

Multisensor Data-Fusion-Based Gas Hazard Prediction Using DSET and 1DCNN for Underground Longwall Coal Mine

Mayank Sharma^{ID}, Member, IEEE, and Tanmoy Maity^{ID}, Member, IEEE

Abstract—The combustible and noxious gases are among the prominent issues affecting the life and the mining operations of underground coal mines (UCMs). Commonly adopted hazard monitoring methodologies in UCM are fuzzy, rule-based systems, statistical methods, and other expert systems, but these models are not reliable for highly complex and nonlinear systems. The neural network's ability to learn and create nonlinear relationships is beneficial to making hazard prediction models. Especially, convolutional neural networks (CNNs) auto feature extraction capabilities to make it more suitable for the task. But, UCM's harsh and crucial environment may result in sensor malfunctioning and faults, giving rise to data uncertainty. Like other data-driven models, data uncertainty significantly affects CNN's performance. This study involves designing an effective and reliable gas hazard monitoring system using a hybrid of Dempster–Shafer evidence theory (DSET)-based filter and one-dimensional CNN (1DCNN) classifier. The novelty of this study is the integration of DSET and 1DCNN to predict the UCM hazard more reliably, even in malfunctioning node scenarios. Inherent usage limitations on traditional communication techniques restrict the application of cloud-based machine learning (ML) methods and this study uses novel edge implementation of the filter and the classifier using edge ML (EML) technology. The proposed model's hazard classification accuracy is 99.6%, even in the faulty node scenarios, where the traditional approaches fail.

Index Terms—Classifier, convolutional neural network (CNN), Dempster–Shafer evidence theory (DSET), edge machine learning (EML), gas hazard prediction, underground coal mine (UCM).

I. INTRODUCTION

THE LONGWALL coal mining operation has the advantage of a higher production rate, but it produces more methane (CH_4) gas than the conventional board and pillar method. Studies in [1] and [2] revealed that for larger longwall faces, methane emission increases, and the highest CH_4 concentration is observed at the tailgate side caused by the outflow of methane from the goaf. Tutak *et al.* [1] have used an automatic methane monitoring model by installing separate sensing units at different locations. The study in [2] has adopted a section-based monitoring method where a longwall face is divided into three or four equal sections, and

Manuscript received 9 March 2022; revised 19 April 2022; accepted 13 May 2022. Date of publication 17 May 2022; date of current version 24 October 2022. (Corresponding author: Mayank Sharma.)

The authors are with the Department of Electrical Engineering, Indian Institute of Technology, Dhanbad 826004, India (e-mail: mayank.17dr000449@mme.ism.ac.in; tanmoy@iitism.ac.in).

Digital Object Identifier 10.1109/JIOT.2022.3175724

methane concentration is monitored using automatic methane sensors. Another work in [3] has used an ANN-based approach to predict a longwall's gas concentrations using a distributed sensor network. The main drawback of these system types is their blind dependency on individual sensor observation. The single-sensor-based decision model is always prone to sensor malfunction, resulting in false alarms and misinterpretation of the actual event. For the underground coal mine (UCM) scenarios, such a wrong interpretation of hazardous situations may result in fatal consequences. The solution to this issue is a fusion of multiple heterogeneous or homogeneous type sensors that leads to a refined decision and is less affected by individual sensor faults.

Dempster–Shafer evidence theory (DSET)-based data-fusion models are the most efficient to tackle data uncertainties [4], but the DSET has the disadvantage of higher computational cost as the number of parameters increases. The nonlinearity and adaptive learning features of deep neural network (DNN) models make them suitable for almost every task where the data related to the study are available. Convolutional neural networks (CNNs) are among the most popular algorithms for classification tasks. The CNN-based models have the advantage of less storage requirement due to the shared kernel weights. On the other hand, CNN or any DNN-based model suffers critically from data uncertainty [5]. This study proposes an edge machine learning (EML)-based reliable gas hazard prediction (EGHP) system for UCM longwall operations using a hybrid of the DSET-based filter and the one-dimensional CNN (1DCNN) classifier. The proposed model utilizes the advantages of DSET to filter out the malfunctioning sensor observations and 1DCNN to predict the hazard class using the filtered raw sensor observations. This hybrid multisensor data-fusion approach is one of the first attempts in the field of UCM. This approach also counters the drawbacks of the DSET and DNN-based 1DCNN. In this study, the functionality of DSET is limited to identifying any sensor observation as healthy or malfunctioning (faulty). Hence, the computational cost is fixed and very low. The 1DCNN classifier uses the healthy sensor data from the DSET filter to identify the hazard class.

