of problems and hazards with this conventional detection technique. For instance, it takes time for smoke to rise to the ceiling, which causes the alert to sound later than necessary and delays the fire alarm's notification. Moreover, a lot of sensors are challenging to utilize in open areas and work best in small regions [1]. Machine vision-based fire detection and monitoring systems have improved recently at a reduced cost, with enhanced real-time performance and the capacity for a great degree of detecting precision. By utilizing multimodal sensor data, it becomes possible to detect fires based on visual cues, temperature changes, and the presence of smoke or hazardous fire. Multimodal sensors refer to devices that can capture and collect data from multiple sources, such as visual, thermal, and fire sensors [2].

By combining different types of sensors, the accuracy of fire identification can be significantly improved. This approach allows for a more reliable detection system that can distinguish between false alarms and actual fire incidents, ensuring timely response and minimizing potential damage. Additionally, multimodal sensors enable real-time monitoring and analysis of fire-related data, enhancing overall safety measures in indoor environments such as homes, offices, and public buildings. With the ability to detect not only smoke but also heat and fire emissions, multimodal sensors provide a more holistic view of fire conditions. This enables faster decision-making and evacuation procedures, ultimately saving lives and reducing property loss.

Image-distributed data refers to the collection and distribution of visual information from various sources, such as surveillance cameras, thermal imaging devices, and drones, to identify and analyze fire incidents in indoor environments. These images capture crucial details like the location, intensity, and spread of the fire, providing valuable insights for effective response strategies. However, handling distributed image data poses challenges such as data synchronization, storage capacity, and real-time analysis. Potential solutions include utilizing cloud-based platforms for seamless data integration and implementing advanced image recognition. Federated learning is a decentralized approach to machine learning that allows models to be trained across multiple devices or locations without the need for data to be centralized. Furthermore, as the model is trained locally on each device, only aggregated results are shared with the central server [3]. This not only improves efficiency but also enhances scalability, allowing for the deployment of indoor fire identification models in a wide range of environments without significant infrastructure requirements. In this paper, the data security and privacy concerns in fire detection are successfully addressed by the application of federated learning techniques. The latest iteration of federated learning can efficiently mitigate these challenges.

The conventional approach to training centralized deep learning models necessitates the transfer of substantial volumes of video and image data to the cloud. This not only consumes considerable network bandwidth but also presents challenges in guaranteeing the confidentiality and privacy of image data. Federated learning fulfills the prerequisites for collaborative machine model learning training while ensuring the confidentiality of client data, as it operates on a distributed machine learning framework supported by secure encryption technology. Federated learning ensures user privacy by accomplishing the decentralization of data from the central server to the client. In the context of indoor fire identification, federated learning can be used to train models using image and sensor data collected from various locations, ensuring privacy and data security.

This collaborative approach enables the development of robust and accurate fire detection models that can be deployed across different buildings or environments. Additionally, federated learning enables continuous learning and adaptation to new fire patterns as more devices contribute to the training process, resulting in improved accuracy over time [4,5]. Multimodal sensor data and distributed image data can be integrated with federated learning to enhance the accuracy and efficiency of fire detection systems. By combining data from various sensors such as temperature, smoke, and humidity, along with distributed image data captured by surveillance cameras, a comprehensive understanding of the fire

situation can be achieved. This integration allows for more robust and reliable fire detection models, enabling quicker response times and better mitigation strategies.

The paper's contributions span from traditional sensor-based fire detection to leveraging thermal imaging, exploring multimodal approaches, and implementing privacy-centric federated learning. These contributions collectively aim to advance the field of fire detection in urban areas, offering potential improvements in accuracy, efficiency, and privacy protection. The primary contributions of this paper are as follows:

- **Identification of Fire Using Sensors Dataset with Deep Learning Models:** This contribution involves utilizing a dataset collected from gas-detecting sensors for the identification of fire. The use of fundamental deep learning models indicates that the research employs established and widely used techniques in the field of artificial intelligence. The significance lies in the exploration of how sensor data alone, which are traditionally used for fire detection, can be effectively processed and classified using deep learning models. This could contribute to improving the accuracy and speed of fire detection systems in indoor environments.
- **Identification of Fire Using Thermal Image Dataset with Deep Learning Models:** This contribution focuses on the use of thermal imaging data for fire identification, employing fundamental deep learning models. Thermal imaging can capture temperature changes associated with fires before smoke particles become visible, offering an additional dimension to fire detection. The research highlights the potential of thermal imaging in combination with deep learning models, showcasing how visual information can be a valuable source for fire detection, especially in scenarios where traditional sensors might have limitations.
- **Multimodal Fire Identification Using Both Sensors and Image Datasets:** This contribution represents the integration of data from both sensors and thermal imaging cameras for fire identification. By combining these modalities, the research aims to create a more robust and comprehensive fire detection system. The multimodal approach addresses the limitations of individual data sources, potentially improving the accuracy and reliability of fire detection by considering multiple aspects such as gas presence and temperature changes.
- **Fire Identification Mechanism Based on Federated Learning:** The incorporation of federated learning (FL) in fire identification is a significant contribution. FL allows model training across multiple devices without centralized data, enhancing privacy and security. The emphasis on safeguarding the privacy of consumers' private information is crucial in scenarios like fire detection where sensitive data might be involved. FL provides a solution by training models collaboratively without exposing raw data to a central server. Indoor fire detection is vital and requires new techniques to surpass conventional detection methods. This research uses federated learning to integrate multimodal sensor data with distributed information to improve urban fire detection system accuracy, efficiency, and privacy. Conventional sensor datasets, thermal imagery, and a privacy-centric federated method of learning are used to progress the field and meet the growing need for quicker and more precise interior fire detection.

## 2. Related Work

The integration of multimodal sensor data and image analysis using federated learning for fire detection in urban areas represents a cutting-edge approach with significant potential for enhancing the efficiency and accuracy of fire detection systems. This approach addresses the challenges posed by the complex and dynamic nature of urban environments, where traditional fire detection methods may fall short. The integration of multimodal sensor data and image analysis through federated learning holds great promise for advancing fire detection capabilities in urban areas. As technology continues to evolve, addressing the challenges and refining the implementation of this approach will be crucial for its successful deployment in real-world scenarios. The literature review evaluates

research in two dimensions: sensor-based fire detection systems and image-based fire detection systems.

It is essential to broaden the literature evaluation of the paper to include a wider range of fire detection systems. This will help place the present study within the larger framework of research in this field. Although the study utilizes sophisticated technologies such as multimodal analysis and federated learning, a more thorough examination of conventional and cutting-edge fire detection approaches would improve the reader's comprehension. This may involve analyzing traditional technologies, such as photoelectric and ionization detectors, as well as modern techniques, like CNN-based models and multisensor fusion methods. The incorporation of literature on fire detection methods that protect privacy and applications for real-time video surveillance, in addition to recent progress in federated learning, processing of images, and security concerns, will enhance the comprehensive understanding of the development of fire detection technology. This comprehensive literature review will offer significant insights into the setting of the current study within the wider scope of fire detection research.

### *2.1. Fire Detection*

The research findings on fire detection indicate that vision-based sensors outperform traditional sensor types, such as light, heat, humidity, and fire sensors, exhibiting higher accuracy and fewer false alarms. Conversely, fire detection utilizing chemical sensing may provide faster alarm signals, particularly in scenarios where fire-indicating elements are emitted before smoke particles. Recognizing that the majority of fire-related casualties result from toxic emissions rather than burns, chemically based fire detection could offer an additional layer of protection for individuals within a structure. The evaluation of research in this field considers two distinct dimensions: sensor-based fire detection systems and image-based fire detection systems.

The sensing principle is typically the determining factor for how sensitive the fire alarm is, how quickly it reacts, and how reliable it is. Photoelectric and ionization fire alarms were extensively contrasted against one another under controlled conditions [6] to create formal benchmarks between the sensing principles that they employ. Particle detectors that are sensitive to a certain distribution of particle sizes are what smoke detectors are, and they can be thought of in this context. In most cases, the fire alarm goes off when the signal from the sensor hits a certain predetermined threshold. Because of this, these systems have difficulty distinguishing between particles that are the product of fires and particles that are not the result of fires when the particles are of comparable size or have similar refractive indices. For instance, smoke detectors exhibit sensitivity not just to smoke but also to dust and water vapor [7].

In addition, they are unable to differentiate between combustion products created under controlled conditions, such as cigarette smoke or certain cooking operations, and combustion products produced under situations when there is a risk of a fire. There is the potential for additional sensors to be added to smoke detectors to increase the level of specificity of the fire alarm. For instance, frequent annoyance scenarios like cooking aerosols, water vapor (from cooking or showers), and dust sources all contribute to an increase in light obscuration, but they do not result in an increase in CO concentration. As a result, CO detection can be utilized to enhance the tolerance to false alarms and to reject false warnings brought on by circumstances that do not generate CO. Systems that rely on a single measurement from a single fire sensor are not suited for fire detection because they produce an unacceptably high number of false alerts. This contrasts with smoke-based fire alarms, which produce fewer false alarms. For instance, a fire detection system that relied solely on CO measurements would fail to detect flame fires and would be overly sensitive to the exhaust fire produced by fire or oil furnaces. Because of this, fire-based systems call for the use of many sensors or multicriteria approaches, both of which necessitate more complicated data processing techniques.

In the field of research dealing with image-based fire detection systems, it has been found that the effectiveness of fire detection systems can frequently be improved by focusing on algorithms that analyze data. Traditional methods of fire detection based on vision rely on three basic classification steps: (1) characterization of the fire zone, (2) detection of edges, and (3) classification. Table 1 comprises various methods used for fire detection in different scenarios.

**Table 1.** Research related to fire detection methods.

Ref.	Contributions	Methods Used	Results
[8]	Multiple fire detection methods utilizing scalar and vision sensors have been discussed.	Scalar sensor-based approaches analyze data from flame, smoke, temperature, and particle sensors to detect fires.	Scalar sensor-based approaches are cost-effective and simple to implement but are only suitable for indoor scenarios and require human interaction for alarm confirmation
[9]	Proposed an efficient VGG-based model (E-FireNet) for fire detection. Conducted comprehensive experiments and compared performance with state-of-the-art models.	Preprocessing of collected fire images to increase the number of samples. - Utilization of an efficient CNN model for fire detection and classification.	E-FireNet achieves 0.98 accuracy, 1 precision, 0.99 recall, and 0.99 F1-score. - The proposed model shows convincing performance in terms of accuracy, model size, and execution time.
[10]	The provided paper proposes a modified YOLOv5s model for efficient fire detection in smart cities, achieving promising results with lower complexity and smaller model size.	Modified YOLOv5s model with integrated Stem module, smaller kernels, and P6 module  - Re-implementation of 12 different state-of-the-art object detection models for comparison.	The proposed modified YOLOv5s model achieves promising results with lower complexity and smaller model size.  - The proposed model had better detection performance compared to other state-of-the-art object detection models.
[11]	The optimized fire attention network (OFAN) is proposed as a lightweight and efficient convolutional neural network (CNN) for real-time fire detection. - It uses dilated variants of convolution layers and additional dense layers to capture global context and optimize weight.	The OFAN is calibrated for real-time processing using a lightweight feature extractor backbone model.	The OFAN outperforms state-of-the-art fire detection models, achieving high accuracies on three widely used fire detection datasets. It achieves accuracies of 96.23, 96.54, and 94.63 on BoWFire, FD, and the newly proposed DiverseFire dataset, respectively.
[12]	The paper introduces the optimized dual fire attention network (DFAN) for efficient fire detection and provides a medium-scale fire classification benchmark dataset.	Dual fire attention network (DFAN) for effective and efficient fire detection. - Modified spatial attention mechanism to enhance discrimination potential of fire and non-fire objects.	The DFAN provides the best results compared to 21 state-of-the-art methods. The proposed dataset advances traditional fire detection datasets by considering multiple classes.
[13]	The authors propose a novel efficient lightweight network called FlameNet for fire detection in smart city environments.	FlameNet works in two steps: first, it detects the fire using the FlameNet network, and then an alert is sent to the fire, medical, and rescue departments.	The newly developed Ignited-Flames dataset is used for analysis, and the proposed FlameNet achieves 99.40% accuracy for fire detection. The empirical findings and analysis of factors such as model accuracy, size, and processing time support the suitability of the FlameNet model for fire detection.

Table 1. Cont.

