

## Research Article

# Convolutional Autoencoder-Based Deep Learning Approach for Aerosol Emission Detection Using LiDAR Dataset

**Mazhar Hameed** <sup>1</sup>, **Fengbao Yang** <sup>1</sup>, **Sibghat Ullah Bazai** <sup>2</sup>, **Muhammad Imran Ghafoor** <sup>3</sup>,  
**Ali Alshehri** <sup>4</sup>, **Ilyas Khan** <sup>5</sup>, **Shafi Ullah** <sup>2</sup>, **Mehmood Baryalai** <sup>6</sup>, **Fawwad Hassan Jaskani** <sup>7</sup>,  
**and Mulugeta Andualem** <sup>8</sup>

<sup>1</sup>School of Information and Communication Engineering, North University of China, Taiyuan, China

<sup>2</sup>Department of Computer Engineering, Balochistan University of Information Technology, Engineering, and Management Sciences, Quetta, Pakistan

<sup>3</sup>Department of Electrical Engineering, Superior University, Pakistan

<sup>4</sup>Department of Computer Science Applied College University of Tabuk Saudi Arabia, Saudi Arabia

<sup>5</sup>Department of Mathematics, College of Science Al-Zulfi, Majmaah University, Al-Majmaah 11952, Saudi Arabia

<sup>6</sup>Department of Information Technology, Balochistan University of Information Technology, Engineering, and Management Sciences, Quetta, Pakistan

<sup>7</sup>Department of Computer Systems Engineering, The Islamia University of Bahawalpur, Pakistan

<sup>8</sup>Department of Mathematics, Bonga University, Bonga, Ethiopia

Correspondence should be addressed to Mazhar Hameed; [mazhar\\_hameed@hotmail.com](mailto:mazhar_hameed@hotmail.com) and Mulugeta Andualem; [mulugetaandualem4@gmail.com](mailto:mulugetaandualem4@gmail.com)

Received 3 December 2021; Revised 3 April 2022; Accepted 29 April 2022; Published 1 June 2022

Academic Editor: Zhenxing Zhang

Copyright © 2022 Mazhar Hameed et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Quantifying atmospheric aerosols and their linkages to climatic repercussions is necessary to understand the dynamics of climate forcing and enhance our knowledge of climate change. Because of this reactivity to precipitation, temperature, topography, and human activity, the atmospheric boundary layer (ABL) is one of the most dynamic atmospheric regions: ABL aerosols have a big impact on the evolution of climate change's radiative forcing, human health, food security, and, eventually, the local and global economy. Continuous monitoring and instrumental and computational approaches are required for the detection and analysis of ABL pattern behavior. This paper provides a deep learning-based outer layer aerosol detection system based on Light Detection and Ranging (LiDAR) data fusion. The suggested method applies sequential models to turn low-level data into compressed features using object-based analysis, feature-level fusion, and autoencoder-based dimensionality reduction. Convolutional neural networks (CNNs) were used to convert compressed data into high-level properties that could be used to categorize air particles in the outer layer. This research describes deep learning approaches that allowed for detecting 40% more atmospheric features at a horizontal resolution of 5 km during daytime operations when applied to LiDAR data. Compared to existing deep learning algorithms for edges and complicated near-surface sceneries during the day, a convolutional autoencoder (CAE) trained using LiDAR dataset standard data products showed the potential for improved aerosol discrimination with 98% accuracy.

## 1. Introduction

In the Earth's climate system, air quality and hydrological cycle with an extent that is mainly dependent upon the atmospheric properties height, thickness, and type and air

characteristics such as clouds and aerosols play a vital role [1]. Clouds of liquid water on the surface of Earth tend to reflect inbound sunlight, cooling the surface of Earth. However, ice clouds in the upper troposphere absorb and reradiate heat emitted from the surface, warming up the surface of

the Earth. Aerosol particles include windblown desert dust, wildfire smoke, sulfurous particles from volcanic eruptions, and fossil fuel particulate matter [2]. Aerosols cool or warm the surface, depending on their size, composition, and location in the atmosphere [3].

A CNN is a supervised machine learning method used for the recognition of picture features. Commercial uses of CNNs include a wide variety of object detection and semanticized image segmentation [3]. CNNs have also been used to predict tropical cyclone intensity precisely using satellite imagery and hailstorm detection in radar images with higher accuracy than previous techniques. After instantiation of CNN's layer architecture, the model is trained with truth data sets, which develops expertise to predict proper characteristics in the image [4]. While the training of a CNN can take a long time, the forecasts are rapid compared to older algorithms or a manual approach. A collection of CNNs has been built to forecast the positions of clouds and aerosols in CATS LiDAR data to increase the speed at which LiDAR data may be distributed and establish feasibility to give real-time layer type products [5].

Natural and artificial aerosol emissions can significantly threaten urban and regional air quality, such as biomass burning. As a result, it is crucial to understand the optical, microphysical, and geometrical characteristics of local or targeted aerosol emissions in the boundary layer. The LiDAR sensor, which uses a laser as its source, may offer highly temporal and spatially vertically resolved profiles of aerosols [6]. As a result, LiDAR remote sensing observations will aid in the research and characterization of aerosol emissions from source to destination and improve air quality. This section welcomes submissions on the most recent results and advancements in LiDAR distant detection of optical, microphysical, and mathematical spray properties from airborne-mounted LiDARs, territorial ground-based LiDAR organizations, global satellite missions, across all instrument platforms (Raman, high-spectral resolution, DIAL, and others), fleeting and spatial scales, and from airborne-moon missions. LiDAR control through a remote and the identification of anthropogenic aerosols that have an impact on air quality from industrial, biomass burning, and agricultural sources, as well as campaigns targeted at giving a full assessment of climate and health consequences, are especially encouraged [7]. Man-made aerosol emissions in cities are linked to their impacts on micrometeorology and the radiative budget, i.e., their function in heating/cooling the atmospheric column and promoting/suppressing convection, and are given specific emphasis [8]. We have used convolutional autoencoder models (CAE) [7] to detect aerosols from fusion LiDAR dataset [8]. CAE's ability to extract the aerosol type is influenced by the optical inputs' physical substance and uncertainty, as well as the CAE structure and training technique, notably the size of the data set employed for this purpose. An aerosol model detailing the optical characteristics of distinct particles was created to provide a consistent depiction of the aerosol types. This model can provide a representative and statistically meaningful synthetic database in order to recreate known aerosol features. This synthetic data collection is important since few

observational data sets are statistically relevant, well-characterized, and representative of the whole range of aerosol species. Normalization is a common practice in data preparation for machine learning. You must normalize your data to a standard scale without distorting the range of numbers or surrendering any information if you want it to be consistent [9]. The aerosol model was constructed in order to train the CAE by simulating a large number of LiDAR observations (i.e., a synthetic data set) [10]. The most likely aerosol type inside the detected layers is the output data from generative adversarial neural networks (GANs) [10].

Deep learning techniques consist of deep layers where feature extraction and classification are not separately performed. Deep learning is a subclass of machine learning that processes data and makes patterns for use in decision-making. The deep learning technique teaches the machine to perform intelligent tasks. Deep learning contains numerous techniques such as a CNN, CAE, and GAN model [11]. The CNN works automatically for detecting features and classifying the raw dataset. Deep learning is a more advanced technique for recognizing hidden features more accurately and efficiently [12].

The detection of aerosols with manual detection is a challenging task due to the various properties of the environment. Hence, an automated and accurate system is required for aerosol emission detection [13]. The research is aimed at designing an intelligent aerosol emission detection system using fusion LiDAR data and applying deep and machine learning techniques to predict the emissions' area [14].

The aerosol emissions recognition has been considered as an essential application for numerous security branches and health systems. Several researchers have applied handcrafted techniques to identify such anomalies in the scene [15]. Using handcrafted features from linear binary configuration from three orthogonal planes, Gaussian mixture model, Markov random field, etc. for irregularity recognition is not acceptable as they solely rely on human assumption. Hence, the training data is not explained correctly to learn discriminative features characteristic of aerosols [16]. The acquired data from remote sensing and satellites are the key characteristics and contributions of this research. We used the convolutional autoencoder (CAE) neural network to process the data, which takes photos and extracts the hidden patterns of the input images before reconstructing the features from the hidden pattern. We then established a sequential model in the autoencoder model, which allows us to simply build sequential layers of the network from input to output. Then, we have applied GAN, which helps to solve such tasks as pattern recognition from descriptions, getting high resolution of images from low-resolution ones and predicting which is the aerosol emission area or not [1].

All of this research [17] have revealed a diverse group of people. A wide range of aerosols is challenging to categorize due to many flaws (e.g., many aerosol types have identical optical characters). Other challenge in the categorization of aerosols is the difficulty in linking their optical qualities to their physical properties source [18]. In actuality, atmospheric aerosols are made up of a variety of substances. There are a lot of sources, and data on pure aerosol kinds

TABLE 1: Comparative analysis of previous studies conducted on LiDAR.

Ref	Technique	Source of data	Outcome	Accuracy
Xiu et al. [1]	Semantic segmentation	LiDAR dataset	The entire waveform of LiDAR	89%
Zhang et al. [2]	Computer vision techniques	Land SAT8 data	Retrieval of Forest aboveground biomass	90.3%
Melotti and Premebida [5]	Multimodal deep learning	Self-created dataset	Object recognition combining camera	93%
Zhang et al. [7]	Multisource data fusion	LiDAR dataset	Data fusion	88%
Wahid et al. [6]	Object exploration vision techniques	Self-created dataset	Distributed soft actor critics	86.45%

are scarce. Systematic measurements and intense measurement campaigns employing various aerosol measurement methods have been carried out to address these difficulties.