The dynamic and complex UCM environment imposes many restrictions on the traditional hazard monitoring systems. On the completion of coal extraction in one of the mine galleries, there is a requirement to move the whole machinery and infrastructure to the next operating region. This dismantling

and remounting process imposes limitations on the fixed communication infrastructures. Moreover, wired communication networks are prone to mechanical and other damages. The subsurface operations of the UCMs put restrictions on traditional wireless systems. Additionally, these wireless systems face high signal attenuation due to the coal dielectric properties and UCM geological features. Hence, cloud-based machine learning (ML) models are not reliable and appropriate for the UCM and, therefore, it makes an ideal ground for EML-based systems, where the whole neural network is implemented directly on edge devices like microcontrollers [6]. Hence, the Dempster–Shafer filter (DSF) and 1DCNN are implemented on the 32-bit ARM CORTEX M4 microcontroller, making the system more independent and secure than other models.

The proposed EGHP model consists of a sensor-node array and a DSET-based filter (DSF) for the different sections of the longwall panel and a 1DCNN classifier which takes the output of DSFs from each section and classifies the hazard class. A sensor-node array consists of five identical sensor nodes. Each sensor node has three sensors for methane (CH_4), carbon dioxide (CO_2), and carbon monoxide (CO) gas. The reason behind selecting these target gas is the majority of the gas hazard associated with underground (UG) coal mines are related to these gases [7], [8]. Another reason is that the data related to these gases is available in the public domain. Moreover, the approach for the hazard prediction of the other gases can be the same. The proposed model is compared with the other ML models based on their accuracy in the healthy node and faulty node scenarios. The whole experimentation is realized for a single section of a longwall which consists of five sensor nodes, one DSF controller of this section, and a 1DCNN controller. Among five nodes, three are healthy nodes, and two are test nodes to emulate unhealthy conditions during the experimentation. The remainder of this article is organized as follows. Section II is a literature review, Section III is for methodology, Section IV deals with results and discussion, and the last section is for the conclusion.

II. LITERATURE REVIEW

The gas hazard prediction model presented here is mainly a classification problem where a type of gas hazard is to be identified using the raw sensor data. This section deals with the previous studies related to ML and other expert system-based gas hazard prediction in UG coal mines and the application of DSET for classification problems. Muduli *et al.* [9] and Danish and Onder [10] have proposed a fuzzy-logic-based multisensor data-fusion system to predict fire intensity in UCMs using temperature, CO_2 , CO, and oxygen sensors from the various locations of UG mines. These fuzzy systems implement *IF – Then* rules to classify the fire intensity in the UG mine. The study in [11] has proposed a framework for the UCM early warning and event detection system based on environmental attributes and miner localization. They developed a mine warning index to interpret any environment-related event in the UG coal mine. They also incorporated the K means algorithm for anomaly detection in the data. Wu *et al.* [12] has proposed a deep belief network-based methane intensity

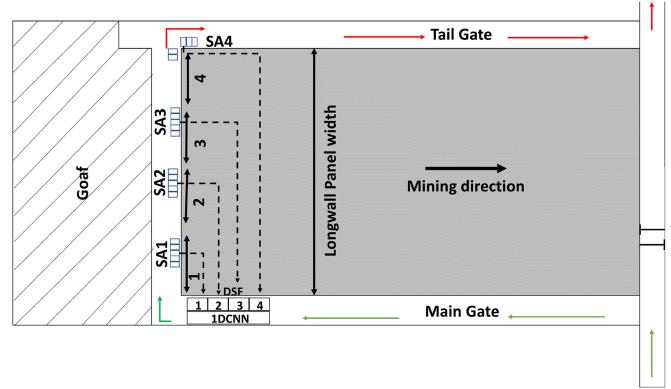


Fig. 1. Gas monitoring system schematic.