Ref.	Contributions	Methods Used	Results
[14]	Proposed an improved federated learning algorithm (FedVIS) for fire detection and localization.	Improved federated learning algorithm incorporating computer vision: FedVIS - Federated dropout and gradient selection algorithm to reduce communication overhead.	The proposed FedVIS outperforms other federated learning methods in terms of detection effect and communication costs. The model's robustness and generalization to heterogeneous data are improved.
[15]	The paper is about the construction of a large-scale Flame and Smoke Detection Dataset (FASDD) for deep learning in fire detection.	Construction of a 100,000-level Flame and Smoke Detection Dataset (FASDD). Formulation of a unified workflow for preprocessing, annotation, and quality control of fire samples.	Most object detection models trained on FASDD achieve satisfactory fire detection results. YOLOv5x achieves nearly 80% mAP@0.5 accuracy on heterogenous images.
[16]	The paper focuses on using an Improved convolutional neural network (ICNN) and LGBM Classifier for real-time fire recognition.	Improved convolutional neural network (ICNN). - LGBM Classifier.	The suggested technique effectively recognized and alerted the public to the occurrence of devastating fires. The suggested technology proved to be effective in protecting smart cities and detecting fires in the urban environment.
[17]	The paper is about the development of a deep learning model called SmokeyNet for detecting smoke from wildland fires using multiple data sources.	SmokeyNet: Baseline model for smoke detection using image sequences. - SmokeyNet Ensemble: Combines baseline model with GOES-based fire predictions and weather data.	The paper presents the results of experiments on the SmokeyNet model. The results show that incorporating weather data improves performance in terms of accuracy and time-to-detect.
[18]	Proposed a method to reduce false positive fire alarms and designed an electronic circuit with 6 sensors to detect 7 physical sensory inputs.	Implementation of fusing and classifying sensor data using machine learning models. - Comparison of multilayer perceptron, support vector machine, and radial basis function network.	Multilayer perceptron is the best model with 96.875% classification accuracy.
[19]	A vision-based fire detection framework for private spaces is proposed. - The framework preserves the privacy of occupants using a near infra-red camera.	Vision-based monitoring with convolutional neural network and other machine learning algorithms - Near infra-red camera for image capture while preserving privacy.	Developed a novel system incorporating spatial and temporal properties of fire. Validated the lightweight nature of the system through a real-world implementation.
[20]	The paper proposes an indoor fire video recognition method based on a multichannel convolutional neural network.	Designing a convolutional neural network (CNN) model. Recognition training on image features of each channel. Fire identification using flame color feature, circularity feature, and area change feature.	Solves the problem of low recognition accuracy in existing fire video recognition technology. Can be applied to indoor fire video recognition.
[21]	Proposed a multisensor fusion indoor fire perception algorithm named TCN-AAP-SVM. Considered time dimension information through trend extraction and sliding window. Addressed shortcomings of existing fire classification algorithms.	Improved temporal convolutional network (TCN). Adaptive average pooling (AAP)—support vector machine (SVM) classifier.	Proposed algorithm improves fire classification accuracy by more than 2.5%. Proposed algorithm improves fire detection speed by more than 15%. Outperforms TCN, BP neural network, and LSTM in accuracy and speed.

Table 1. Cont.

Ref.	Contributions	Methods Used	Results
[22]	Proposed system achieves high precision, recall, and F1 scores for fire detection. System reduces false alarms and improves early fire detection.	Multimodal sensors are integrated to acquire data of carbon monoxide, smoke, temperature, and humidity. Support Vector Machine (SVM) is used for data analysis and classification.	Precision: 99.8%—Recall: 99.6%—F1 score: 99.7%.
[23]	Effective fire detection using deep learning techniques in smart cities. Use of unmanned aerial vehicles (UAVs) for wide area coverage. Highlighting the most important fire regions using multiheaded self-attention.	Deep multiscale features from a backbone model are employed. Attention mechanism is applied for accurate fire detection. Features fusion is used to represent the image effectively. Multiheaded self-attention enhances the fused features.	Preliminary experimental results demonstrate effective performance of the proposed model. The proposed model outperforms rivals in fire detection accuracy.
[24]	Upgraded the classic model by adding LGBM in the final layer. Developed a real-time fire catastrophe monitoring system. Altered the network structure for effective fire recognition under different weather conditions.	Improved convolutional neural network (ICNN)—LGBM Classifier—Data augmentation methods. Automated color enhancement—Parameter reductions.	The technique effectively detects fire areas and provides early warnings. The suggested technology is effective in protecting smart cities and detecting fires in urban environments. Tested the system against previously published fire detection methods.
[25]	Proposed a novel optimized Gaussian probability-based threshold convolutional neural network (GTCNN) model for fire detection.	Sensor-based methods for fire detection. Computer vision-based approaches using surveillance camera-based video (SV).	The proposed optimized GTCNN achieves a detection accuracy of 98.23%. The optimized GTCNN outperforms other deep learning networks in terms of accuracy.

From the research related to fire detection, it is observed that none of the approaches is geared towards protecting the confidentiality of fire detection in any way. Research related to federated learning is given in the next section.

## 2.2. Federated Learning

When it comes to image processing, federated learning ensures the confidentiality of the data that is necessary to train the model [26]. Real-time prediction, protection of data privacy and security, the ability to do offline prediction, and the provision of an intelligent framework are the key advantages brought about by the utilization of federated learning in image processing applications [27]. In the case of federated learning, each prediction is executed on the edge device; hence, there is no need to be concerned about delays in the data transfer process. In addition, because federated learning organizes and runs its own training, the only thing that must be transferred is the model.

By decentralizing data collection from a centralized server to individual clients, federated learning provides a way to protect the privacy of individual users. This paradigm was developed because of the combination of two important elements [28]: The inability to keep sufficient data centrally on the server side due to limits placed on direct access to such data and, if network asynchronous communication is involved, securing sensitive data by utilizing local data from clients rather than forwarding it to the server to protect it.

The process of prediction continues even if the gadget is not connected to the internet. Therefore, there is no reason to worry about the gadget, whether one is connected to the internet or not. If the model has access to the various input devices, it will be able to successfully perform the assignment. Because federated learning does not rely on any certain kind of complicated hardware to function, the requirements for the infrastructure needed to support it are extremely simple. The federated learning algorithms that are most used are as follows.

After completing the local training with the data that is available there, the FedSGD algorithm [29] transfers the results of the training to the server side from the client side. It waits before contributing to the average joint aggregate, which it does. This process of waiting could result in a lengthy wait for a variety of reasons; each batch selects a subset of the nodes to participate in an epoch training and then uploads all the nodes to the server. To obtain a new weight, the server will first add and then sum up all the weights, and then it will distribute the new weight to each node. To train the new epoch, each node will replace the distributed node it was assigned with the weight that was calculated by the previous epoch. It continues with the previous three procedures until the server notifies the user that the weights have converged.

On the client side, the original data is split up into numerous portions using the FedAVG algorithm [30], which is derived from the FedSGD algorithm. It is the method that is used the most frequently in federated learning for the purpose of model optimization [10]. Before updating and dispersing information locally, this method calculates an average of the data that was locally uploaded for the stochastic descent gradient. It has been demonstrated to be effective at learning across multiple tasks at once. On the other hand, the FedAvg algorithm itself has a few flaws, including global model instability and a slow convergence rate when applied to diverse datasets.

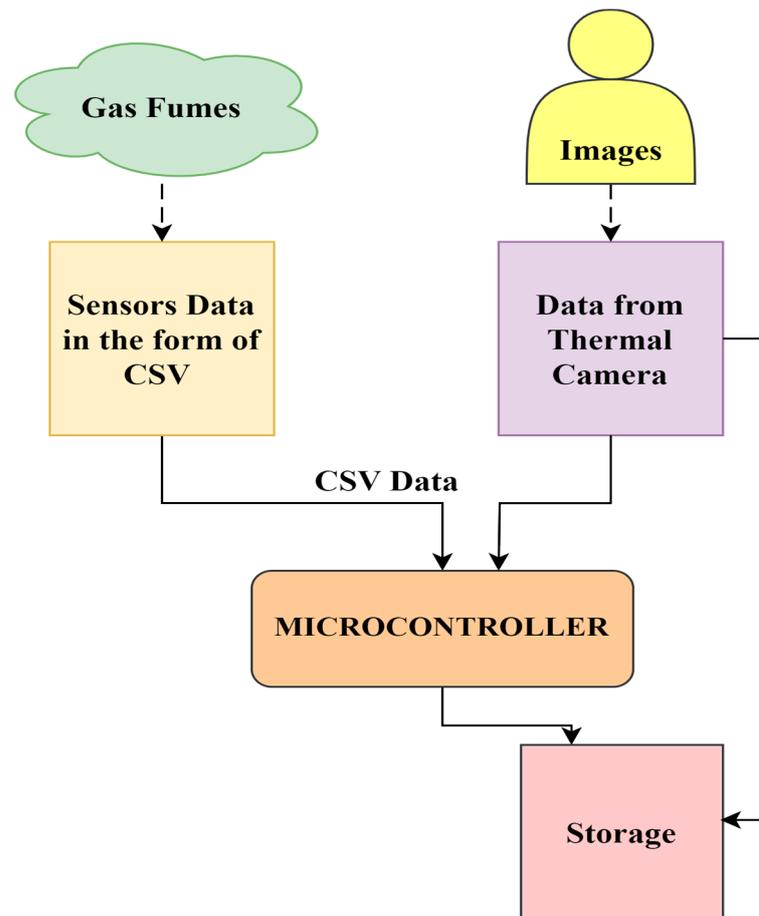
Furthermore, when applied to the scene of a fire monitoring system, the current federated learning methods (including the traditional federated learning algorithms FedSGD and FedAVG, among others) struggle with low computational efficiency as a result of their lengthy training period and poor cooperative training effect. Both of these elements diminish the overall effectiveness of the training process, which presents a challenge. The communication bandwidth is significantly burdened by the substantial quantity of redundant and irrelevant information present in the local model updates uploaded by the client. Consequently, to comply with the algorithm, the client is obligated to filter the local model updates using the previous round of global model correlation and refrain from uploading local model updates that do not satisfy the threshold.

Communication mitigated federated learning (CMFL) is an algorithm that Wang et al. [31] proposed for the utilization of local model updates uploaded by the client. To train machine learning models with private data from numerous dispersed devices, federated learning is implemented. This is achieved using federated learning. A variety of factors can contribute to data heterogeneity in the context of fire detection, such as the utilization of multiple cameras, monitored environments, and differing levels of illumination. This decentralized strategy guarantees that the initial data remains only on individual devices, thereby resolving privacy concerns linked to centralizing sensitive information.

FL ensures robust user privacy by employing collaborative learning without sharing raw data. Regarding the classification of fire leaks, this methodology prioritizes privacy by adhering to ethical data management standards. Additionally, it encourages wider participation by enabling varied datasets to contribute while safeguarding individual privacy.

### 3. Proposed Work

The proposed work involves compiling the dataset by creating a controlled environment as mentioned in Figure 1. Despite this, emphasizing the importance of fire prevention is crucial, as deliberately starting a fire is considered risky and, in many regions globally, potentially illegal. The ideal approach for training in fire safety involves safe and controlled conditions, access to educational materials, and guidance from trained professionals. The system setup configurations used for the proposed work are shown in Table 2.



**Figure 1.** Block-level connections for dataset collection setup.

**Table 2.** System setup configurations.

S.No.	Component	Configurations
1	1 Server Computer	Core i7, 32 GB RAM, NVIDIA 3070 8 GB Graphics Memory
2	5 Client Computer's	Core i5, 16 GB RAM, NVIDIA 1650 4 GB Graphics Memory
3	Python Programming	Version 3.7
4	Keras	Version 3.0
5	TensorFlow	Version 2.14
6	TensorFlow Federated	Version 1.0
7	Camera	5 MP HD

The block-level connections within this dataset collection setup enable real-time data acquisition, preprocessing, and logging. By integrating information from fire detectors and thermal imagery, this setup ensures a holistic dataset that captures diverse aspects of fire-related phenomena. The well-established connections contribute to the dataset's reliability and richness, providing a robust foundation for training and testing multimodal fire detection models. The dataset collection setup involves a sophisticated integration of multiple fire detectors and a thermal camera, strategically positioned to ensure comprehensive coverage of fire-related data.

At the block level, the connections within this setup can be described in the following key components: Eleven metal-oxide fire detectors (Sensor1 to Sensor12) are deployed, each with specific sensitivities to diverse fire-indicating elements. Block-level connections involve wiring and communication interfaces to gather data from these sensors. These

connections ensure the transfer of real-time information on the presence of various fire parameters. A thermal camera, a crucial component of the multimodal setup, captures thermal signatures associated with fires. Block-level connections include power supply and data communication channels to transfer thermal images to the central processing unit.

### 3.1. Dataset

Experimentation often involves recreating scenarios with ignited combustible materials in controlled settings, such as those found in specialized fire training facilities. These facilities are equipped with safety measures, fire suppression systems, and trained personnel to ensure the well-being and safety of participants. In various industries, including the IoT, home automation, and scientific research, among others, it is a common practice to collect data using sensors and Arduino. Thermal cameras capture images of fires, hotspots, and smoke, facilitating early fire detection and monitoring. The datasets obtained from thermal photography can be valuable for training models focused on fire prevention and control. The dataset utilized in this research was created during the NASA Space Apps Challenge in 2018.