Many Earth systems, such as temperature, air quality, and hydrology, are affected by the atmospheric properties of clouds and aerosols, and their effects are highly influenced by their height, thickness, and kind. At ground level, liquid water clouds tend to reflect incoming sunlight, which helps chill the surface [18]. Additionally, clouds in the upper troposphere that comprise ice are capable of absorbing heat from the surface and reemitting it, therefore contributing to the rise in surface temperature [19]. Dust, smoke, sulfur, and particles of fossil fuel burning are all examples of aerosol particles. The ability of aerosols to cool or heat the surface depends on their size, composition, and position in the atmosphere [20]. Many types of aerosols, including dark-colored ones, such as black carbon from fossil fuel combustion, are known to absorb radiation. Table 1 shows the previous studies' comparative analysis.

For this reason, the current study presents and contributes as follows:

- (i) The use of LiDAR and orthophotofusion combined with a deep learning (DL) strategy to detect aerosols
- (ii) DL has progressed past multilevel perceptron and now includes the following

This particular study employs an autoencoder framework and a convolutional neural network (CNN) to accomplish feature dimensionality reduction and object classification of aerosols and no aerosols in LiDAR and orthoimage data after segmentation.

## 2. Materials and Methods

An aerosol model was used to investigate the optical characteristics of pure aerosols produced by a single source (e.g., dust produced by the deserts and marine particles produced by the oceans). Continental, continental polluted, dust, marine, smoke, and volcanic are the six forms of pure aerosols addressed in this article. The aerosol model combines the global aerosol dataset with iterative computations of each aerosol type's intensive optical properties, as well as a numerical technique for  $T$ -matrix. The OPAC software application was used to determine the chemical makeup of each pure aerosol type (aerosol and cloud optical proper-

ties). To replicate the vast spectrum of particles in the atmosphere, the chemical composition of each aerosol type was modified within specified boundaries. For sound wavelengths of 350, 550, and 1000 nm, the aerosol model was utilized to create a synthetic database. These wavelengths were selected from OPAC's 61 wavelengths (0.25–40 m) for which GADS possesses microphysical aerosol parameters. After that, the wavelengths are rescaled in angstroms to match the traditional LiDAR wavelengths (i.e., 355, 532, and 1064 nm). This was deemed to be an acceptable assumption for all aerosol types, given the minimal difference between the LiDAR and model wavelengths. The aerosol model can be expanded to cover more wavelengths if necessary.

Every type of pure aerosol is made up of an internal combination of fundamental components in variable mix ratios that do not interact physically or chemically. Water-soluble, insoluble, soot, mineral, sulfate, and sea salt are all collected by OPAC (accumulation, coarse). The microphysical properties of each component are stored in the GADS database. Smoke and continentally contaminated kinds, however, cannot attain values above 1.2 for angstrom (550 to 350 nm) with the present GADS soot refractive index values.

Figure 1 depicts the workflow of the suggested technique combining a convolutional autoencoder, a sequential algorithm, and a GAN. The following are the details of the block diagram.

*2.1. Data Acquisition of UAV LiDAR Datasets.* The input photos are from the remote sensing LiDAR data of the aerosol emission dataset,  $128 \times 128 \times 3$  input form. The convolutional autoencoder uses this image as an input (CAE). To recover the hidden patterns, CAE separates the input image into convolutional and pooling layers. It will then be supplied into the deconvolutional and unspooling layers, which will reconstruct the features of the hidden pattern. We used a sequential approach, which made it simple to build subsequent network layers in the order of input to output. Then, we employed a GAN to tackle problems like picture generation from descriptions or features, converting low-resolution image frames to high-resolution image frames, detecting which emission activity is active or not, and recovering image frames containing a given pattern.

One of the most well-known datasets in the field of aerosol detection is the LiDAR fusion dataset. It contains data from an aerial view, LiDAR, and other sensors attached to

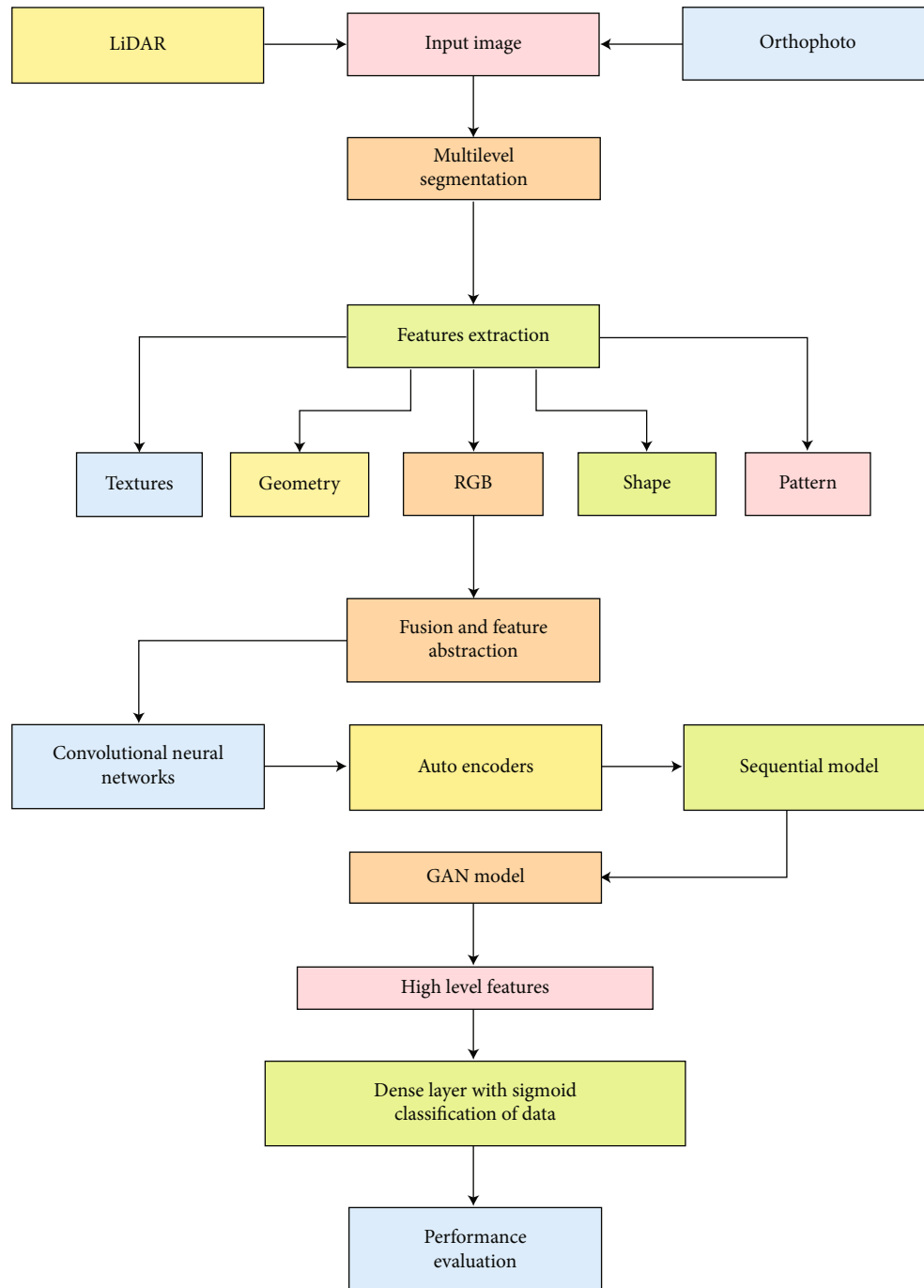


FIGURE 1: Proposed methodology.

the top of a drone that flies through various environments and scenarios.

This collection contains LiDAR frames that have been converted to 2D depth images. These 2D depth images show the same scene as the corresponding LiDAR frame but are more user-friendly.

The 360 LiDAR frames, like those in the dataset, are arranged in a cylinder around the sensor. The 2D depth images in this dataset might be represented as if the cylinder of the LiDAR frame had been split in half and straightened

into a 2D plane. The distance of the reflecting item from the LiDAR sensor is represented by the pixels in these 2D depth photographs. The number of laser beams utilized to scan the surroundings is represented by the vertical resolution of the 2D depth image (64 in our case). These 2D depth images could be utilized for segmentation, detection, recognition, and other tasks, drawing on a large body of computer vision literature on 2D images. We have compared our model with a hybrid model of GAN and autoencoder to compare the performances.