prediction model. A Least square support vector machine (LS SVM)-based methane hazard classifier is proposed in [13]. All these models, including the fuzzy-based models, are suitable for the scenario when data sample dimensionality is very low. For higher dimensional data with multiple attributes, the performance of these systems suffers greatly. A solution to this issue is a DNN-based CNN model. The study in [14] is dedicated to the CNN and long short-term memory (LSTM)-based hybrid hazard monitoring system for UG coal mines. The model forecasts toxic and flammable gas concentrations inside the goaf in a coal mine. But all the models discussed above are prone to sensor malfunction. Hence, there is a requirement to identify such erroneous observations, and DSET is a proven way to deal with such conditions. Although the DSET has an extensive operational scope, it is still not implemented in UG coal mine scenarios. Some significant studies which inspired the design of DSF filters are as follows.

Nesa and Banerjee [15] presented a Dempster–Shafer (DS)-based room occupancy detection model. The model used evidence based on room temperature, light, humidity, and CO_2 sensor readings. In [16], the study is focused on the smart environment using heterogeneous data. They proposed a DS-based combination of heterogeneous sensor data with an improvised basic belief assignment (BBA) function using modified belief entropy. Ghosh *et al.* [17] is dedicated to applying DS theory to identify faulty node data in an Internet of Things (IoT) application. Many articles propose DS theory is more suitable where uncertainty and erroneous sensor data are involved.

III. METHODOLOGY

For this study, a longwall panel is considered with its face width divided into four sections represented by numbers 1–4. Four arrays (SA1–SA4) consisting of five sensor nodes each are placed in each section, as shown in Fig. 1. Each node comprises CH_4 , CO_2 , and CO sensors sampling every second. A double layer sensor fusion model is adopted for the EGHP model. The first layer consists of four DSFs for each section (DSF 1–4), and the second layer has a 1DCNN block. Each section's sensor array (SA) transmits its observations to its respective DSF, represented by a dotted line. Then, the output of the DSF filter goes to the 1DCNN classifier for hazard

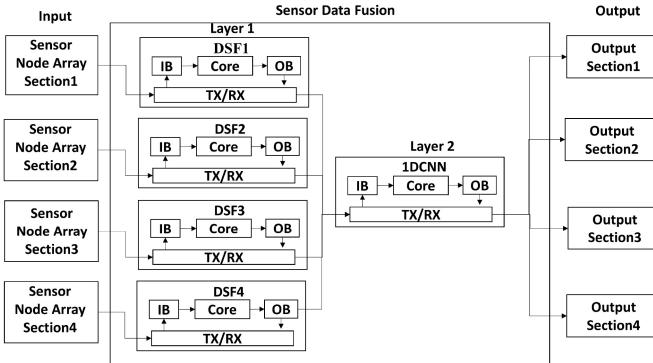


Fig. 2. Block diagram for EGHP.

prediction. For each DSF, the 1DCNN classifier input consists of three data frames. Each data frame represents the gas concentrations for CH_4 , CO_2 , and CO. The rise or fall of the gas concentration in UCM cases for seconds or minutes-based data frames is gradual as compared to the hours or days-based data frames, where the rise and fall can be sudden. Since, in this study, the data frame consists of 60-s data samples, concentration changes are gradual and can be detected within one or two data frames; hence, the proposed model can tackle such occurrences effectively. The whole operation can be understood with the help of a block diagram, as shown in Fig. 2.

The block diagram consists of three main blocks, i.e., the sensor-node array of each section as the input, sensor data-fusion block, and output block. The raw data from sensor-node arrays come to the sensor data-fusion block, which is further divided into two layers. Layer 1 consists of four DSF blocks for all four sections, and layer 2 has a 1DCNN block.