The primary goal was to facilitate the development of a model capable of distinguishing images containing fire (*fire\_images*) from regular images without fire (*non-fire\_images*). The dataset is structured for binary classification and is divided into two folders, with 5000 outdoor-fire augmented images in the *fire\_images* folder and 5000 augmented images in the *non-fire\_images* folder. Notably, the dataset exhibits class imbalance, necessitating careful consideration during model training and validation [32]. The sensors dataset is available in [33,34] having 5000 sensor data entries for the fire category and 5000 sensor data entries for the non-fire category as shown in Table 3.

**Table 3.** Dataset entries of images and sensors.

Category	Number of Image Entries	Number of Sensor Data Entries
Fire	5000	5000
Non-Fire	5000	5000

### 3.2. Multimodal Fire Detection Dataset

The focal point of the present study is the dataset compiled using multiple fire detectors and a thermal camera. This dataset forms the basis for a multimodal fire detection investigation, where the fire sensors and a thermal camera work collaboratively to gather comprehensive information about the presence of a fire. At the heart of this research is the development and application of a multimodal fire detection dataset, a critical element in advancing the capabilities of fire detection systems. The dataset is meticulously curated, drawing on diverse sources and specifically incorporating data from numerous fire detectors and a thermal camera. This multimodal strategy plays a pivotal role in augmenting the accuracy and reliability of fire detection models by providing a holistic perspective on fire-related phenomena.

#### 3.2.1. Fire Sensors Integration

The dataset's enrichment is a result of deploying a sophisticated array of eleven metal-oxide sensors and a thermal camera, strategically positioned in the experimental environment. This combination of sensors is meticulously designed to capture crucial information related to various fire parameters, providing a holistic view of fire-related phenomena. The incorporation of these sensors addresses the limitations associated with individual sensor types, ensuring a comprehensive coverage of fire-related data for the multimodal dataset. The deployed sensors, including Sensor1/MQ-2, Sensor2/MQ-3, Sensor2/MQ-4, Sensor5/MQ-5, Sensor6/MQ-6, Sensor7/MQ-7, Sensor8/MQ-8, Sensor9/MQ-9, Sensor10/MQ-135, Sensor11/MQ-138, and Sensor12/MQ-139, exhibit sensitivity to a wide range of fire-indicating elements, as outlined in Table 4. This diverse set of sensors allows

the dataset to capture variations in chemical volatiles and other early indicators that precede the formation of smoke particles during a fire event.

**Table 4.** Sensors and corresponding sensitive fire.

Used Sensor	Fire Sensitive to Sensor
Sensor1/MQ-2	Liquefied petroleum fire, methane fire, butane fire, smoke
Sensor2/MQ-3	Smoke, ethanol, alcohol
Sensor2/MQ-4	Methane, CNG fire
Sensor5/MQ-5	Liquefied petroleum fire, natural fire
Sensor6/MQ-6	Liquefied petroleum fire, butane fire
Sensor7/MQ-7	Carbon monoxide fire
Sensor8/MQ-8	Hydrogen fire
Sensor9/MQ-9	Carbon monoxide fire, flammable fire
Sensor10/MQ-135	Air quality (CO, ammonia, benzene, alcohol, smoke)
Sensor11/MQ-138	Benzene, toluene, alcohol, acetone, propane, formaldehyde, hydrogen
Sensor12/MQ-139	Infra-red flame

The specific attributes of each sensor, such as sensitivity to liquefied petroleum fire, methane fire, smoke, and various other fire-indicating elements, contribute to the dataset's richness. This diverse sensor array, combined with the thermal camera, ensures that the multimodal dataset encompasses a broad spectrum of fire-related information. The dataset, thus curated, becomes a powerful resource for training models to recognize and respond to the complex array of cues associated with fire occurrences in different environments.

The dataset gathering setup's sensors are sensitive to fire-indicating factors, enhancing the fire detection system. Liquefied petroleum, methane, butane, and smoke are detected by Sensor1/MQ-2. Sensor2/MQ-3 detects smoke, alcohol, and ethanol, whereas Sensor5/MQ-5 detects natural and liquefied petroleum fire. MQ-7, MQ-8, and MQ-9 sensors detect the presence of carbon monoxide, the gas hydrogen, and ignited fire. Sensor10/MQ-135 also samples air for gases such as carbon monoxide, alcohol, ammonia, benzene, and smoke. The sensors themselves, including MQ-138 for toluene, benzene, and alcohol and MQ-139 with infrared flame, enrich the dataset and allow the model to gather a wide range of related data.

### 3.2.2. Thermal Camera

In the dataset collection setup, a thermal camera stands as a pivotal component, contributing essential data for the multimodal fire detection dataset. The thermal camera utilized in this study employs advanced infrared technology to measure temperature fluctuations. Unlike conventional cameras, every pixel on the image sensor of a thermal camera serves as an infrared temperature sensor, allowing simultaneous temperature measurement for each point within the camera's field of view. The thermal camera operates based on the principles of infrared light detection, capturing temperature variations in the environment. Each pixel functions as a discrete temperature sensor, generating images in a temperature-based format rather than traditional RGB. This approach provides a direct representation of temperature differences across the captured scene.

Thermal cameras can operate effectively in diverse environments, unaffected by factors such as shape or texture. Unlike conventional image cameras, thermal cameras are not limited by darkness, making them suitable for applications in low-light conditions. The thermal camera employed in this study features a 36-degree field of view, allowing for a wide coverage area during data collection. It has a measurement range spanning from 40 °C to 330 °C, enabling the detection of temperature variations associated with fire events. The camera operates at a framerate of 9 Hz, providing real-time data acquisition capabilities.

With a total of 32,136 thermal pixels and 206,156 thermal sensors, the camera ensures detailed and accurate thermal imaging for the creation of the multimodal dataset. Data for training and testing the fusion model are collected simultaneously using both the thermal camera and the deployed fire sensors.

The thermal camera captures thermal signatures, contributing valuable temperature-related information to the dataset. The thermal camera operates in tandem with the fire sensors, providing a complementary source of information for the multimodal dataset. The combined data from thermal imaging and fire sensors enhances the dataset's richness, allowing the model to learn from both visual and chemical cues associated with fires. The thermal camera, along with fire sensors, plays a crucial role in the gathering of data for model training and testing. Detailed data gathering and preprocessing methodologies are elaborated upon in subsequent sections of the paper, ensuring transparency and reproducibility in the research process. The incorporation of a thermal camera in the dataset collection setup adds a crucial dimension to the multimodal dataset, enabling the model to learn from temperature changes associated with fire events. This combination of thermal imaging and fire sensor data enhances the dataset's comprehensiveness, contributing to the development of a robust and accurate fire detection model.

### *3.3. Preprocessing of Multimodal Data*

The preprocessing of multimodal data is a crucial step in preparing the dataset for effective model training. In this study, the preprocessing pipeline involves transforming the numerical fire readings from the seven metal-oxide (MOX) sensors into heatmap images, followed by scaling these images along with the infrared (IR) thermal images to fit the input layer sizes of various convolutional neural network (CNN) variations. The fire readings from the seven MOX sensors are initially transformed into heatmap images. At regular intervals of 2 s, each numerical fire measurement is converted into an RGB image, where the numerical values are mapped to color intensity values on the RGB scale. This mapping process creates a colormap pattern (heatmap) for each sensor, resulting in RGB images, which are then saved with the .jpg extension. Both the generated MOX sensor heatmap images and IR thermal images are scaled to fit the input layer sizes of six different CNN variations. Scaling is a crucial step to ensure uniformity in input dimensions across different CNN architectures, facilitating consistent model training and evaluation.

The preprocessed data is divided into training and testing portions using a 70–30% split. This division ensures that a significant portion of the data is allocated for training the model, while a separate portion is reserved for evaluating its performance. Augmentation is employed as a critical step to enhance the training performance of the CNNs. By increasing the number of images in the training dataset through augmentation, the models become more robust, less prone to overfitting, and better equipped to generalize to a variety of fire scenarios. The dataset, comprising both MOX sensor heatmap images and IR thermal images, is now prepared for feeding into CNN variations for training and testing.

The combination of these multimodal data sources enriches the learning process, enabling the model to capture both visual and chemical cues associated with fires. The scaled images are adapted to the specific input layer sizes of the chosen CNN variations. This adaptation ensures that the multimodal data is effectively processed by each CNN architecture, optimizing their ability to learn and extract meaningful features for fire detection. The preprocessing steps involve the transformation of MOX sensor readings into heatmap images, scaling of these images and IR thermal images, data splitting for training and testing, and the augmentation of training data. These steps collectively enhance the dataset's suitability for training robust CNN models capable of effectively detecting fires based on both visual and chemical information [35].

### *3.4. Data Classification for Multimodal System for DL Models*

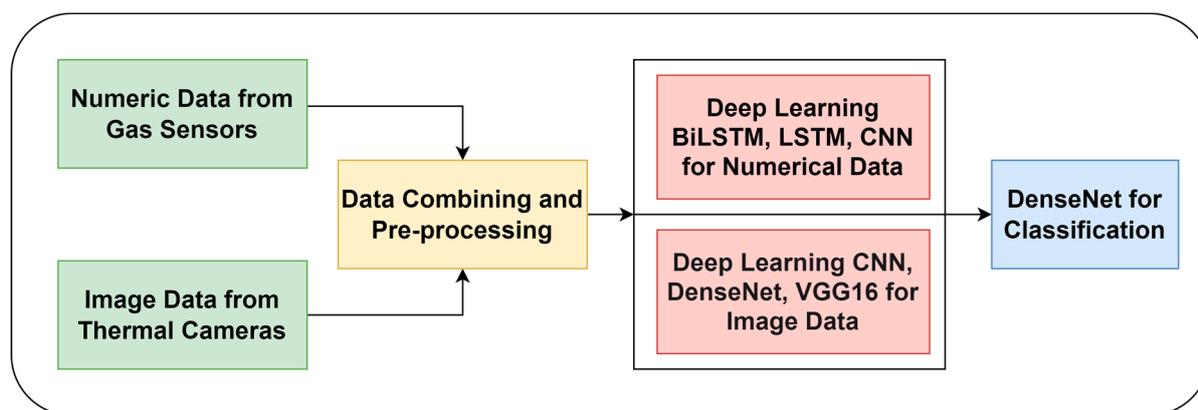
In the data classification phase of the multimodal system, the objective is to integrate information from both the thermal camera and the fire sensors to create a comprehensive

dataset for deep learning models. The process involves gathering data on the presence of a fire, extracting relevant information from numerical data obtained from fire sensors, and visual images from the thermal camera. The subsequent steps include feature extraction, data normalization, and data cleaning to prepare the data for multimodal input representation in deep learning models. The thermal camera and fire sensors are combined to collect data on the presence of a fire. Information from each data source, including numerical data from fire sensors and visual images from the thermal camera, is extracted to form a multimodal dataset. The next step involves feature extraction, where pertinent information is identified and extracted from both types of data. Data normalization and cleaning procedures are applied to ensure consistency and eliminate noise from the dataset.

A multimodal input representation for deep learning models is created by combining the extracted features from numerical data and visual images. This representation is designed to capture both chemical and thermal aspects of fire occurrences.

Different deep learning techniques are employed for training on numerical data from fire sensors and image data from thermal sensors. LSTM, BiLSTM, and CNN are utilized for training on numerical data, while CNN, DenseNet, and VGG16 are employed for training on thermal image data. The utilization of Long Short-Term Memory is successful in processing numerical data, providing a method for accurately storing and understanding sequential relationships in fire sensor readings. This is crucial for the identification of emerging patterns. The Bidirectional LSTM improves temporal modeling by incorporating information from both preceding and succeeding sequences, hence facilitating a more holistic comprehension of sequential data obtained from fire detectors in the fire detection mechanism.

Convolutional neural networks are highly effective in analyzing visual data, which makes them ideal for extracting hierarchical characteristics from thermal pictures. This model is crucial for identifying unique patterns related to flames in various urban situations. The DenseNet architecture is used due to its dense connectivity structure, which enables the effective reuse of features and optimal utilization of parameters. This enhances the model's ability to learn when trained on thermal imaging information for fire detection. The VGG16 model, a convolutional neural network with a straightforward and consistent design, excels in capturing complex characteristics from thermal pictures. The system's efficacy stems from its capacity to identify intricate visual patterns, enhancing the resilience of the system for detecting fires. The data classification scheme is illustrated in Figure 2, outlining the flow of data processing and classification within the multimodal system.



**Figure 2.** Data classification of multimodal system.

The scheme incorporates the sequential steps of data integration, feature extraction, and multimodal input representation. Model performance is evaluated on a test dataset using metrics such as accuracy, precision, recall, and loss. These metrics provide insights into how well the trained model can detect fire leaks in the given dataset. To evaluate the performance of multimodal data, both slow and fast learning rates are employed. In

the slow learning rate approach, each modality generates a set of feature representations capturing information specific to that modality. These features are processed individually using specialized neural network designs. In contrast, the rapid learning rate neural network architecture fuses or combines features from different modalities at an early layer, allowing for faster integration of information.