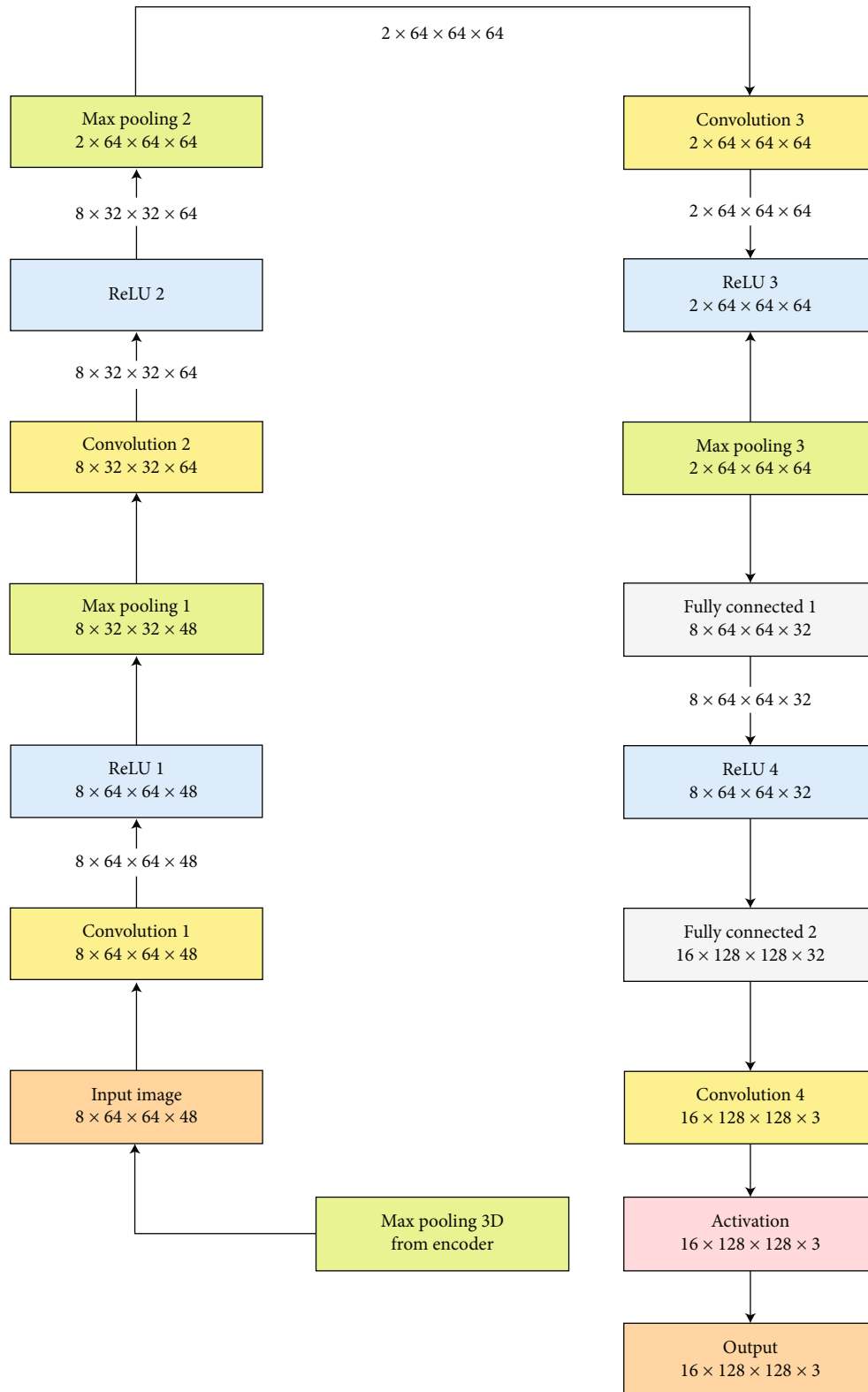


FIGURE 2: Generative adversarial network (GAN).

2.2. *Model Training.* The proposed technique defined the rule when emissions occurred. We trained the model, which contains spatial feature descriptors. The image description explains the visual feature of each frame. Each frame has

its characteristics such as shape, color, and texture. This description provides a feature vector. The convolutional autoencoder model is adequately trained with blocks of pixels that contain only standard segments. The frames'

input and output volume mistakes are reduced. The model is trained correctly on the regular images, and then the model shows the low reconstruction error. Each testing input image produces a reconstruction error. The reconstruction error depends upon custom loss. We set the threshold on the value. If the value crossed a threshold limit, it shows an aerosol emission and represents a regular event below the threshold limit. Thus, the system will be able to recognize the rare events that occur in the images.

**2.3. Model Parameters.** The training model is used for reducing the reconstruction error of the input volume. The proposed Model used an Adam optimizer, and the learning rate automatically depends upon the updated history of the model's weight. The minimum patch size is 64. Every training image size is trained for a maximum of 50 epochs, or until the aerosol layers are lost and the 10 consecutive epochs are reduced. The spatial autoencoder activation goal is chosen to be the hyperbolic curve. Despite its regularization capacity, we did not use the rectified linear unit (ReLU) to guarantee the regularity of the encoding and decoding functions because triggered values from ReLU have no upper bound.

**2.4. Convolutional Autoencoder Model.** An autoencoder is an encoder-decoder system that reconstructs the input as the output. We achieved autoencoder by two subsystems: the encoder converts the input image frame into a feature vector for internal representation [6]. The decoder, on the other hand, translates the internal representation back to the original reconstructed image. Autoencoder provides a reconstruction error [19]. The minimum reconstruction error means a slight difference between the input and the reconstructed image frames [20].

**2.5. Sequential Model.** The sequential model was used, which makes it simple to stack sequential network layers from input to output. Figure 2 shows generative adversarial network (GAN).

**2.6. Generative Adversarial Network Model.** GAN is a generative modelling method based on the CNN method. In machine learning, generative modeling is an unsupervised learning problem [21]. It comprises pitting two neural networks against each other to automatically find and learn regularities or patterns in incoming data. Adversarial competition consists of two parts: generator: replicate authentic data in order to create fictitious data and discriminator: detecting the generator by distinguishing between accurate and fictitious data [12].

As a result, we used GAN to perform tasks such as image generation from descriptions or features, obtaining high resolution image frames from low resolution ones, predicting which emission activity is active and which is not and retrieving image frames containing a given pattern. Figure 3 shows classification using the generative adversarial model.

**2.7. Model Description.** As shown in Figure 1, the UAV fused LiDAR dataset utilized in this investigation was collected

above the Universiti Sains Malaysia campus on February 3, 2018, at midday. A Canon PowerShot SX230 HS (5 mm) camera was used to collect data from a UAV flying at a height of 353 meters (5 mm). The photographs were created using three channels (RGB) with a ground resolution of around 9.95 cm/pixel, a resolution of 4000 3000-pixels, and an 8-bit radiometric resolution. An orthomosaic snapshot of the collected image series was produced with an average root mean square errors (RMSE) of 0.192894 m. (1.08795 pix). The DSM was also created with Agisoft PhotoScan Professional. The chosen subset spans a total area of 1.68 km<sup>2</sup>. The DSM's resolution was 79.6 cm/pixel, while Agisoft's point clouds had a point density of about 1.58 points/m<sup>2</sup>. Figure 1 depicts the operational flow of the suggested technique using a convolutional autoencoder, a sequential algorithm, and a generative adversarial network (GAN). The following are the details of the block diagram:

The input photos come from the UAV aerosol real-world aerosol collection and have a 128x128x3 input shape. The convolutional aAutoencoder uses this image as an input (CAE). CAE extracts latent patterns from input pictures using convolutional and pooling layers (128x128x3). It will then be fed into the deconvolutional and max-pooling layers, which will recreate the hidden pattern's characteristics. We used the sequential model, which allows us to stack sequential network layers from input to output effortlessly. Then, we used GAN to help with tasks like picture generation from descriptions or features, obtaining high resolution image frames from low-resolution ones, predicting whether aberrant activity is there or not, and retrieving image frames containing a given pattern.

Feature descriptors output feature descriptors/feature vectors from an input image frame. Feature descriptors are a set of integers that encode useful information. To validate the results, the UAV data was divided into two categories: testing (20%) and training (80%). Convolutional autoencoder and GAN model are two deep learning algorithms. The purpose of each model is to generate reconstructed images in a hybrid way by using an output layer from previous models. The sequential model has been used in CAE for sequencing the stack layers.

**2.8. Raw Image Data Processing.** The UAV aerosol dataset is used for testing and evaluation of the offered method. The aerosol dataset contains 13 different real-world anomalies. The real-world anomalies are abused, arrest, assault, and explosion, etc. We know that images are combinations of frames; so, we have converted the images into frames for preprocessing and feature extraction. The converted image frames in the form of JPEG and applied image resizing are as follows: the image resizing is essential because the dimension of each image's frame is not the same. The resized images are fed into the temporal volumes.