In DSF, the raw data initially come to the input buffer (IB) through the TX/RX block. The TX/RX block provides a way of communication between controllers and sensor nodes, e.g., USART. Then, the process of faulty node data filtering is done in the core, which is the controller's processing unit that does all computations dedicated to DSF. Then, the mean of the filtered data samples for each section is stored in their respective output buffers (OBs). The DSF outputs from their OBs are transferred to the 1DCNN block. In the 1DCNN block, IB stores the DSF outputs of each section. The core then uses the data in IB to classify the hazard class of each section. The classification result is stored in the OB of the 1DCNN block. The prediction output of the 1DCNN is provided to the output block representing the hazard indicators.

A. Data Sample Description

The dataset to train the proposed model is compiled from the sources [18]–[21] representing different UCM ventilation air (VA) data. These sources consist of the various VA gas concentrations and their statistical features. Especially, the data source [19] represents the 9 199 930 s of data samples from a Polish coal mine. Additionally, Schatzel *et al.* [2] described the methane gas flow patterns across the face of longwall operations. Using the VA gas distribution properties by the data sources mentioned above, 78 000 input data samples of 60 s

TABLE I
CONCENTRATION OF ATTRIBUTE GASES BY AVERAGING
EVERY 60 S DATA POINTS

Sample number	CH_4 ppm	CO_2 ppm	CO ppm
1	11605	3455	11
2	11064	4630	7
3	22910	905	4
4	15932	1306	22
5	50463	3361	26
6	9305	8964	99
7	23773	790	12
8	44328	1921	157
9	9453	2746	32
10	126	6508	418

TABLE II
GAS HAZARD CLASSES FOR UG LONGWALL MINE

Hazard Class	Gases beyond their hazardous level
Class1	All gases are within safe limit
Class2	CH_4
Class3	CO_2
Class4	CO
Class5	$CH_4 + CO_2$
Class6	$CO_2 + CO$
Class7	$CH_4 + CO$
Class8	$CH_4 + CO_2 + CO$

each are sampled for each attribute, i.e., for CH_4 , CO, and CO_2 . Some portions of the mean value of data samples are shown in Table I.

Based on these three attributes, eight hazard classes are defined in Table II.

The hazardous level of gas is subjected to the threshold limit value (TLV) of gas, which may result in accidents. The reference or TLV for various gases is presented in [22]. The TLV for CH_4 , CO_2 , and CO are 50 000, 20 000, and 50 ppm. The alarm setting limit (ASL) is the gas concentration normally set below the gas TLV as a safety measure in the underground mine. The authorized body of the government defines the ASLs based on the mine type, location, and hazard history. In India, the Directorate General of Mines Safety (DGMS) is responsible for such decisions. In this study, ASL for CH_4 is 12 500 ppm, for CO_2 , it is 5000 ppm, and for CO is 35 ppm. The data beyond the ASL and TLV are critical concentrations of gases that may result in hazards if not treated well. The whole data are divided into three categories, i.e., training, validation, and test sets; 70% is the training set, and 30% belongs to the validation and test set. Z-score normalization is used to normalize the samples. For the hardware emulation, calibration gases of CH_4 , CO_2 , and CO are used. The different concentration levels are prepared by diluting them with air. One separate data sample is prepared for various malfunctioning node scenarios using faulty sensors and purposefully injected faults. These datasets are dedicated to testing the classifier's performance integrated with DSF for the cases when the model encounters anomalies due to sensor-node malfunctioning. These samples are not involved in training the

classifier. The detailed description of DSF1 and 1DCNN is as follows.

B. Dempster-Shafer Filter

The DSF proposed in this study utilizes the concept of DSET to come up with a combined decision about any hypothesis. The fundamental requirement of this method is a dedicated mass function, also known as the basic belief function (BBA) and the Dempster combination rule. For the basic terminologies and theory of DSET, one is referred to [15]–[17]. A detailed description of DSET in this context is as follows.