The classification scheme involves the fusion of information from multiple modalities, emphasizing the integration of both chemical and thermal cues. This holistic approach ensures that the deep learning models can effectively recognize and classify fire occurrences with a comprehensive understanding of multimodal data. The data classification process, as outlined in Figure 2 and described in Figure 3, represents a systematic approach to leveraging both numerical and visual information for robust fire detection. The incorporation of multimodal learning rates and model fusion techniques enhances the versatility and accuracy of the deep learning models in detecting fire leaks across various scenarios.

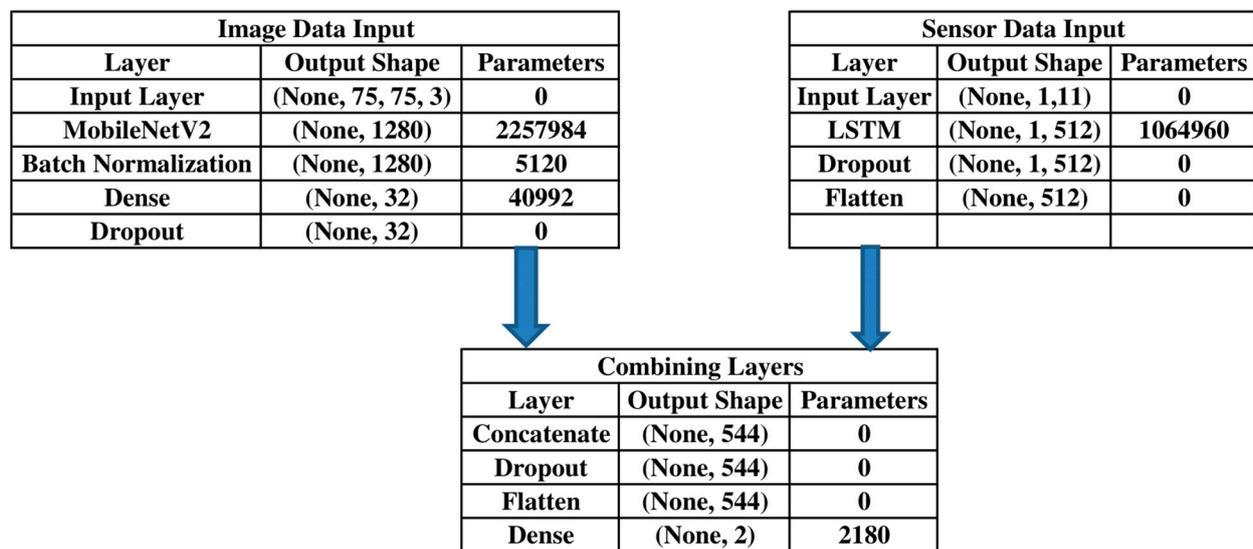


Figure 3. Configuration parameters of multimodal system.

### 3.5. Multimodal Data Classification in Federated Eco-System

In the realm of multimodal data classification, federated learning (FL) emerges as a cutting-edge distributed machine learning paradigm, gaining prominence in both academic and business settings. FL addresses challenges related to data ownership, localization, and privacy by training a high-quality centralized model using data dispersed across various locations and devices. This section explores the mathematical principles and procedures at the core of the FL paradigm and its potential application in the context of fire detection. Using data that is dispersed across a variety of places and devices, a high-quality centralized model is trained when the FL paradigm is used [36–38]. Google introduced a technique in 2016 that would take data from each site to independently compute an update to the current ML model, which is how the phrase was initially used [36].

This update is then sent back to a central service, which compiles it into a new global model and distributes it to the various locations [36]. The code is therefore brought to the data in this paradigm rather than the data being brought to the code. As a result, the FL paradigm solves issues with data ownership, localization, and privacy [37]. What follows presents and discusses the mathematical principles and procedures at the heart of the FL paradigm as well as its potential for addressing the water leak detection issue. The general architecture is first given before the FL paradigm is explained. A centralized FL server that can communicate with a collection of devices that are prepared to carry out the necessary FL task makes up the FL architecture in most cases. Six major steps make up the workflow [36,37]:

- a. The group of devices transmits a message of availability indicating that they are prepared to finish a FL task.
- b. At time  $t_i$ , the FL server selects a portion of these available devices and distributes the deep learning (DL) model to them.
- c. Following that, each device runs a training procedure using the local data to create a new local ML model.
- d. Based on the aforementioned training procedure, each device communicates the updated parameters of its machine learning model.
- e. The updated global DL model for time  $t_i$  is then calculated by the FL server by combining the local models.
- f. All devices receive the updated global DL model from the FL server.
- g. Every round, this process is repeated, with the FL server deciding how frequently to update it.

In terms of mathematics, the FL paradigm tries to learn the  $W$  matrix-representable parameters of the global ML model. To do this, a portion of the total number of  $D_{tot}$  devices is sent the model  $W_{t_{i-1}}$  by the FL server. Every device  $D_{t_i}$  goes through a training procedure to establish an updated local model  $W_{t_i}^j$ . Each device then transmits its update to the FL server using the formula  $H_{t_i}^j = W_{t_i}^j - W_{t_{i-1}}$ . The FL server then combines these local modifications to create the following global model [36,37]:

$$W_{t_i} = W_{t_{i+1}} + \alpha_{t_{i+1}} H_{t_i} \quad (1)$$

where  $\alpha_{t_i}$  is the learning rate chosen by the FL server and  $H_{t_i}$  is the average aggregated device-shared update given by

$$H_{t_i} = \frac{1}{|D_{tot}|} \sum_{j \in D_{t_i}} H_{t_i}^j \quad (2)$$

For particular implementations, the fact that the term  $H_{t_i}$  can be computed as the weighted sum of the device-shared updates as opposed to the average is of no consequence [37]. This paradigm is applicable to facilities that are dispersed across multiple regions. This is suitable for organizations that possess a multitude of manufacturing facilities dispersed across diverse geographical areas.

The FL paradigm facilitates the transfer of insights gained from a fire leak incident at one facility to other production sites, capitalizing once more on the rarity of multiple concurrent breaches occurring at distinct locations. In this scenario, each facility would operate as a FL device through the deployment of a collection of servers that are employed to conduct the local training. All of the facilities would be connected by a centralized cloud server (such as the Amazon cloud service) that acts as the FL server and compiles local models prior to returning the utilized global ML models. As mentioned earlier, machine learning detection models such as support vector machines (SVM) and artificial neural networks can be trained at the facility level.

The centralized FL server then provides the aggregated global ML model, which has been demonstrated to be an effective method for detecting leaks. It is important to acknowledge that the implementation of this architecture is anticipated to result in heterogeneous or non-iid distribution of the data (non-iid) hardware equipment capacities or capabilities of the facilities. There exist multiple approaches that can be implemented to tackle this issue. One approach is to organize comparable facilities into clusters and appoint one of them to provide updates on behalf of the group. By adopting this approach, it would be possible to mitigate the computational capacity at each site as well as the diversity of the data.

The “esgrsecond approach” entails the global sharing of a subset of the data obtained from each facility. Local models being trained at each location could thus be able to view and analyze data from other internet domains. For example, Zhao et al. [39] demonstrated

that sharing only 5% of the local data globally could significantly improve the accuracy of the global model. To ensure that the centralized FL server delivers a global model of superior quality, a comparable approach may be implemented for the multi-facility design. A variety of architectures are depicted in Figure 4, with the FL design on the left depicting it within a single facility and the FL design across multiple facilities on the right.

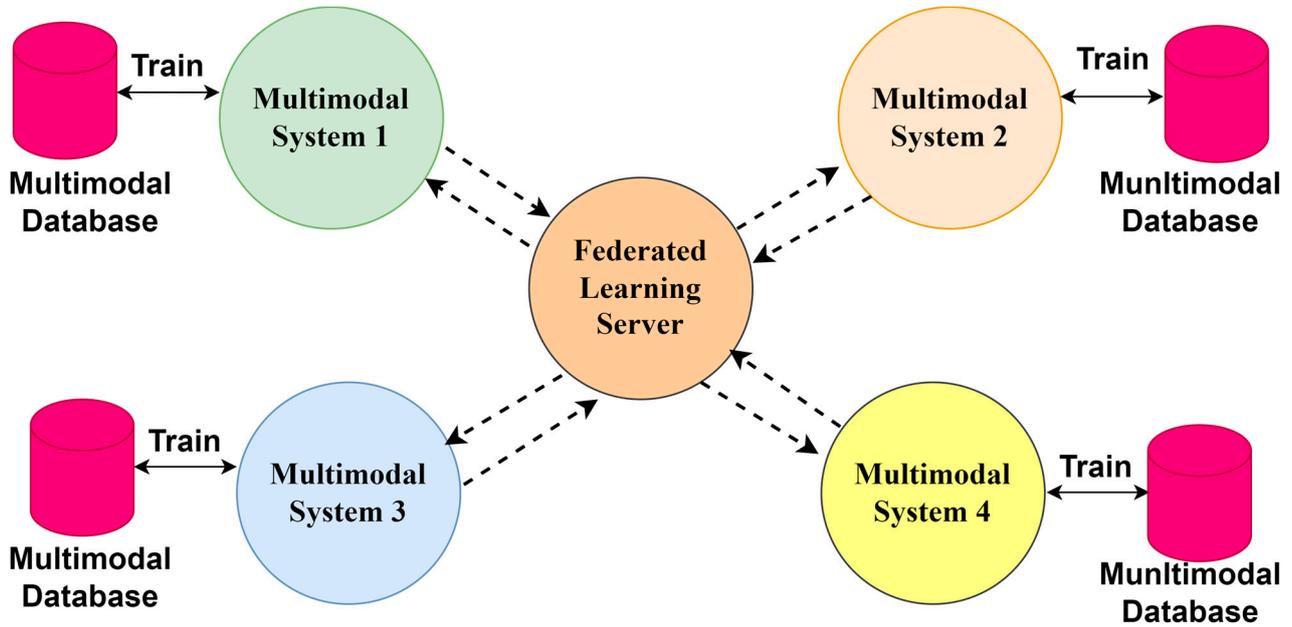


Figure 4. Proposed multimodal federated system.

#### 4. Experimental Results, Analysis, and Discussion

For each model that has been attempted, values for precision, recall, F1 score, accuracy, and loss function are shown. To validate the performance of classification, precision and recall were used, and accuracy was used to analyze as a single numerical study of a system’s completion. Categorical accuracy is another name for training accuracy. It indicates how accurately models classify the practice data. One of the most important elements of a deep neural network is the model loss function. It shows how much the models are off from the final outcome. It is possible to gauge how better CNN models are performing when predicting from a dataset using the values of the model loss. The purpose of the test accuracy is to assess how well the models perform. Following training, the top-performing CNN architectures [40,41] were chosen based on the outcome metrics as given below:

- Accuracy indicates how well the model can identify the correct label for each sample, and it is calculated using Equation (3) [4,5,9,13,14].

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN) \tag{3}$$

where, True Positive (TP) is the count of samples that are correctly classified as positive. True Negative (TN) is the count of samples that are correctly classified as negative. False Positive (FP) is the count of samples that are wrongly classified as positive. False Negative (FN) is the count of samples that are wrongly classified as negative.

- Precision is a performance metric that calculates the proportion of correctly identified positive samples to the total number of positive samples predicted by the model. It measures how accurate the model is in identifying the relevant samples. The formula for precision is shown in Equation (4) [4,5,9,13,14].

$$\text{Precision} = TP / (TP + FP) \tag{4}$$

- Recall, also known as sensitivity or true positive rate, is a metric that measures the proportion of actual positive samples that are correctly identified by the model. It is calculated by dividing the number of true positive predictions by the total number of actual positive samples in the dataset, as shown in Equation (5) [4,5,9,13,14].

$$\text{Recall} = \text{TP}/(\text{TP} + \text{FN}) \tag{5}$$

- The F1 score is a metric that considers both precision and recall by taking their harmonic mean. This score is useful for evaluating the performance of a model when the dataset is imbalanced. The formula for the F1 score is given by Equation (6) [4,5,9,13,14].

$$\text{F1-Score} = 2 \times (\text{Precision} \times \text{Recall})/(\text{Precision} + \text{Recall}) \tag{6}$$

Our main goal was to improve test accuracy while reducing model loss function. Adam served as the optimizer for all models, which were run for 10 iterations with a learning rate of 0.0001. Multiple data from the same instance are available at once in the multimodal system. In addition, the data is of various types, including photos and sensor data obtained from fire sensors and thermal imaging, respectively. To train models for both types of data concurrently, multimodal data is fed to multimodal deep learning models, which then produce findings in the form of common classification.

#### 4.1. Analysis on Unimodal Data

Collecting and analyzing individual sensor data is fundamental for effective fire leakage detection systems. This individual data is crucial for the accurate and timely identification of potential hazards. Here, image data from thermal cameras and sensors data from fire sensors for fire leakage detection have been collected.