**2.9. Model Training.** The proposed technique defined the rule when abnormal events occurred. The maximum regular frames are different as compared to the abnormally frames. We trained the model, which contains spatial feature descriptors. The image description explains the visual feature

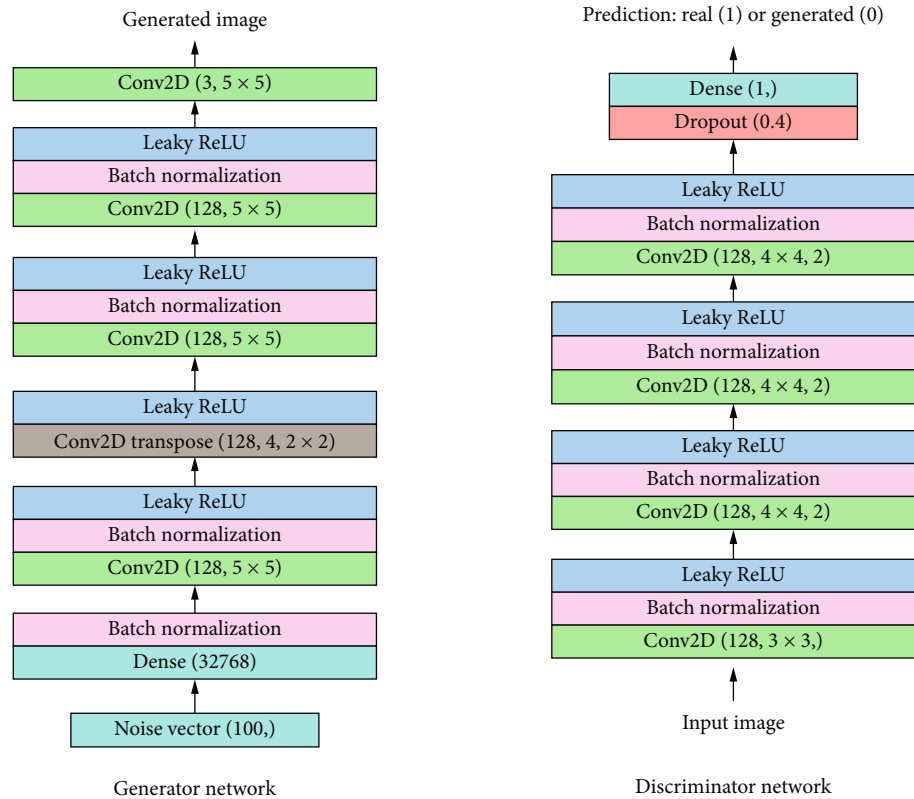


FIGURE 3: Classification using the generative adversarial model.

of each image's frame. Each frame has its characteristics such as shape, color, texture, and motion. This description provides a feature vector. The convolutional autoencoder model is adequately trained with images blocks contains only regular segments. The error between the input and output volume of the frames is reduced. The model is trained correctly on the standard image's frames, and then the model shows the low reconstruction error. Each testing input images volume produces a reconstruction error. The reconstruction error depends upon custom loss. We set the threshold on the value. If the value crossed a threshold limit, it shows an abnormal event, and below the threshold limit, it represents a typical event. Thus, the system will be able to recognize the rare events that occur in the images. In the following Table 2, we have shown the features that are extracted in the model.

One can find a wide variety of information about surface elements like topography, texture, and shape by studying them on images and LiDAR surveys. Using a lot of different characteristics might lead to overfitting, and that it is especially true when the training set is quite small. The other downsides of using several characteristics are that they increase the level of noise, the volume of redundant information, and the time it takes to compute. To deal with the problem of high-dimensional feature space, an autoencoder-based technique is proposed that reduces feature space dimensionality and improves low-level features by translating them into fewer features (i.e., reduced low-level features). The redesigned features will most likely be more informative than the initial raw features, assisting the full detection model creation pro-

cess. CNN models were built to identify key architectural attributes, which were then processed using a series of convolution and pooling procedures to convert low-level characteristics into high-level ones. This section discusses the process after using autoencoders and CNN models to abstract low-level properties.

**2.10. Model Parameters.** The training model is used for reducing the reconstruction error of the input volume. The proposed model used an Adam optimizer; the learning rate automatically depends upon the updated history of the model's weight. The minimum patch size is 64. Depending on the aerosol layers, each training image size is trained for a maximum of 50 epochs. Following the loss of authentication data, the 10 consecutive epochs are no longer reduced. The spatial autoencoder activation goal is chosen to be the hyperbolic curve. Despite its regularization capacity, we did not use the rectified linear unit (ReLU) to guarantee the regularity of the encoding and decoding functions because triggered values from ReLU have no upper bound. An autoencoder is an encoder-decoder system that reconstructs the input as the output. We achieved autoencoder by two subsystems: the encoder converts the input image frame into a feature vector for internal representation.

The decoder, on the other hand, uses the internal representation to translate it back to the reconstructed images. Autoencoder provides a reconstruction error. The minimum reconstruction error means a slight difference between the input images frame and the reconstructed image frame.

TABLE 2: Features extraction from dataset.

Data	Feature	GLCM	Description
Orthophoto	Spectral	Angular Contrast Correlation	$\sum_{i,j=0}^{N-1} P2ij$ $\sum_{i,j=0}^{N-1} Pij(i-j)$ $\sum_{i,j=0}^{N-1} Pij((i-u)-(j-u))$
Orthophoto	Texture	Entropy Homogeneity Mean	$\sum_{i,j=0}^{N-1} P \ln (ij)$ $\sum_{i,j=0}^{N-1} P \frac{j}{1+(Pij)}$ $\varrho_i \sum_{i,j=0}^{N-1} f(i)Pij(i) = \varrho_j \sum_{i,j=0}^{N-1} f(j)Pij(j)$
LiDAR	Shape	Area Compactness Density	Area of segments Compactness of polygon Density of holes
LiDAR	LiDAR	DEM DSM nDSM	Digital elevation model Digital surface model Object high (DEM-DSM)

2.11. *Sequential Model.* The sequential model was employed in Figure 4, which allows us to easily stack sequential network layers from input to output.

2.12. *Generative Adversarial Network (GAN).* GAN is a generative modeling method based on the CNN method. In machine learning, generative modeling is an unsupervised learning problem. It comprises pitting two neural networks against each other to find and learn regularities or patterns in incoming data. As a result, we used a generative adversarial network (GAN) to solve tasks like image generation from descriptions or features, obtaining high resolution image frames from low-resolution ones, predicting whether abnormal activity is abnormal or not, and retrieving image frames that contain a given pattern.

### 3. Results and Discussion

The neural complexity addresses the lesser limits of neural resources (neuronal counts) a network needs to do a specific task within a certain tolerance. Lower limits on the information required for the intended input-output function are measured by the complexity of the information (i.e., number of examples). This study suggests a superresolution convolutional neural network (CNN) with a minimal level of complexity (SR). The computational complexity of the suggested strategy is 71.37 percent lower in CPU, TPU, and GPU than the very-deep SR (VDSR) technique, with a peak signal-to-noise ratio loss of 0.49 dB.

Autoencoder, a generative model, was used in the proposed model. Image samples are used to train the autoencoder, and testing images are used to predict the aerosol. An encoder and a decoder make up the autoencoder. For the reconstructed pictures, the trained model's loss function is

calculated. At the feature extraction stage, as shown in Figure 5, A total of 21 features, including spectral, form, textural, and LiDAR-based attributes, were retrieved to recognize aerosol layers objects in LiDAR and orthophoto data. Spectral features were used to evaluate the mean pixel values in the orthophoto bands. Shape attributes are the geometric information of meaningful things that is determined from the pixels that make up these objects. To make sure that these features are used effectively, the map must be segmented accurately. Haralick texture characteristics were also used to construct texture features based on the grey-level co-occurrence matrix (GLCM) or the grey-level difference vector. Alternatively, the topography and height of objects were described using LiDAR-based characteristics. The identification and description of aerosol layers are critical elements in the reconstruction of aerosol layers objects. The preceding alluded to a method for distinguishing aerosol layers items among various objects [20]. The last, on the other hand, is concerned with defining the mathematical limit of aerosol layer objects so that their computation and concentration information can be displayed as attributes connected with the objects in a geographic information framework (GIS). From one perspective, orthophoto has a critical spatial objective limit and exhibits solid reflectance around the limits of aerosol layers. In any event, the uncanny similarity of distinct ground objects complicates orthophoto extraction of aerosol layers. However, because of the relatively tiny footprint size of the laser bar and unfavorable backscattering from lighted targets, collecting aerosol layers edges with tallness discontinuities is difficult in LiDAR [20]. The use of orthophoto and LiDAR together can increase the precision of aerosol layer detection and description measurements in this way.

Information combination is the process of using or combining data from multiple sources to frame a new dataset