1) *Basic Belief Assignment*: The initial step for the BBA function is to define a frame of discernment (FOD), a mutually exclusive and exhaustive set of all possible hypotheses. Since the sensor node consists of three sensors for CH_4 , CO_2 , and CO gases, their FODs are created as follows:

$$FOD_x = \{\{H\}, \{N\}\} \quad (1)$$

where subscript x signifies sensor type, i.e., CH_4 , CO_2 , and CO. Apart from the node-level hazard classes given in Table II, the BBA considers two categories of hazards at the sensor level. Therefore, in FOD, H and N indicate a hazardous class and a nonhazardous class of gas concentrations, respectively. The power set of any sensor x is defined as all possible subsets of elements in FOD_x , i.e.,

$$POW_x = 2^{FOD_x} = \{\{H\}, \{N\}, \{H, N\}, \text{null}\}. \quad (2)$$

The BBA is a function that maps all elements of POW_x into $[0, 1]$ scale based on the sensor x observation. It can be represented as follows:

$$m : e \rightarrow [0, 1] \quad \forall e \in POW_x \quad (3)$$

where e represents the elements of POW_x . An important fact that needs to be mentioned here is the mass of the null set in POW_x shall be 0, and the summation of masses of all elements in POW_x except null set shall be 1, i.e.,

$$\begin{aligned} m\{\text{null}\} &= 0 \\ m\{H\} + m\{N\} + m\{H, N\} &= 1. \end{aligned}$$

The BBA function for each sensor in this study depends on the difference between ASL and sensor reading. It is derived as follows:

$$m_x\{N\} = \begin{cases} 0.5 + |DF1_x|, & \text{if } DF1_x > 0 \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

$$m_x\{H\} = \begin{cases} 0.5 + |DF2_x|, & \text{if } DF2_x \leq 0 \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

$$m_x\{H, N\} = \begin{cases} 0.5 - |DF1_x|, & DF1_x > 0 \\ 0.5 - |DF2_x|, & DF2_x \leq 0. \end{cases} \quad (6)$$

In (4)–(6), $m_x\{N\}$ represents the mass assigned to N class for sensor x , similarly, $m_x\{H\}$ and $m_x\{H, N\}$ represent the mass assigned to class H and uncertainty, respectively. $DF1$ and $DF2$ are derived as follows:

$$DF1_x = \frac{Diff_x \times SF1_x}{Ref_x} \quad (7)$$

where

$$Diff_x = Ref_x - Obs_x. \quad (8)$$

$DF1_x$ in (7) represents the difference factor, and $SF1_x$ represents the scaling factor for sensor x . In (8), $Diff_x$ represents the difference between the threshold value (Ref_x) and reading (Obs_x) of sensor x . $SF1_x$ is used to scale the $DF1_x$ value from 0 to 0.5 for the variation in Obs_x from 0 to Ref_x . For this study, $SF1_x$ is found to be 0.5 for each sensor

$$DF2_x = \frac{Diff_x \times SF2_x}{Ref_x} \quad (9)$$

where

$$SF2_x = \frac{Ref_x \times 0.5}{Ref_x - Obs_x(\text{Full Scale})}. \quad (10)$$

In (9) and (10), $DF2_x$ is the difference factor and $SF2_x$ is the scaling factor for $Diff_x < 0$ cases. $SF2_x$ scales the value of $DF2_x$ from 0 to 0.5 for the variation in Obs_x from Ref_x to its full-scale reading ($Obs_x(\text{Full Scale})$).

2) *Combination of Mass Using Dempster Combination Rule*: Since each sensor-node array consists of five nodes and each node has three sensors, there are five BBA for each sensor. The combined mass of each element of the power set is given as follows:

$$\begin{aligned} M_x\{H\} &= \frac{1}{1-K} \sum_{A1 \cap A2 \cap A3 \cap A4 \cap A5 = H} m1_x\{A1\} \times m2_x\{A2\} \\ &\quad \times m3_x\{A3\} \times m4_x\{A4\} \times m5_x\{A5\} \end{aligned} \quad (11)$$

where

$$\begin{aligned} K &= \sum_{A1 \cap A2 \cap A3 \cap A4 \cap A5 = \text{null}} m1_x\{A1\} \times m2_x\{A2\} \\ &\quad \times m3_x\{A3\} \times m4_x\{A4\} \times m5_x\{A5\}. \end{aligned}$$

Similarly

$$\begin{aligned} M_x\{N\} &= \frac{1}{1-K} \sum_{A1 \cap A2 \cap A3 \cap A4 \cap A5 = N} m1_x\{A1\} \times m2_x\{A2\} \\ &\quad \times m3_x\{A3\} \times m4_x\{A4\} \times m5_x\{A5\} \end{aligned} \quad (12)$$

where, $M_x\{H\}$ and $M_x\{N\}$ represent the combined mass for classes H and N , respectively, and $m1_x$, $m2_x$, $m3_x$, $m4_x$, and $m5_x$ are masses related to nodes 1, 2, 3, 4, and 5 for sensor x .