##### 4.1.1. Image Data Analysis

CNN model and six pre-trained CNN models for fire detection from image data—DenseNet201, InceptionResNet, MobileNetV2, VGG16, VGG19, and Xception—have been chosen. Table 5 shows the accuracy, precision, recall, and loss of our six stated CNN models. Excellent outcomes were provided by all models. After checking the precision value, all six models outperformed CNN because they all received precision values of greater than 0.99. This indicates that the models were able to forecast the greatest number of predictions for no fire leakage that truly fall into the no fire leaking class. In comparison to other techniques, the DenseNet201, VGG19, and Xception models have better loss values.

**Table 5.** Comparative analysis of various methods used for image data.

	Accuracy	Val Accuracy	Loss	Val Loss	Precision	Val Precision	Recall	Val Recall
<b>Convolutional Neural Network</b>	94.95	90.6	0.396	0.3878	94.95	90.6	94.95	90.6
<b>DenseNet201</b>	99.66	100	0.0843	0.0641	99.66	100	99.66	100
<b>MobileNetV2</b>	100	99.33	0.1018	0.1113	100	99.33	100	99.33
<b>XceptionNet</b>	99.92	97.99	0.09	0.1148	99.92	97.99	99.92	97.99

Figure 5 illustrates the comparison of training accuracy, test accuracy, model loss function, validation loss, precision, validation precision, recall, and validation recall to demonstrate why seven models perform better. The comparison of training accuracy for each of the seven designs is displayed in the training accuracy graph. The upward movement of the lines of the DensNet201 and Xception models in this example demonstrated that the models were picking up new information from the training data relatively quickly. CNN is indicated on the blue line as having the lowest training accuracy (83.00%).

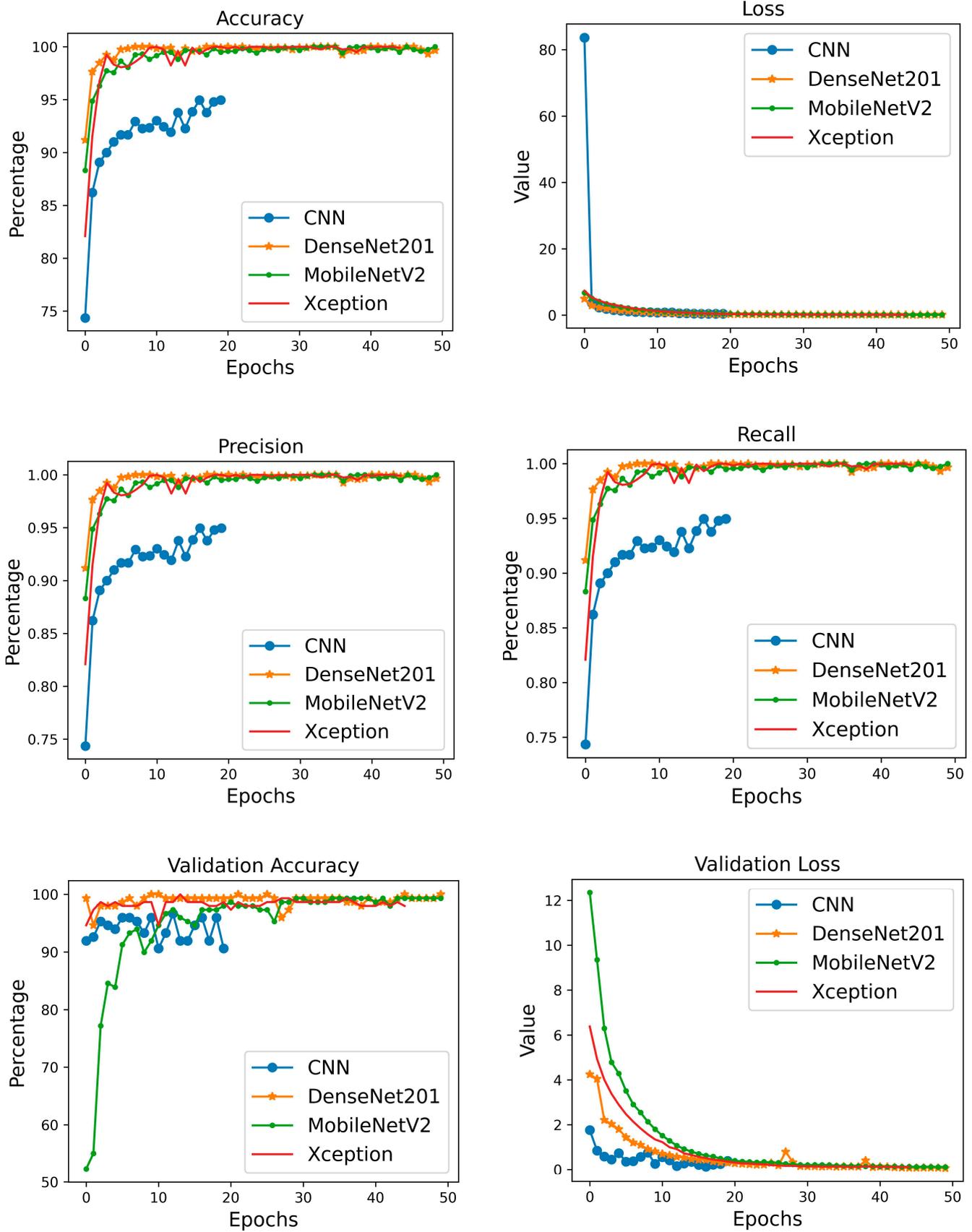
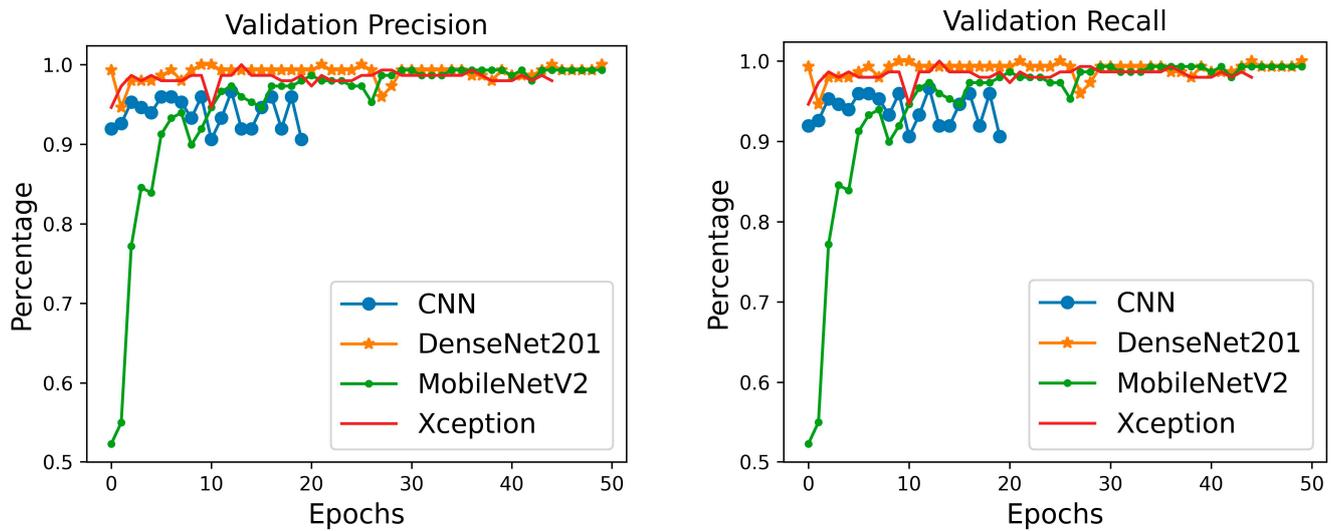


Figure 5. Cont.



**Figure 5.** Comparison of various methods used for validating image data.

Every model, however, began at a low value and ended up at a greater one. Lower loss functions were achieved as training time on the models increased. The comparison of seven architectures for test accuracy is shown in the second figure. Every model's graph is moving upward in each epoch, which shows that our proposed model did a great job of identifying fire leaks. These numbers show that our models were correctly trained on the fire dataset. The models accurately identified fire from the test dataset without being either over-fit or under-fit. These are the explanations for why the models' output in the classification report and confusion matrix was so outstanding.

The evaluation metrics showcase promising trends in the performance of the models. Regarding training accuracy, both the DensNet201 and Xception models exhibit consistent upward trends, indicating their quick adaptation to new information from the training data, suggesting effective learning capabilities. In terms of test accuracy, all models consistently demonstrate upward trends, showcasing their ability to effectively identify fire leaks without encountering issues of overfitting or underfitting. The observed lower loss functions with increased training time across all models signify improved overall performance, as the models minimize errors and enhance their predictive capabilities. Additionally, precision and recall values for all models are notably high, highlighting the models' accurate identification of fire incidents and reflecting a robust performance in both precision (minimizing false positives) and recall (minimizing false negatives).

These combined results indicate the effectiveness of the models in learning and accurately identifying fire incidents, affirming their potential for practical deployment in fire detection scenarios. The comparative analysis illustrates the superior performance of pre-trained models, particularly DenseNet201, VGG19, and Xception, in comparison to the baseline CNN model. These models exhibit high accuracy, precision, and recall values, indicating their efficacy in accurately detecting fire incidents. The graphical representation in Figure 5 provides a visual understanding of the models' training and validation performance, further emphasizing their robustness in fire leakage detection.

#### 4.1.2. Sensors Data Analysis

Three deep neural networks, namely BiLSTM-Dense, Dense, and LSTM-DenseDenseNet201, for the task of identifying fire leakage from sensors data have been chosen. Table 6 shows the accuracy, precision, recall, and loss of these three deep learning models. Excellent outcomes were provided by all models. All six models outperform BiLSTM-Dense after checking the precision value; they all received a precision value of over 93.39, indicating that the model could correctly forecast the greatest number of no fire leakage predictions that really fall into the no fire leakage class. The best of the three approaches, BiLSTM-Dense, has a loss value of 0.15. BiLSTM-Dense

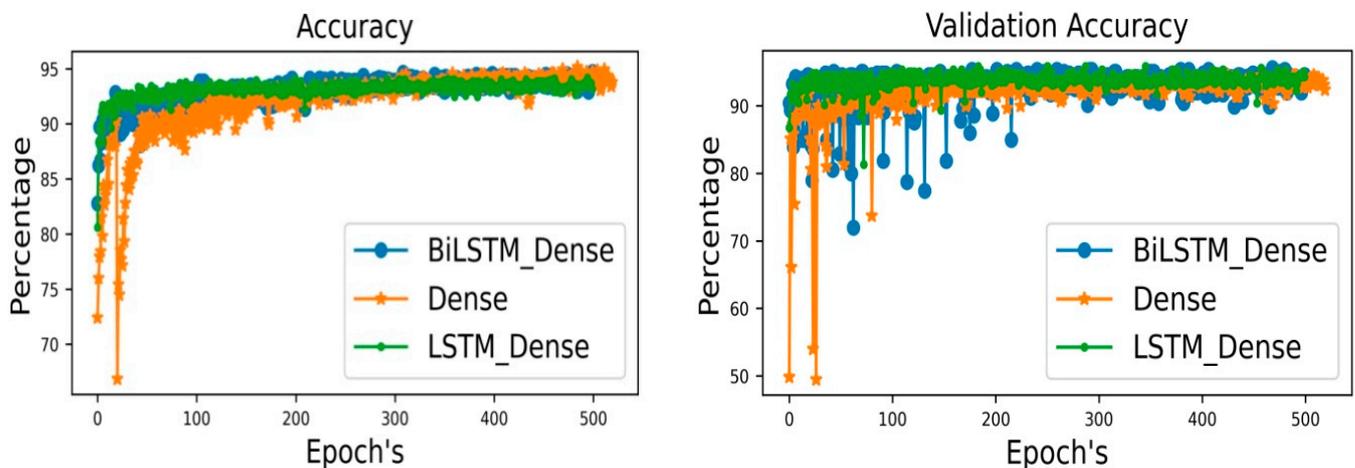
has the highest accuracy, making it the most effective method for finding fire leaks. Figure 6 illustrates the comparison of training accuracy, test accuracy, model loss function, validation loss, precision, validation precision, recall, and validation recall to demonstrate why seven models perform better.