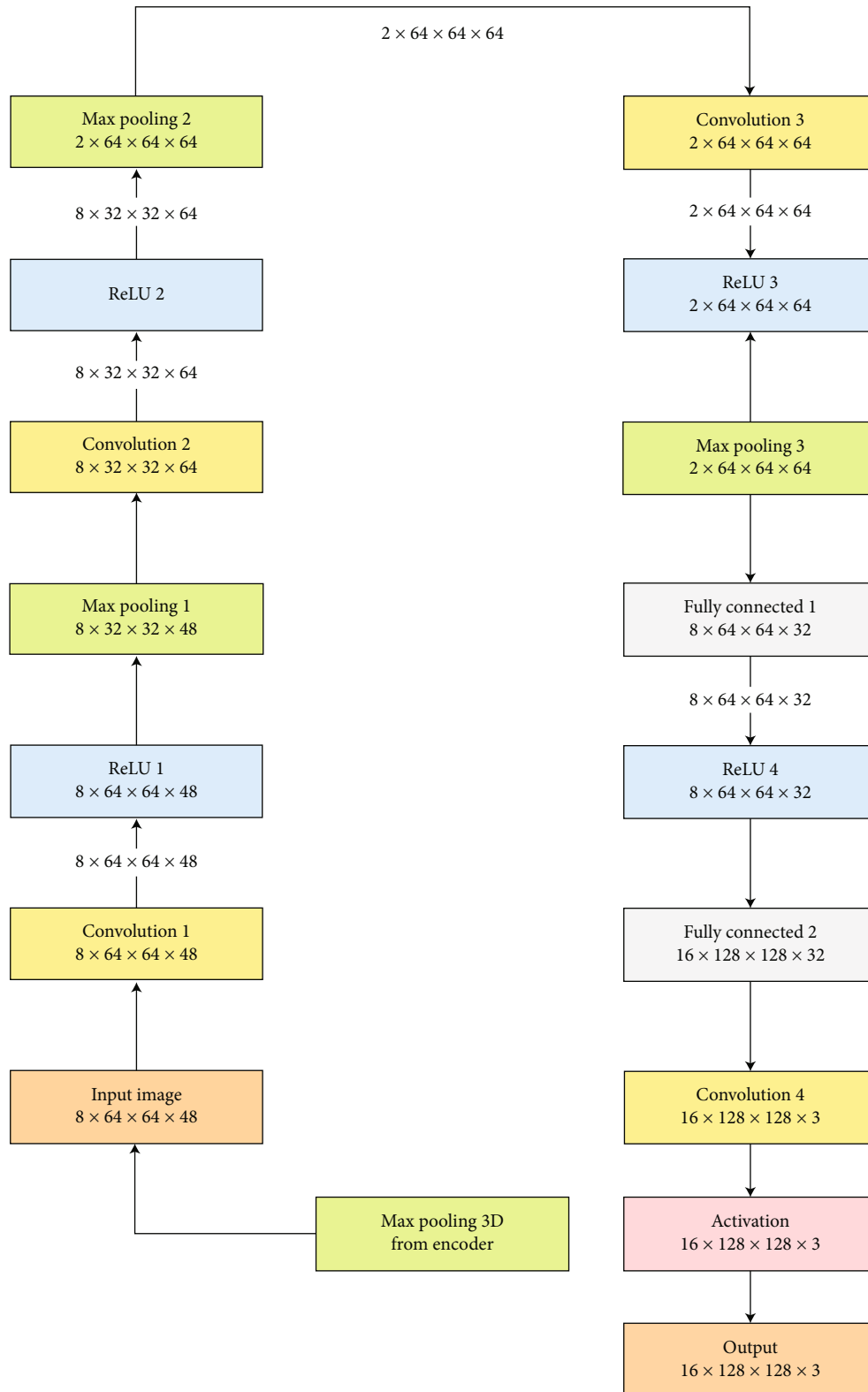


FIGURE 4: Sequential model.

and achieve a certain aim. [21]. Pixel, highlight, and choice combinations are the three layers of combination that can combine information from numerous sources. Because aerosol layer identification and description using object-based

inquiry are more basic and proficient, the current research receives the component level. Low-level highlights for aerosol layer detection are framed using orthophoto highlights (e.g., phantom and textural highlights) and LiDAR

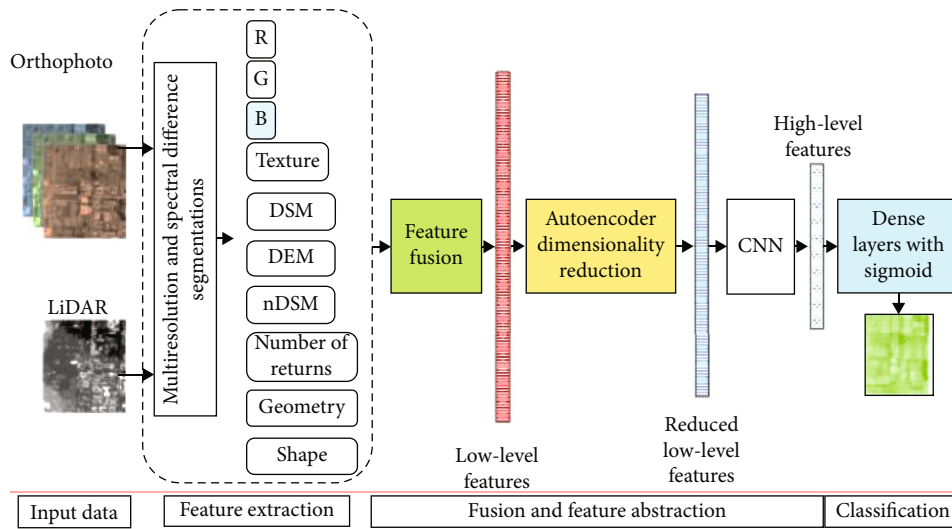


FIGURE 5: Experimental approach.

highlights (e.g., DSM, DEM, nDSM, and spatial highlights) (Table 1).

Many of the features associated with ghastry, textural, geological, and shape collections can be separated from orthophoto and LiDAR data. Overfitting can occur when several highlights are used, especially when the training tests are minimal. Commotion, extra information, and more computer time are some of the drawbacks of employing a large number of highlights. The current study uses an autoencoder-based technology to minimize space dimensionality and improve low-level highlights by reducing them into fewer highlights (i.e., diminished low-level highlights). The new highlights should be more informative than the old ones, and they should improve the overall system work process for recognizing aerosol layers. A CNN model is also evolved by executing several convolution and pooling actions to choose the right highlights for identifying aerosol layers and to turn lowered low-level highlights into significant level highlights. The autoencoder and CNN models are used to decrease (or abstract) low-level highlights in the next sections.

The model has been properly trained when the reconstruction error is modest. The model was not sufficiently trained if the inaccuracy was significant. Testing photos are used to evaluate the model after it has been trained. All of the layers in the autoencoder with the dense layer are fully connected. The information is passed through the bottleneck layer between the encoder and the decoder. We only use one frame at a time in a simple autoencoder. Figure depicts the autoencoder visual layer structure and properties. Each convolution layer has a  $3 \times 3$  filter size with 128, 192, and 256 filters. The convolution's filtering processes are combined in the max-pooling layer, which is  $3 \times 3$ . The image volume is normalized using the normalization layers. The activation function is performed using RLU layers. The number of aerosol frames is detected on the softmax layer using the loss function. For training, the loss function is utilized. The sigmoid response value varies between 0.5 and 0.7. Figure 6 shows the structure of the max pooling layer.

Figure 6 shows an autoencoder that uses layers to translate the input image frame into a feature vector for internal representation (batch normalization, ReLU activation function, and Conv3D). The internal representation is used by the decoder. In the third column, it reverts to the original reconstructed picture frames, and in the second column, it expresses the shape in vector form.

**3.1. Model Sequential.** The sequential model is made by applying a 3D convolutional neural network, varying the number of filters in convolutional layers. It will make it suitable for a basic stack of layers where each layer has accurately one input tensor and one output tensor. It will create its weights the first time it is called on an input image since the shape of the weights depends on the shape of the image frames. Before completing training, a model configures the learning process, which is done via the compile function. It receives three arguments optimizer, loss function, and a list of metrics.

An optimizer should be the string identifier or call to an optimizer function. In the sequential model, the main aim is to minimize the loss function. It is a string identifier to call a loss function, e.g., a loss means squared error.

The output of the fully connected layer CNN is a softmax, and the sigmoid function is used for combining the result of each layer in the sequential model that is shown in Figure 7. For classification, a median filter of size three is applied to the output conclusion to smooth out variations in the classification of anomalies.

Figure 7 above shows the creation of a sequential model by applying a 3D convolutional network, changing the number of filters in convolutional layers. It will make a suitable plan for input to output weights of the shape depend upon the image frame.

**3.2. Generative Adversarial Network (GAN).** The GAN model has been used for the reconstruction of the image with HD resolution. Here, we use the model.