3) *Faulty Node Filter*: The result of (11) and (12) is used to decide the combined decision of each sensor in each node. The result with the higher mass will be considered as the final decision about the hazard class, i.e.,

$$\text{class}_x = \begin{cases} H, & \text{if } M_x\{H\} > M_x\{N\} \\ N, & \text{if } M_x\{H\} < M_x\{N\}. \end{cases} \quad (13)$$

Based on the decision in (13), the node whose sensor x does not obey this decision is identified as a faulty node. Hence, after removing erroneous observations of faulty nodes, the mean of the healthy sensor readings of all five nodes is transferred to the 1DCNN block for further process. Although this approach can accurately identify the sensor whose decision does not match with the combined decision, it cannot segregate the sensors whose decision matches with the combined decision, but its raw data differs significantly from the

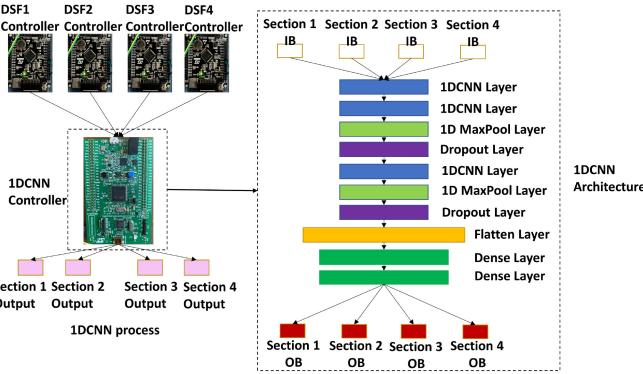


Fig. 3. Second-layer fusion model with architecture.

rest of the healthy sensors. These nodes may result in wrong decisions about any hazard in the future. To identify those nodes which are indicating the correct hazard class but their measurement value differs more than 10% of the healthy nodes sensors, an additional decision vector is designed as follows:

$$D_x = [d12_x, d13_x, d14_x, d15_x, d32_x, d24_x, d25_x, d34_x, d35_x, d45_x]. \quad (14)$$

In (14), D_x is the difference vector of the resultant class obtained from (13) for sensor x . $d12_x$ represents the difference between the sensor x mass among the nodes 1 and 2 for the resultant class obtained in (13). Similarly, other elements of D_x can be interpreted. Now, if the hazard class obtained in (13) is H , then $d12$ for CH_4 sensor is as follows:

$$d12_{CH_4} = |m1_{CH_4}\{H\} - m2_{CH_4}\{H\}|. \quad (15)$$

Similarly, the other elements of the difference vector in (14) can be calculated. Now, the next step is to calculate the support between each node pair masses as follows:

$$S_x = 1 - D_x. \quad (16)$$

The S_x represents the similarity between the masses of each node pair for sensor x . In this study, it is found that for the similarity of more than 97% between the node pair masses, the value of S_x for that pair shall be greater than 0.97. Hence, a highly similar node pair can be obtained as follows:

$$S_x = S_x > 0.97. \quad (17)$$

Using (17), a faulty node having an observational difference of more than 10% from the healthy nodes can be detected accurately. After this step, the mean of healthy node sensor data of all the five nodes from each section is transferred to the 1DCNN block to identify the hazard class of each section, as represented in Table II.

C. 1DCNN-Based Fusion Model

In this layer, the refined data from the various sections of the longwall face using DSF are combined to detect the class of hazard mentioned in Table II. The multi-input–multi-output-based 1DCNN architecture is used for this purpose, shown in Fig. 3. The model consists of 1DCNN layers, MaxPool layers, Dropout layers, and Dense layers.