**Table 6.** Comparative analysis of various methods used for sensors data.

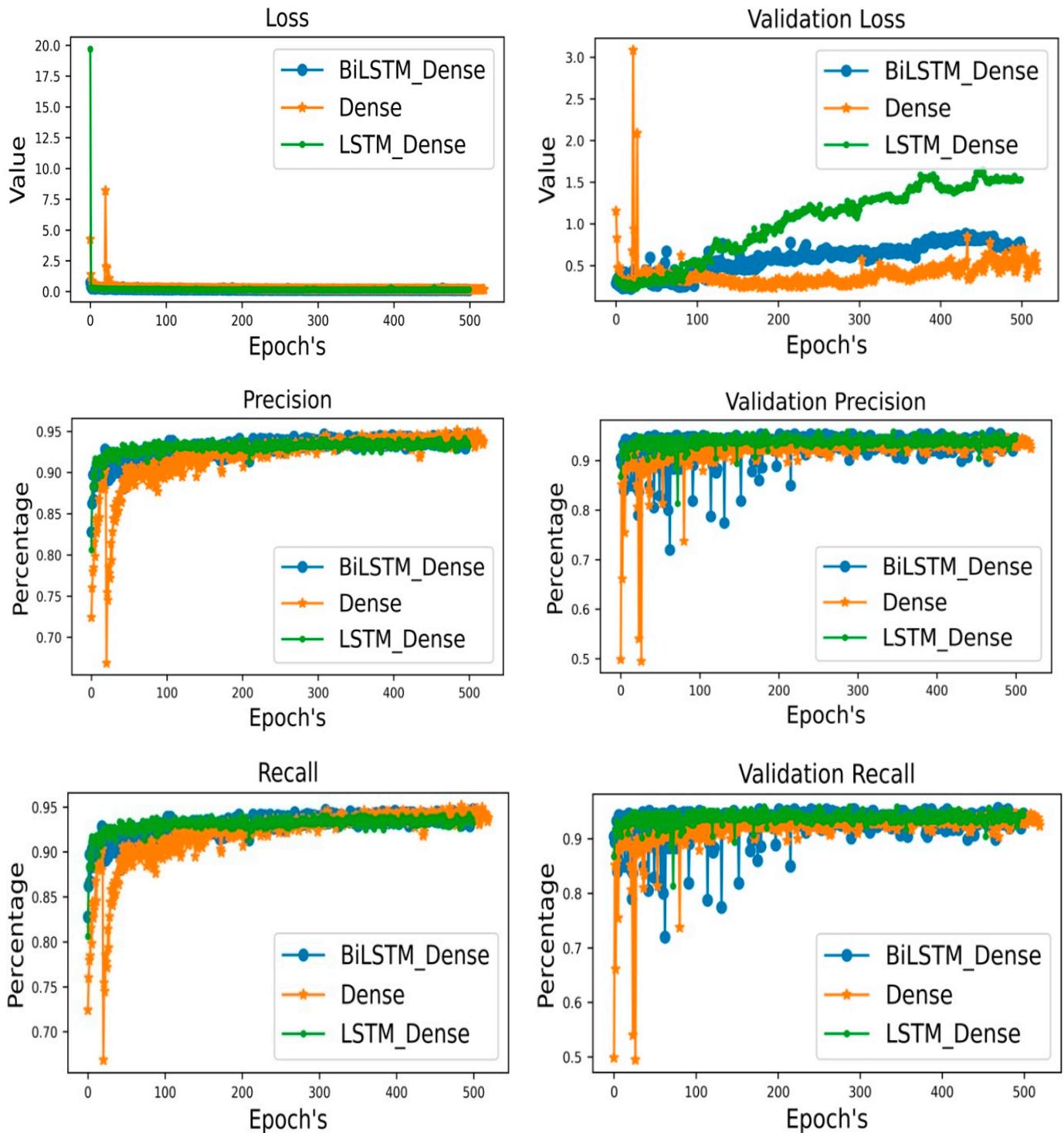
	Accuracy	Val Accuracy	Loss	Val Loss	Precision	Val Precision	Recall	Val Recall
BiLSTM_Dense	94.71	95.58	0.13	0.22	94.71	95.58	94.71	95.58
Dense	95.15	94.46	0.14	0.22	95.15	94.46	95.15	94.46
LSTM_Dense	94.19	95.84	0.14	0.21	94.19	95.84	94.19	95.84

The comparison of training accuracy for each of the seven designs is displayed in the training accuracy graph. The upward movement of the lines of the DensNet201 and Xception models in this example demonstrated that the models were picking up new information from the training data relatively quickly. CNN is indicated on the blue line as having the lowest training accuracy (83.00%). Every model, however, began at a low value and ended up at a greater one. Lower loss functions were achieved as training time on the models increased. The comparison of seven architectures for test accuracy is shown in the second figure. Every model’s graph is moving upward in each epoch, which shows that our proposed model did a great job of identifying fire leaks. These numbers show that the models were correctly trained on the fire dataset. The models accurately identified fire from the test dataset without being either over-fit or under-fit. These are the explanations for why the models’ results in the classification report and confusion matrix were so great.

The comprehensive evaluation of the models reveals positive trends across key metrics. Notably, the training accuracy, exemplified by the upward trajectory of lines for BiLSTM-Dense, indicates the models’ swift adaptation to new information from the training data, suggesting efficient learning capabilities. Furthermore, the test accuracy demonstrates consistent upward trends across all models, underscoring their effectiveness in identifying fire leaks without succumbing to overfitting or underfitting issues. The observed decrease in loss functions with prolonged training time signifies enhanced model performance, emphasizing their ability to minimize errors and improve predictive capabilities. Additionally, the high precision and recall values across all models affirm their accuracy in identifying fire incidents, highlighting a robust performance in both precision (minimizing false positives) and recall (minimizing false negatives). These collective findings underscore the models’ effectiveness in learning, adapting, and accurately identifying fire incidents, emphasizing their potential for practical deployment in fire detection scenarios.



**Figure 6.** Cont.



**Figure 6.** Comparison of various methods used for validating sensors data.

The comparative analysis reveals the effectiveness of the selected models for fire leakage detection using numerical data. BiLSTM-Dense emerges as the most accurate and precise method, with a low loss value, demonstrating its superiority in identifying potential fire incidents. The graphical representation in Figure 6 visually highlights the models' training and validation performance, emphasizing their robustness in fire leakage detection based on numerical data.

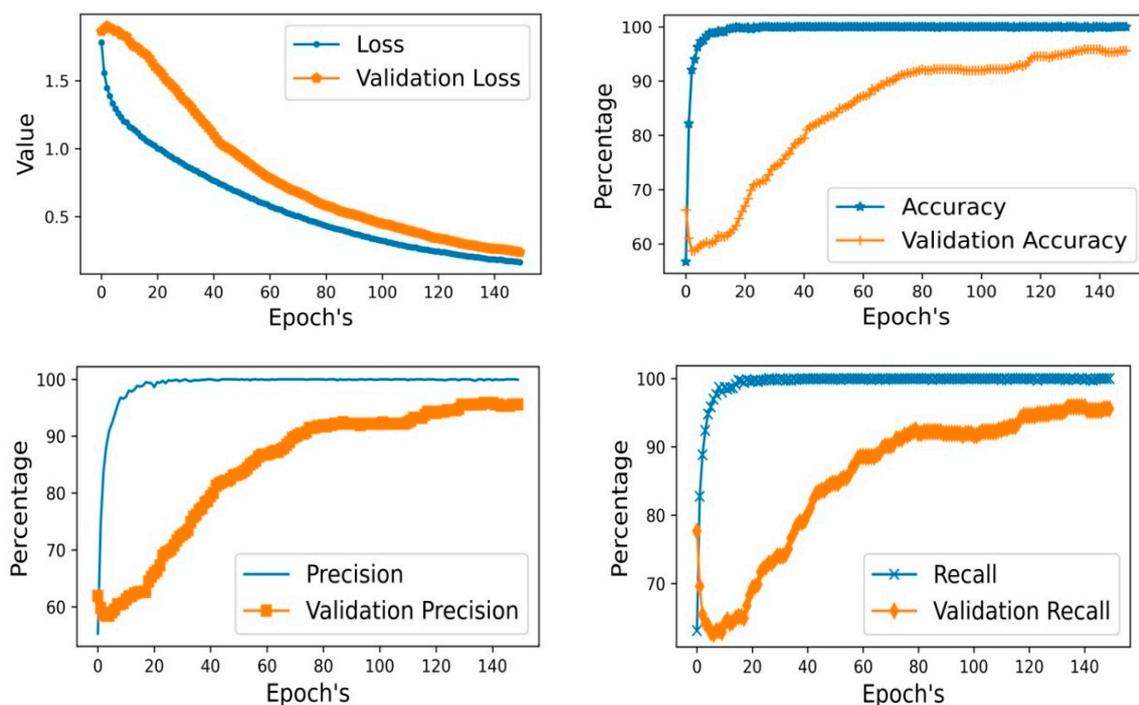
#### 4.2. Multimodal (Image and Sensors Data Analysis)

In this stage of the work, the integration of features from fire sensor measurements and thermal image extraction was undertaken to enable precise decision-making. It was discovered that the use of data from multiple modalities significantly enhances the classifier’s accuracy compared to relying solely on data from a single modality. The combined multimodal classifiers, trained on labeled data from one modality, proved effective when applied to data from another modality, achieving an acceptable accuracy score with the support of multimodal representations. The multimodal model demonstrates exceptional performance, achieving perfect accuracy on the training set (1.00) and a high accuracy of 0.92 on the validation set. The low loss value (0.06) on the training set and a slightly higher value (0.20) on the validation set indicate robust learning without overfitting as shown in Table 7.

**Table 7.** Comparative analysis of multimodal (image and sensors data) data.

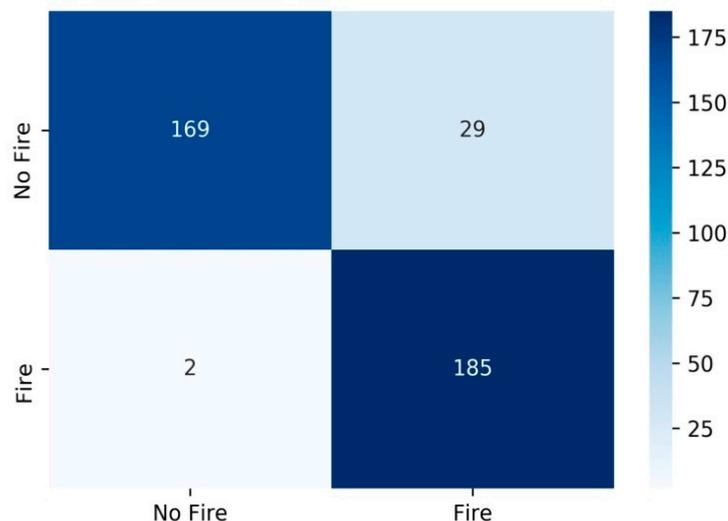
	Accuracy	Val Accuracy	Loss	Val Loss	Precision	Val Precision	Recall	Val Recall
<b>Multimodal</b>	100	92	0.06	0.2	100	92	100	92

In Figure 7, the model performance comparison reveals compelling results across various metrics. Notably, the multimodal model achieves perfect accuracy on the training set, signifying its adept learning from the amalgamated features of fire sensor measurements and thermal images. This proficiency extends to the validation set, where the model maintains high accuracy at 0.92, demonstrating its robust ability to generalize well to unseen data. The low loss on the training set suggests effective convergence, while a slightly higher loss on the validation set indicates good generalization without succumbing to overfitting. Additionally, the model attains perfect precision and recall on both training and validation sets, underscoring its capability to accurately identify fire incidents. These findings collectively highlight the comprehensive and proficient performance of the multimodal model in fire detection scenarios, showcasing its potential for practical application and deployment.



**Figure 7.** Comparison of various parameters used for validating multimodal data.

The confusion matrix provides a visual representation of the model's performance, showcasing its ability to correctly classify instances of fire and non-fire incidents as shown in Figure 8. The high values on the diagonal of the confusion matrix indicate accurate predictions, while off-diagonal values highlight instances of misclassification.



**Figure 8.** Confusion matrix for validating multimodal data.

The multimodal model, combining information from fire sensors and thermal images, emerges as a powerful approach for fire leakage detection. The model demonstrates high accuracy, precision, and recall on both training and validation sets, indicating its effectiveness in making accurate predictions based on the combined features from different modalities. The confusion matrix further confirms the model's ability to correctly classify fire incidents, contributing to its robustness in real-world applications.