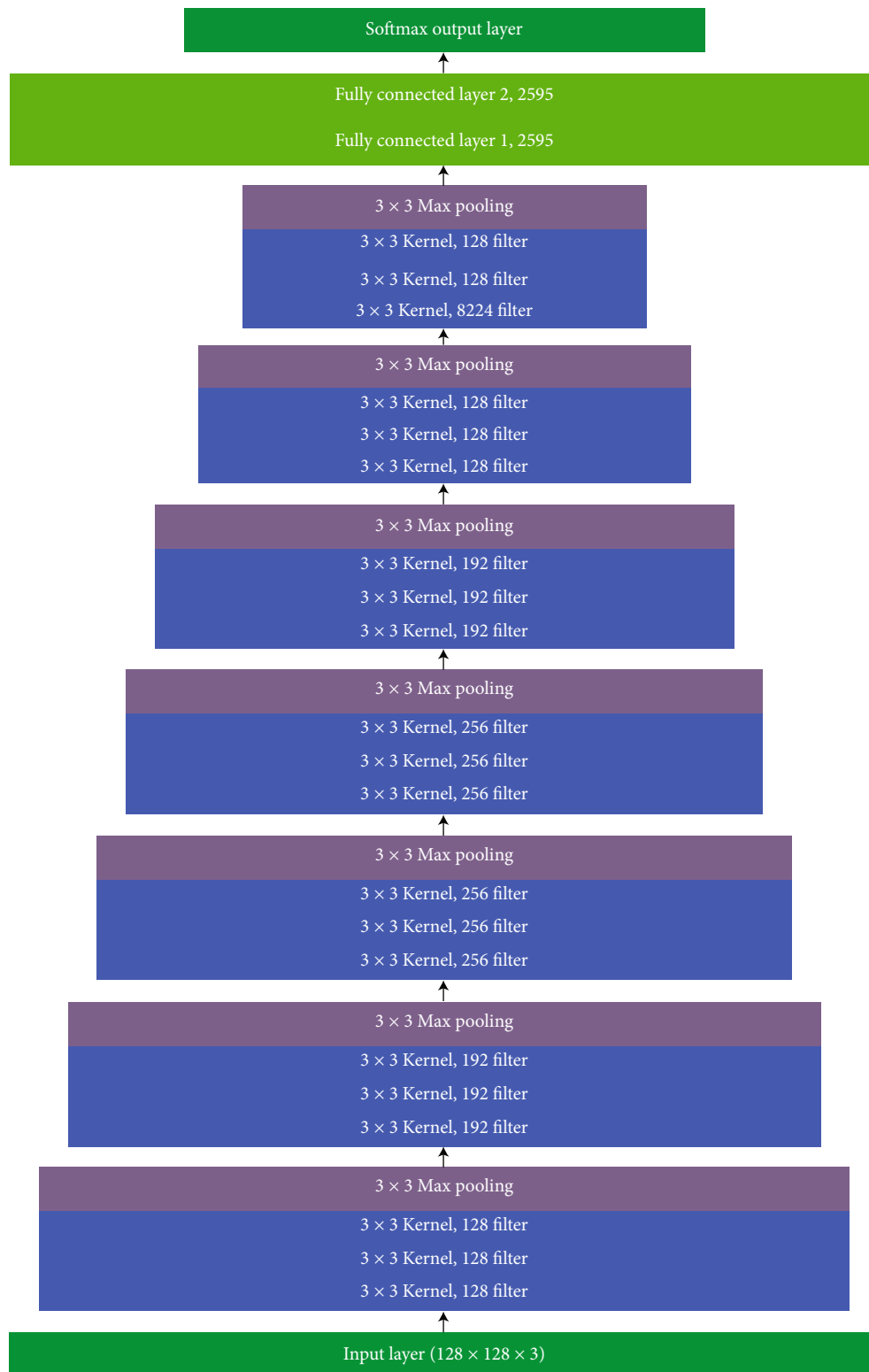


FIGURE 6: The structure of the max pooling layer.

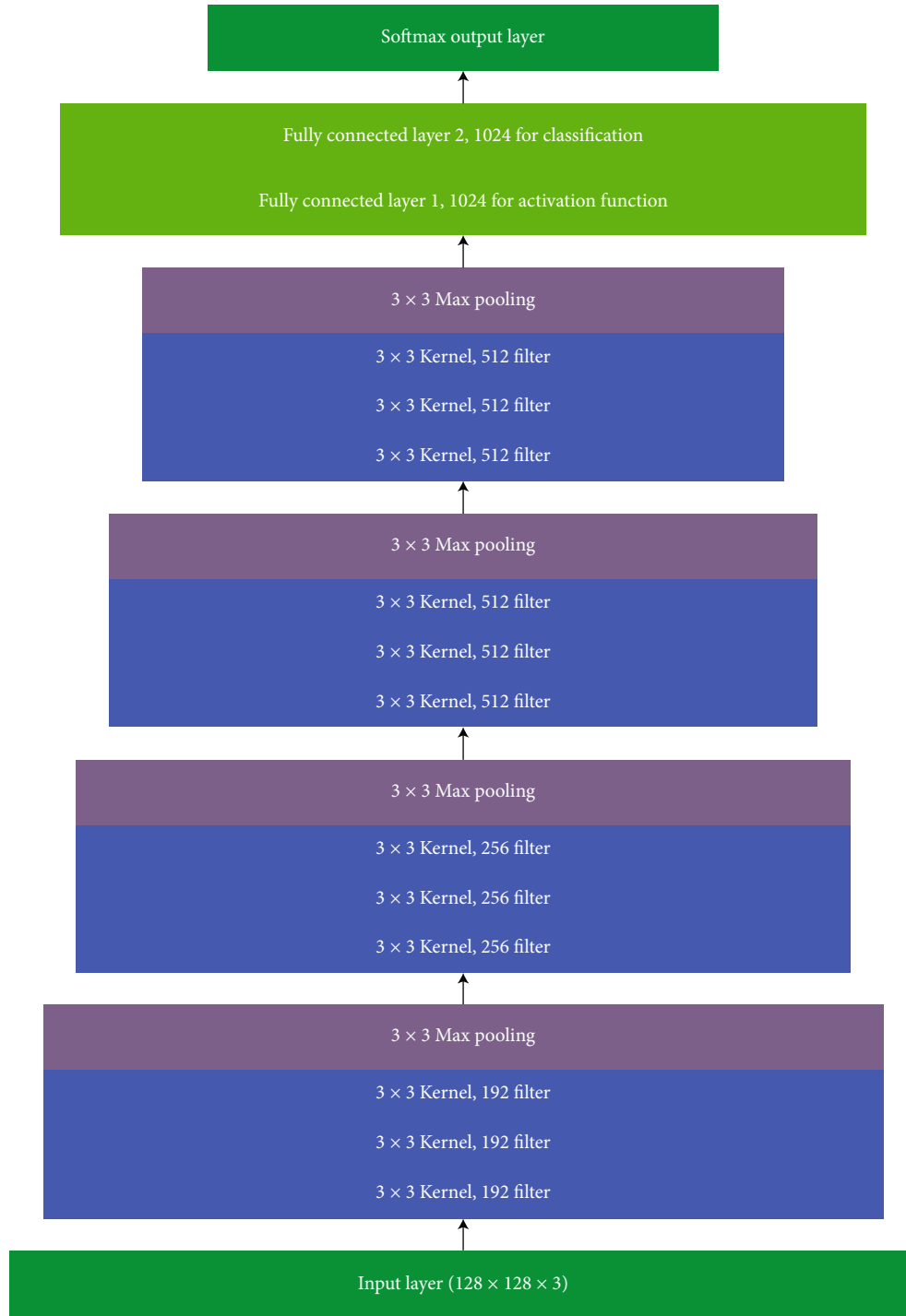


FIGURE 7: The internal structure of the sequential model.

Following are the parameters on which basis we have to evaluate our work. To validate the suggested technique, performance measures such as accuracy, sensitivity, specificity, and AUC are determined. The following are the performance parameters of the suggested technique:

$$\text{Specificity} = \frac{\text{TN}}{(\text{TN} + \text{FP})},$$

$$\text{Accuracy} = \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{TN} + \text{FP} + \text{FN})},$$

$$\text{Sensitivity} = \frac{\text{TP}}{(\text{TP} + \text{FN})}$$

$$\text{AUC} = \frac{\text{Sensitivity} + \text{Specificity}}{2},$$

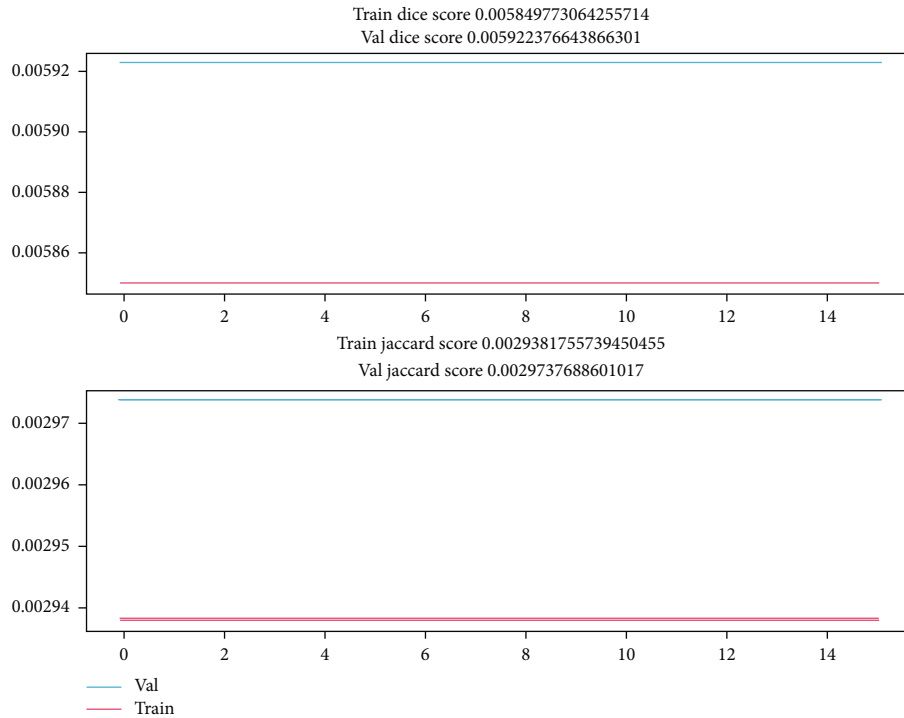


FIGURE 8: Training of data.

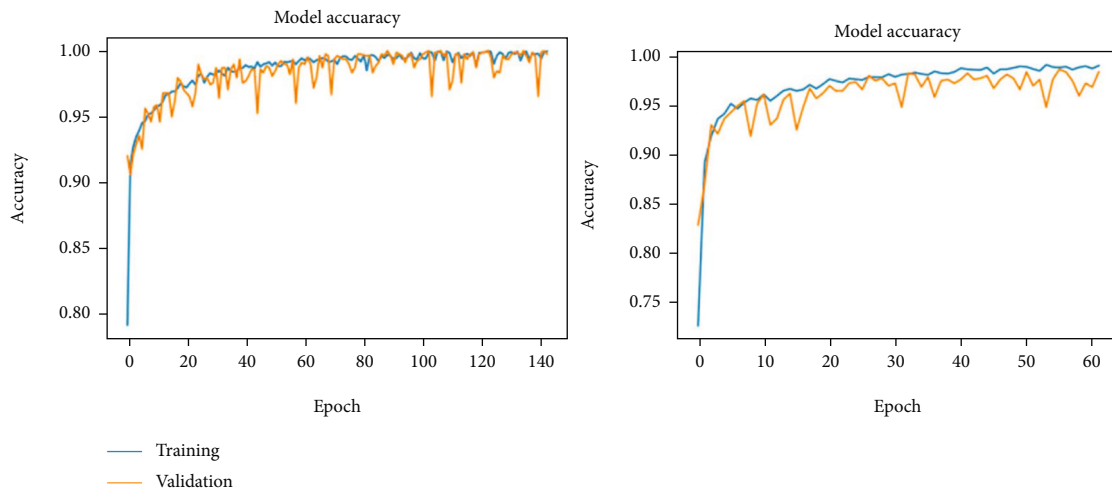


FIGURE 9: Model accuracy and loss.

$$\text{Recall} = \frac{TP}{TP + FN},$$

$$\text{Precision} = \frac{TP}{TP + FP},$$

$$F1_{\text{Score}} = \frac{2 * (\text{Recall} * \text{Precision})}{(\text{Recall} + \text{Precision})}. \quad (1)$$

(i) False-negative (FN): the feature detected result is 0, and predictive powers are present

(ii) True-negative (TN): the feature detected result is 0, and predictive power is absent

(iii) False-positive (FP): the feature detected result is 1, and predictive power is absent

(iv) True-positive (TP): the feature detected result is 1, and predictive power is present

3.3. *Training.* In the first step, we have trained our model on 70% data the training loss 0.00186344. Figure 8 shows training of data.

3.4. *Testing.* The testing of data on the GAN model has been carried out at 30% of set. Model accuracy to detect aerosols

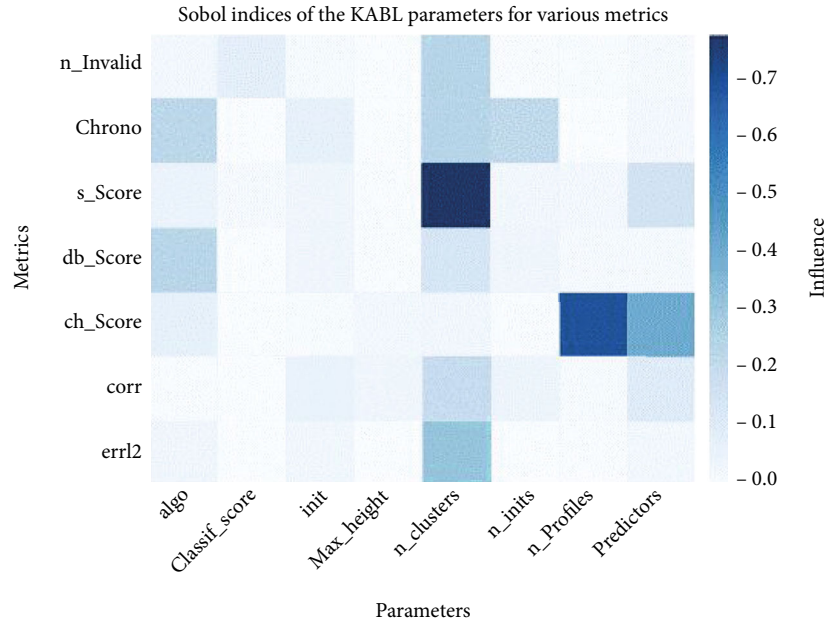


FIGURE 10: Confusion matrix for classification.

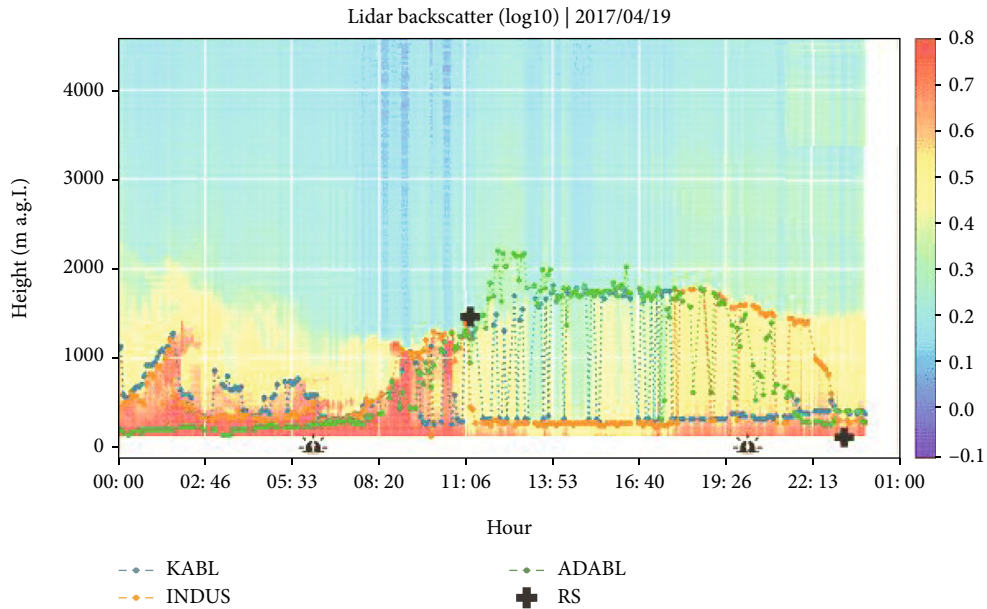


FIGURE 11: Aerosol detection using the GAN model.

is almost 98% at 140 epochs while at 60 epochs, it is 97%. Figure 9 shows model accuracy and loss.

Figure 10 above shows the training and validation accuracy and loss of the CAE model in classifying the images of aerosols from the datasets. The model has shown 98% accuracy of training and 98.7% accuracy of validation during experiments. Figure 11 shows aerosol detection using the GAN model.

Figure 12 shows aerosol outlier detection using at different time spans. While comparing with the GAN model on the right side at 60 epochs, CAE has shown the accuracy of 98% on training and 99% on testing with the inclusion of

the GAN model. Accuracy curves for training and validation have no dropout when combining DSM and RGB (left) and loss of information when using RGB alone (right), both without dropout. Thus, the parameters were examined and hypertuned to maximize detection accuracy. According to the findings of the sensitivity analysis of these parameters, the best 10-fold crossvalidation accuracy for aerosol detection was achieved for the area in which the tests were done. The study concluded that 128 filters delivered an accuracy of 98.76% percent. The poorest results (15.5% accuracy) were seen when using 64 filters. Adam also delivers an accuracy of 81.41%, which is much superior to other optimization

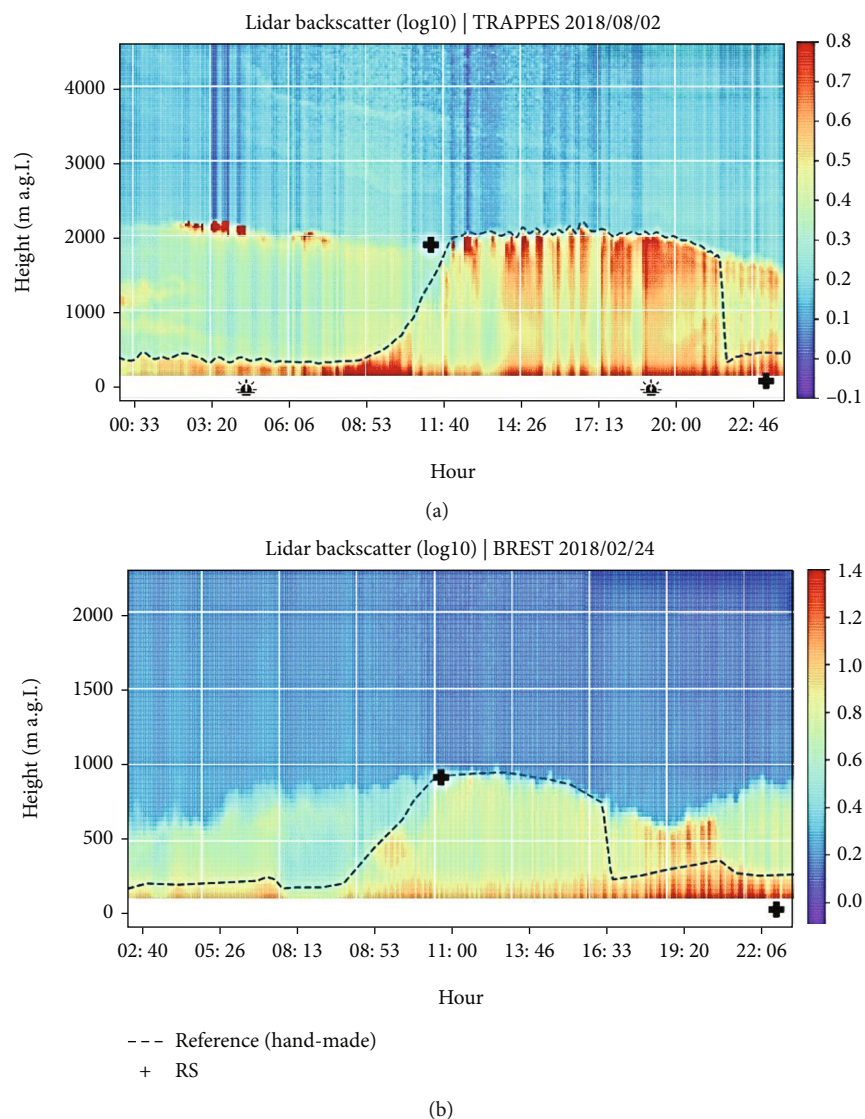


FIGURE 12: Aerosol outlier detection using at different time spans.