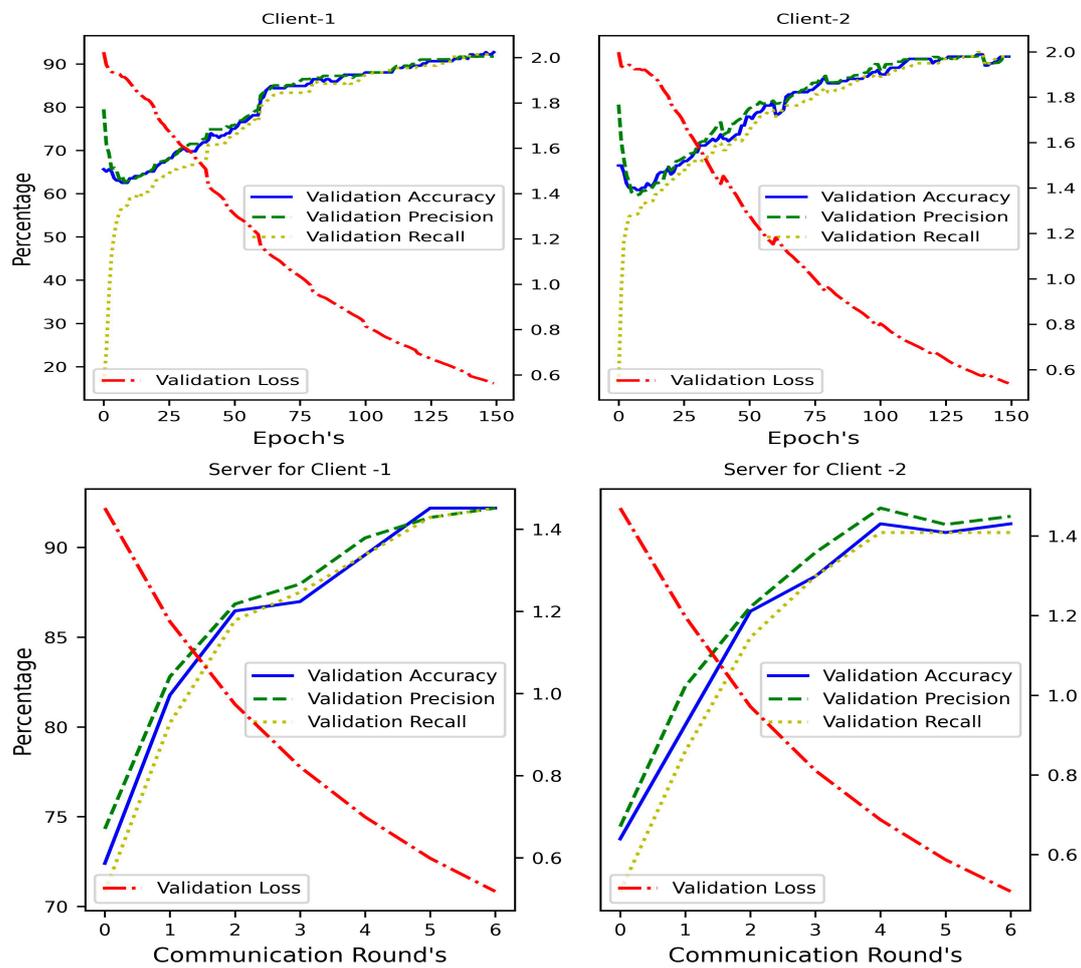
#### 4.3. Analysis of Multimodal Data on Federated Learning Ecosystem

Federated learning, a machine learning technique, facilitates the training of a model across decentralized devices (clients) while preserving data on those devices rather than transferring it to a centralized server. This approach ensures data privacy by allowing clients to locally train a model using their own data. The central server then combines the model changes from each client, incorporating collective intelligence without exposing individual data. Federated learning proves advantageous for fire leak detection, enhancing model accuracy while maintaining the privacy of sensitive information.

In the federated learning framework, the server engages in a maximum of six communication rounds with participating clients, strategically selecting 10% of them for local training in each round. During local training, clients execute 100 epochs with a learning rate of either 0.01 or 0.001, contingent on their individual performance. Notably, the client's local data size aligns with the size of the server's labeled dataset, with the local data being randomly drawn from the overall training dataset. This approach ensures that clients contribute meaningfully to the model's training process while maintaining consistency with the server's labeled data, promoting effective collaboration in the federated learning environment.

In Figure 9, the federated multimodal aggregated results for IID data are presented, illustrating the loss and validation loss curves for both the client side and the server side. Notably, the server's validation results showcase an impressive aggregated accuracy of 99.7% and an exceptionally low validation loss, indicative of superior performance in the context of Independent and Identically Distributed (IID) data. Moving to Figure 10, which focuses on federated multimodal aggregated results for non-IID data, the graphs display loss and validation loss curves for both the client side and the server side. The server's validation loss results emphasize the effectiveness and security of the proposed

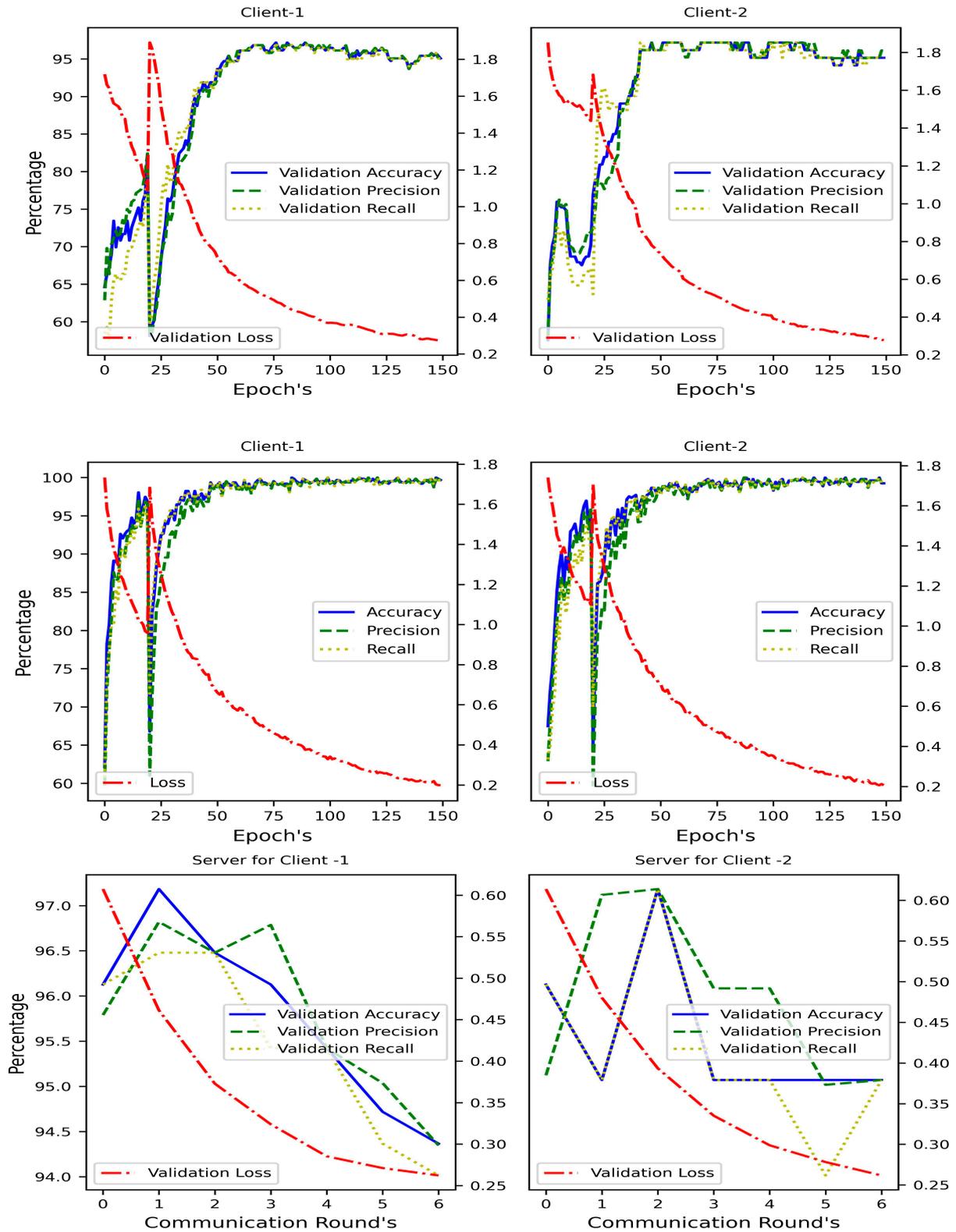
federated multimodal system, surpassing conventional frameworks while maintaining cost-efficiency. These figures collectively underscore the robustness and efficiency of the federated multimodal approach in handling both IID and non-IID data scenarios.



**Figure 9.** Federated multimodal results for IID multimodal data.

In the realm of individual data analysis, specifically focusing on image and sensor data, our study unveils noteworthy insights. For image data, a suite of CNN models, including DenseNet201, InceptionResNet, MobileNetV2, VGG16, VGG19, and Xception, showcased remarkable performance. Precision values consistently surpassed 0.99, affirming precise predictions for scenarios devoid of fire leakage. Notably, the DenseNet201, VGG19, and Xception models outshone their counterparts, exhibiting superior outcomes in terms of loss values. The training accuracy comparison depicted accelerated learning in the DensNet201 and Xception models, while all models exhibited refinement over epochs, with CNN initiating at the lowest training accuracy of 83.00%. Equally compelling results emerged from sensor data analysis, where the BiLSTM-Dense, Dense, and LSTM-DenseDenseNet201 models excelled, outperforming BiLSTM-Dense with precision values exceeding 93.39.

Training and test accuracy comparisons highlighted efficient learning in the DensNet201 and Xception models, showcasing their efficacy in fire leak detection. Furthermore, multimodal data integration proved pivotal, enhancing classifier accuracy through the amalgamation of features from fire sensor measurements and thermal images. This multimodal approach exhibited superior results without additional resource requirements. Finally, the adoption of a federated learning approach underscored its significance, ensuring model training across decentralized devices while safeguarding data privacy.



**Figure 10.** Federated multimodal results for non IID multimodal data.

The aggregated results showcased exceptional validation accuracy (99.7%) and minimal validation loss for both IID and non-IID data, affirming the effectiveness and security of the federated multimodal system. This research underscores the effectiveness of leveraging multimodal data and federated learning for the detection of fire leakage. The integration of convolutional neural networks (CNN) with sensor data yields promising results, and the in-

corporation of multimodal information further enhances the overall accuracy of the system. Crucially, the implementation of federated learning ensures collective intelligence among devices without compromising data privacy, rendering the system robust and secure. In comparison to traditional frameworks, the proposed approach outperforms, providing a cost-efficient and privacy-preserving solution for real-world fire detection scenarios. These findings collectively highlight the potential of multimodal data integration and federated learning in advancing the capabilities of fire detection systems for enhanced accuracy, privacy, and efficiency.

Additional empirical evidence is required to validate the effectiveness of the suggested fire leakage detection system. This can be achieved through testing in various urban contexts and potential pilot deployments. Testing the system in several metropolitan settings with unique infrastructure and environmental features will provide a thorough assessment of its adaptability and applicability. Studying various building structures, fire event patterns, and environmental factors in real-world situations will offer useful insights into how the system performs in different contexts.

Furthermore, evaluating a pilot implementation in a controlled yet operational setting will enable a thorough analysis of the system's feasibility, dependability, and capacity for growth. This may need working with pertinent authorities, emergency services, or industrial sites to establish the system and collect live data on its effectiveness. Consistent monitoring and incremental enhancements guided by input from these deployments will help enhance the system and guarantee its optimal performance in real-world, changing environments. Conducting thorough testing in different urban settings and implementing a pilot program would strengthen the reliability of the data, confirming the effectiveness of the system and its viability for practical use.

## 5. Discussion

The research presented showcases the efficacy of different advanced machine learning models in detecting fire leaks by utilizing multimodal data, which includes measurements from fire sensors and thermal images. The thorough assessment of each of the models for imaging and sensor data uncovers significant observations. When it comes to analyzing picture data, CNN models regularly demonstrate exceptional performance, with precision values consistently exceeding 0.99. The DenseNet201, VGG19, and Xception models demonstrate exceptional loss levels, highlighting their effectiveness in accurately detecting fires.

Regarding sensor data, the Dense, BiLSTM-Dense, and LSTM-DenseDenseNet201 algorithms outperform the BiLSTM-Dense model in terms of precision values that are higher than 93.39. The comparison of training and test accuracy demonstrates the efficient learning capabilities of the DensNet201 and Xception models, confirming their efficiency in detecting fire leaks. The integration of multimodal data is crucial since it improves the accuracy of the classifier by combining information obtained from fire sensor readings and thermal images. This multimodal strategy demonstrates improved outcomes without any additional resource demands. The implementation of federated learning highlights its importance by enabling model training on distributed devices while protecting data privacy.

The combined outcomes demonstrate outstanding validation accuracy (99.7%) and negligible loss of validation for both IID and non-IID data, confirming the efficiency and safety of the federated multimodal approach. This study highlights the efficacy of utilizing many modes of information and federated learning for the identification of fire leakage. To summarize, the combination of CNN with sensor data shows encouraging outcomes, and the inclusion of multimodal information improves the overall precision of the system. Federated learning implementation guarantees the sharing of intelligence among devices while maintaining data privacy, resulting in a robust and secure system. The proposed approach offers a cost-effective and privacy-enhancing alternative for real-world detection of fire scenarios, in contrast to conventional frameworks. The combined results emphasize

the possibility of integrating multimodal data and utilizing federated learning to improve the precision, privacy, and efficiency of fire detection systems.

Within the domain of future work, multiple potential areas for study and development might greatly improve the suggested fire detection system. Exploring the system's ability to withstand various environmental conditions, its capacity to operate effectively in large-scale urban environments, and its incorporation of advanced computing technologies are important areas to investigate. Enhancing accuracy and early detection can be achieved by optimizing the fusion of several sensors, investigating human-in-the-loop methods, and evaluating novel sensor technologies. Furthermore, the inclusion of cybersecurity factors and comprehensive assessments will guarantee the system's robustness and ability to withstand challenges over an extended period. Future studies can aid in the ongoing development and advancement of fire detection technologies by focusing on these factors. This will help overcome current limits and stay ahead of emerging issues in this quickly progressing sector.