TABLE 3: Comparisons of current study with previous studies.

References	Techniques	Accuracy	Outcome
[2]	CNN	89%	Object recognition
[1]	Mark CNN	88%	Waveform recognition
[5]	Mod CNN	88.5%	Biomass land recognition
Our proposed	CAE with GAN	98%	Outlier aerosol recognition
	ADAM	81.41%	
	AdaGrad	23.57%	
	AdaDelta	38.95%	
	SGD	78.41%	
	RMSProp	69.55%	

techniques. Conversely, the dense layer has ignored concealed units with no consequence.

In testing, the most accurate (98.76%) results were produced using 10 or 100 units. The results for the 50- and 3-

unit units were slightly less accurate (81.61 percent). The application of a smaller number of units in the fully linked layer is beneficial to the computational performance of the model; hence, it is considered best to utilize 10.

We have also compared our technique with state-of-the-art algorithms, as shown below in Table 3.

#### 4. Conclusions

The researchers employed autoencoders and CNN models to detect aerosols in a LiDAR–orthophoto dataset, resulting in a DL method. The architecture is designed to generate objects using multiresolution and spectral difference segmentations. The identification of 9 distinct features, including spectral, textural, LiDAR, and orthofusion, was completed for the detection of aerosols. Next, they were compressed into 10 features at the feature level, using the autoencoder model. To categorize the items, they employed the high-level features generated from the modified compressed features. Building detection using this design has many advantages, including automated feature selection and removal of redundant characteristics. Convolutional neural networks (CNNs) were utilized to convert compressed information into high-level characteristics that could categorize the outer layer of atmospheric particles. This research describes deep learning approaches that, when applied to Lidar data, allowed for detecting 40% more atmospheric features at a horizontal resolution of 15 km during daytime operations. In comparison to existing deep learning algorithms for edges and complicated near-surface sceneries during the day, a convolutional autoencoder (CAE) trained using LiDAR Dataset standard data products showed the potential for improved aerosol discrimination. However, the dataset including height information (the fused orthomosaic photo and DSM) performed better in most discriminative classifications. This study demonstrated CAE's capacity to accurately categorize lower-resolution UAV-fused LiDAR images in comparison to very-high-resolution aerial shots and also indicated that dataset fusion is promising. The model has shown 98% accuracy of training and 98.7% accuracy of validation during experiments. While comparing with the GAN model on the right side at 60 epochs, CAE has shown the accuracy of 98% on training and 99% on testing with the inclusion of the GAN model. The sensitivity of CNNs with various fusion methods to the training dataset, regularization functions, and optimizers will be the subject of future research.

#### Data Availability

Dear Sir/Madam, as you are concerned with the availability of my code and practical work, it is confidential to my lab and cannot be shared to anyone until it gets published. And secondly, I will publish my code and lab work according to the instructions of my supervisor: Lab name: supervisor name: Regards Mazhar Hameed.

#### Conflicts of Interest

On behalf of all authors, the authors declare that there is no conflict of interest.

#### Acknowledgments

I would like to acknowledge my indebtedness and render my warmest thanks. Thanks are due to my supervisor, Professor Yang Fengbao, who made this work possible. His friendly guidance and expert advice have been invaluable throughout all stages of the work. I would also wish to express my gratitude to Miss Gao Min for her extended discussions and valuable suggestions which have contributed greatly to the improvement of the paper. This work was supported by the National Natural Science Foundation of China (Grant Nos. 61672472 and 61972363), Science Foundation of North University of China, Postgraduate Science and Technology Projects of North University of China (Grant No. 20181530), and Postgraduate Education Innovation Project of Shanxi Province.

#### References

- [1] T. Shinohara, H. Xiu, and M. Matsuoka, "FWNet : semantic segmentation for full-waveform LiDAR data using deep learning," *Sensors*, vol. 20, no. 12, article 3568, 2020.
- [2] L. Zhang, Z. Shao, J. Liu, and Q. Cheng, "Deep learning based retrieval of forest aboveground biomass from combined LiDAR and Landsat 8 data," *Remote Sensing*, vol. 11, no. 12, p. 1459, 2019.
- [3] J. Rock, M. Toth, P. Meissner, and F. Pernkopf, "CNNs for interference mitigation and denoising in automotive radar using real-world data," in *Machine Learning for Autonomous Driving Workshop at the 33rd Conference on Neural Information Processing Systems (NeurIPS 2019)*, pp. 1–9, Inffeldgasse 16c Graz Austria, 2019.
- [4] L. Duan, D. Zhang, F. Xu, and G. Cui, "A novel video encryption method based on faster R-CNN," in *2018 International Conference on Computer Science, Electronics and Communication Engineering (CSECE 2018)*, vol. 80, pp. 100–104, Inffeldgasse 16c Graz Austria, 2018.
- [5] G. Melotti and C. Premebida, "Multimodal deep-learning for object recognition combining Camera and LIDAR data," in *2020 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC)*, Ponta Delgada, Portugal, 2020.
- [6] A. Wahid, A. Stone, K. Chen, B. Ichter, and A. Toshev, "Learning object-conditioned exploration using distributed soft actor critic," 2020, <https://arxiv.org/abs/2007.14545>.
- [7] J. Zhang, "Multi-source remote sensing data fusion: status and trends," *International Journal of Image and Data Fusion*, vol. 1, no. 1, pp. 5–24, 2010.
- [8] B. Major, D. Fontijne, A. Ansari et al., "Vehicle detection with automotive radar using deep learning on range-azimuth-doppler tensors," in *Proceedings -2019 International Conference on Computer Vision Workshop*, vol. 2019, pp. 924–932, Seoul, Korea, 2019.
- [9] M. Hameed, F. Yang, M. I. Ghaffoor et al., "IOTA-based Mobile crowd sensing: detection of fake sensing using logit-boosted machine learning algorithms," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 6274114, 15 pages, 2022.
- [10] C. Sun, J. M. U. Vianney, and D. Cao, "Affordance learning in direct perception for autonomous driving," 2019, <https://arxiv.org/abs/1903.08746>.



- [11] A. Atghaei, S. Ziaeinejad, and M. Rahmati, "Abnormal event detection in urban surveillance videos using GAN and transfer learning," pp. 1–7, 2020, <https://arxiv.org/abs/2011.09619>.
- [12] P. Sedigh, R. Sadeghian, and M. T. Masouleh, "Generating synthetic medical images by using GAN to improve CNN performance in skin cancer classification," in *ICRoM 2019 - 7th international conference on robotics and mechatronics*, pp. 497–502, Tehran, Iran, 2019.
- [13] M. Gan and R. Jiang, "ROUND: walking on an object-user heterogeneous network for personalized recommendations," *Expert Systems with Applications*, vol. 42, no. 22, pp. 8791–8804, 2015.
- [14] V. Costa, J. Correia, N. Lourenço, and P. Machado, "COE-GAN: evaluating the coevolution effect in generative adversarial networks," in *Proceedings of the genetic and evolutionary computation conference*, pp. 374–382, Prague, Czech Republic, 2019.
- [15] M. Aqeel, K. B. Khan, M. A. Azam, M. H. Ghouri, and F. Hassan, *Detection of Abnormal Events in Videos Using Convolutional Autoencoder and Generative Adversarial Network model*, IEEE, 2020.
- [16] K. Kaiyrbekov and M. Sezgin, "Deep stroke-based sketched symbol reconstruction and segmentation," *IEEE Computer Graphics and Applications*, vol. 40, no. 1, pp. 112–126, 2020.
- [17] S. Li, J. Kawale, and Y. Fu, "Deep collaborative filtering via marginalized denoising auto-encoder," in *Proceedings of the 24th ACM international on conference on information and knowledge management*, Melbourne, Australia, 2015.
- [18] P. Kharazmi, J. Zheng, H. Lui, Z. Jane Wang, and T. K. Lee, "A computer-aided decision support system for detection and localization of cutaneous vasculature in dermoscopy images via deep feature learning," *Journal of Medical Systems*, vol. 42, no. 2, p. 33, 2018.
- [19] S. Hamdi, S. Bouindour, K. Loukil, H. Snoussi, and M. Abid, "Two-streams fully convolutional networks for abnormal event detection in videos," in *Proceedings of the 12th International Conference on Agents and Artificial Intelligence (ICAART 2020) - Volume 2*, pp. 514–521, Inffeldgasse 16c Graz Austria, 2020.
- [20] M. Sewak, S. K. Sahay, and H. Rathore, "Comparison of deep learning and the classical machine learning algorithm for the malware detection," in *Proceedings -2018 IEEE/ACIS 19th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing*, pp. 293–296, Busan, Korea (South), 2018.
- [21] Y. Li, Q. Pan, S. Wang, T. Yang, and E. Cambria, "A generative model for category text generation," *Information Sciences*, vol. 450, pp. 301–315, 2018.