## 6. Conclusions

This study stands as a testament to the remarkable progress in fire detection technology, harnessing advancements in sensor technology, microelectronics, and information technologies witnessed over the past decade. Our research specifically delved into the evaluation of intelligent multimodal data for fire leakage detection and identification, culminating in a comprehensive performance and results discussion. By meticulously comparing outcomes derived from disparate data modalities—fire sensor measurements and infrared thermal imaging—we employed a spectrum of deep learning models, including LSTM, BiLSTM, CNN, DenseNet, and VGG16. The performance discussion revealed a significant boost in classifier accuracy through the fusion of these distinct datasets into multimodal data. This underscores the efficacy of harnessing diverse data sources, signifying a pivotal advancement in fire detection systems.

This study proposed a method for assessing the validity of intelligent multimodal data for fire leakage detection and identification. We compared results using the separate data modalities of fire sensor measurements and IR thermal imaging for fire detection and identification. While temperature data is trained on the LSTM, BiLSTM, and CNN methods for fire detection, fire sensor data is trained on CNN, DenseNet, and VG16. The two datasets are then combined to produce multimodal data. According to the results, using data from many modalities increased the classifier's accuracy over just using data from one modality. FL offers a workable option for extracting useful information from the collected data while yet retaining its privacy and localization given the scattered nature of fire monitoring systems with sensors gathering data at multiple geographic locations. The findings reveal the way intelligent multimodal integration of information improves fire detection systems. Federated learning can achieve high accuracy, data privacy, and geographical spread in fire monitoring systems, according to our research. This research sets the groundwork for enhanced, privacy-preserving fire detection technologies, making the future safer and more advanced.

**Author Contributions:** Conceptualization, V.K. and I.K.; methodology, V.K. and R.P.; software, V.K. and R.K.; validation, R.K., I.K. and J.V.; formal analysis, I.K. and J.V.; investigation, R.K. and S.K.; resources, A.S., S.K. and V.K.; data curation, V.K. and I.K.; writing—original draft preparation, I.K. and J.V.; writing—review and editing, R.P., R.K. and V.K.; supervision, V.K.; project administration, I.K.; funding acquisition, R.K. and A.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

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

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Open data used in this paper are available at <https://data.mendeley.com/datasets/f3mjnbm9b3/1> accessed on 10 January 2024, <https://www.kaggle.com/datasets/phylake1337/fire-dataset> accessed on 10 January 2024, <https://www.kaggle.com/datasets/deepcontractor/smoke-detection-dataset> accessed on 10 January 2024.

**Acknowledgments:** We are thankful to all contributors.

**Conflicts of Interest:** The authors declare no conflicts.

## References

- Jain, A.; Srivastava, A. Privacy-preserving efficient fire detection system for indoor surveillance. *IEEE Trans. Ind. Inform.* **2021**, *18*, 3043–3054. [CrossRef]
- Foggia, P.; Saggese, A.; Vento, M. Real-time fire detection for video-surveillance applications using a combination of experts based on color, shape, and motion. *IEEE Trans. Circuits Syst. Video Technol.* **2015**, *25*, 1545–1556. [CrossRef]
- Mothukuri, V.; Parizi, R.M.; Pouriyeh, S.; Huang, Y.; Dehghantanha, A.; Srivastava, G. A survey on security and privacy of federated learning. *Future Gener. Comput. Syst.* **2021**, *115*, 619–640. [CrossRef]
- KhoKhar, F.A.; Shah, J.H.; Khan, M.A.; Sharif, M.; Tariq, U.; Kadry, S. A review on federated learning towards image processing. *Comput. Electr. Eng.* **2022**, *99*, 107818. [CrossRef]
- Caldas, S.; Konečný, J.; McMahan, H.B.; Talwalkar, A. Expanding the reach of federated learning by reducing client resource requirements. *arXiv* **2018**, arXiv:1812.07210.
- Fleming, J.M. Photoelectric and Ionization Detectors—A Review of The Literature Re-Visited. Retrieved Dec. **2004**, *31*, 2010.
- Keller, A.; Rüegg, M.; Forster, M.; Loepfe, M.; Pleisch, R.; Nebiker, P.; Burtscher, H. Open photoacoustic sensor as smoke detector. *Sens. Actuators B Chem.* **2005**, *104*, 1–7. [CrossRef]
- Yar, H.; Ullah, W.; Khan, Z.A.; Baik, S.W. An Effective Attention-based CNN Model for Fire Detection in Adverse Weather Conditions. *ISPRS J. Photogramm. Remote Sens.* **2023**, *206*, 335–346. [CrossRef]
- Dilshad, N.; Khan, T.; Song, J. Efficient deep learning framework for fire detection in complex surveillance environment. *Comput. Syst. Sci. Eng.* **2023**, *46*, 749–764. [CrossRef]
- Yar, H.; Khan, Z.A.; Ullah FU, M.; Ullah, W.; Baik, S.W. A modified YOLOv5 architecture for efficient fire detection in smart cities. *Expert Syst. Appl.* **2023**, *231*, 120465. [CrossRef]
- Dilshad, N.; Khan, S.U.; Alghamdi, N.S.; Taleb, T.; Song, J. Towards Efficient Fire Detection in IoT Environment: A Modified Attention Network and Large-Scale Dataset. *IEEE Internet Things J.* **2023**. [CrossRef]
- Yar, H.; Hussain, T.; Agarwal, M.; Khan, Z.A.; Gupta, S.K.; Baik, S.W. Optimized dual fire attention network and medium-scale fire classification benchmark. *IEEE Trans. Image Process.* **2022**, *31*, 6331–6343. [CrossRef] [PubMed]
- Nadeem, M.; Dilshad, N.; Alghamdi, N.S.; Dang, L.M.; Song, H.K.; Nam, J.; Moon, H. Visual Intelligence in Smart Cities: A Lightweight Deep Learning Model for Fire Detection in an IoT Environment. *Smart Cities* **2023**, *6*, 2245–2259. [CrossRef]
- Hu, Y.; Fu, X.; Zeng, W. Distributed Fire Detection and Localization Model Using Federated Learning. *Mathematics* **2023**, *11*, 1647. [CrossRef]
- Wang, M.; Jiang, L.; Yue, P.; Yu, D.; Tuo, T. FASDD: An Open-access 100,000-level Flame and Smoke Detection Dataset for Deep Learning in Fire Detection. *Earth Syst. Sci. Data Discuss.* **2023**, 1–26. [CrossRef]
- Tamilselvi, M.; Ramkumar, G.; Prabu, R.T.; Anitha, G.; Mohanavel, V. A Real-time Fire recognition technique using a Improved Convolutional Neural Network Method. In Proceedings of the 2023 Eighth International Conference on Science Technology Engineering and Mathematics (ICONSTEM), Chennai, India, 6–7 April 2023; pp. 1–8.
- Bhmra, J.K.; Anantha Ramaprasad, S.; Baldota, S.; Luna, S.; Zen, E.; Ramachandra, R.; Kim, H.; Baldota, C.; Arends, C.; Zen, E.; et al. Multimodal Wildland Fire Smoke Detection. *Remote Sens.* **2023**, *15*, 2790. [CrossRef]
- Nakıp, M.; Güzelış, C. Development of a multi-sensor fire detector based on machine learning models. In Proceedings of the 2019 Innovations in Intelligent Systems and Applications Conference (ASYU), Izmir, Turkey, 31 October–2 November 2019; pp. 1–6.
- Majid, S.; Alenezi, F.; Masood, S.; Ahmad, M.; Gunduz, E.S.; Polat, K. Attention-based CNN model for fire detection and localization in real-world images. *Expert Syst. Appl.* **2022**, *189*, 116114. [CrossRef]
- Yang, Z.; Bu, L.; Wang, T.; Yuan, P.; Jineng, O. Indoor video flame detection based on lightweight convolutional neural network. *Pattern Recognit. Image Anal.* **2020**, *30*, 551–564. [CrossRef]
- Li, Y.; Su, Y.; Zeng, X.; Wang, J. Research on multi-sensor fusion indoor fire perception algorithm based on improved TCN. *Sensors* **2022**, *22*, 4550. [CrossRef]
- Chen, S.; Ren, J.; Yan, Y.; Sun, M.; Hu, F.; Zhao, H. Multi-sourced sensing and support vector machine classification for effective detection of fire hazard in early stage. *Comput. Electr. Eng.* **2022**, *101*, 108046. [CrossRef]
- Hussain, T.; Dai, H.; Gueaieb, W.; Sicklinger, M.; De Masi, G. UAV-based Multi-scale Features Fusion Attention for Fire Detection in Smart City Ecosystems. In Proceedings of the 2022 IEEE International Smart Cities Conference (ISC2), Pafos, Cyprus, 26–29 September 2022; pp. 1–4.
- Tao, J.; Gao, Z.; Guo, Z. Training Vision Transformers in Federated Learning with Limited Edge-Device Resources. *Electronics* **2022**, *11*, 2638. [CrossRef]

25. Sridhar, P.; Thangavel, S.K.; Parameswaran, L.; Oruganti VR, M. Fire Sensor and Surveillance Camera-Based GTCNN for Fire Detection System. *IEEE Sens. J.* **2023**, *23*, 7626–7633. [[CrossRef](#)]
26. McMahan, B.; Moore, E.; Ramage, D.; Hampson, S.; Arcas, B.A. Communication-efficient learning of deep networks from decentralized data. In *Artificial Intelligence and Statistics; Proceedings for Machine Learning Research*; MD, USA, 2017; pp. 1273–1282. Available online: <https://proceedings.mlr.press/v54/mcmahan17a?ref=https://githubhelp.com> (accessed on 15 December 2023).
27. Govil, K.; Welch, M.L.; Ball, J.T.; Pennypacker, C.R. Preliminary results from a wildfire detection system using deep learning on remote camera images. *Remote Sens.* **2020**, *12*, 166. [[CrossRef](#)]
28. Cao, Y.; Yang, F.; Tang, Q.; Lu, X. An attention-enhanced bidirectional LSTM for early forest fire smoke recognition. *IEEE Access* **2019**, *7*, 154732–154742. [[CrossRef](#)]
29. Shi, N.; Lai, F.; Kontar, R.A.; Chowdhury, M. Fed-ensemble: Improving generalization through model ensembling in federated learning. *arXiv* **2021**, arXiv:2107.10663.
30. Sousa, M.J.; Moutinho, A.; Almeida, M. Wildfire detection using transfer learning on augmented datasets. *Expert Syst. Appl.* **2020**, *142*, 112975. [[CrossRef](#)]
31. Wang, L.; Wang, W.; Li, B. CMFL: Mitigating communication overhead for federated learning. In Proceedings of the 2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS), Dallas, TX, USA, 7–10 July 2019; pp. 954–964.
32. Available online: <https://www.kaggle.com/datasets/phylake1337/fire-dataset> (accessed on 10 January 2024).
33. Available online: <https://www.kaggle.com/datasets/deepcontractor/smoke-detection-dataset/discussion> (accessed on 10 January 2024).
34. Available online: <https://data.mendeley.com/datasets/f3mjnbm9b3/1> (accessed on 10 January 2024).
35. Havens, K.J.; Sharp, E.J. *Thermal Imaging Techniques to Survey and Monitor Animals in the Wild: A Methodology*; Academic Press: Cambridge, MA, USA, 2015.
36. Konečný, J.; McMahan, H.B.; Yu, F.X.; Richtárik, P.; Suresh, A.T.; Bacon, D. Federated learning: Strategies for improving communication efficiency. *arXiv* **2016**, arXiv:1610.05492.
37. Bonawitz, K.; Eichner, H.; Grieskamp, W.; Huba, D.; Ingerman, A.; Ivanov, V.; Kiddon, C.; Konečný, J.; Mazzocchi, S.; McMahan, B.; et al. Towards federated learning at scale: System design. *Proc. Mach. Learn. Syst.* **2019**, *1*, 374–388.
38. Yang, Q.; Liu, Y.; Chen, T.; Tong, Y. Federated machine learning: Concept and applications. *ACM Trans. Intell. Syst. Technol. (TIST)* **2019**, *10*, 1–19. [[CrossRef](#)]
39. Zhao, Y.; Li, M.; Lai, L.; Suda, N.; Civin, D.; Chandra, V. Federated learning with non-iid data. *arXiv* **2018**, arXiv:1806.00582. [[CrossRef](#)]
40. Kukreja, V.; Kumar, D.; Kaur, A. GAN-based synthetic data augmentation for increased CNN performance in Vehicle Number Plate Recognition. In Proceedings of the 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 5–7 November 2020; pp. 1190–1195.
41. Dhiman, P.; Kukreja, V.; Manoharan, P.; Kaur, A.; Kamruzzaman, M.M.; Dhaou, I.B.; Iwendi, C. A novel deep learning model for detection of severity level of the disease in citrus fruits. *Electronics* **2022**, *11*, 495. [[CrossRef](#)]

**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.