TABLE III
1DCNN ARCHITECTURE PARAMETERS

Parameters	Value
Number of 1D CNN layers	3
Number of MaxPool layers	2
Number of Fully connected Layers	2
Number of Filters	10
Filter Size	3
Optimizer	ADAM
Number of Mini-batches	64
Activation function for Hidden layers	RELU
Activation Function of the output layer	SOFTMAX
Total Number of Input Layers	4
Total Number of Output Layers	4
Learning rate	0.001

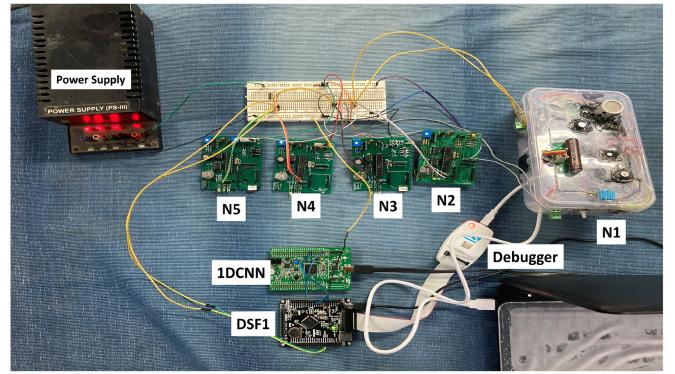


Fig. 4. Single-section hardware prototype.

The complete description of the model architecture is given in Table III. The hyperparameter of the 1DCNN model is obtained using the random search method. The output from the DSF of each section is fed to the IB of the 1DCNN. IB stores each DSF output sampled every second for 1 min. Since there are three sensors per node, there are 180 data points per sample for a section in the IB. After that, the data sample is arranged into a 60×3 shape where 60 represents the data samples of every second and 3 illustrates the number of sensors.

The prepared data sample from every section is fed to the 1DCNN input layer. The automatic feature extraction is done using 1DCNN and 1D max-pooling layers using a convolution process across the 60 samples in each channel. Then, based on the extracted features, the decision is made using fully connected layers.

IV. RESULT AND DISCUSSION

The proposed EGHP model mentioned in Fig. 2 has been designed and tested for a single longwall section case. The hardware setup is shown in Fig. 4. The setup consists of five sensor nodes, i.e., N1, N2, N3, N4, and N5, where N1 and N2 are test nodes, and the rest are healthy nodes. Test node N1 is an expanded version of other nodes with the provision of turning on and off of its all sensors separately while testing. The STM32F407 microcontroller is used to implement DSF1, and the 1DCNN model is deployed on STM32F411. Both the controllers are from the ARM CORTEX M4 family with AI support. For 1DCNN model implementation, Cube

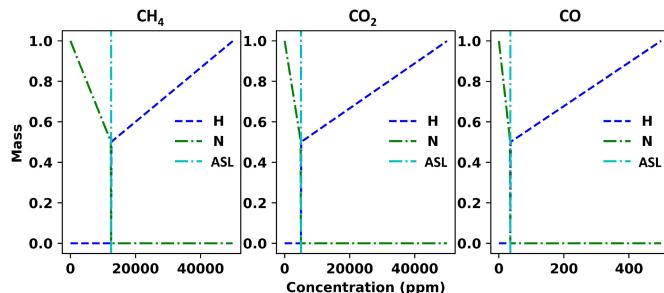


Fig. 5. Variation of mass for various levels of gas concentration.

MX, a proprietary application from ST Microelectronics, is used. The 1DCNN model is initially trained and validated in a system with Intel Core i7-9750H CPU, 8-GB RAM, and Windows 11 operating system. Additionally, the X-CUBE-AI software package is used to convert the trained 1DCNN model for the microcontroller. The driver for the DSF model for the controller is created using Cube MX IDE. The detailed descriptions of DSF and 1DCNN models are as follows.

A. Performance Analysis of DSF

The DSF is designed based on DSET to filter out the observations of malfunctioning sensors from the rest of the healthy ones. Hence, a dedicated BBA function is created, which depends on the scaled difference between the ASL and sensors observation. The variation of mass for CH_4 , CO_2 , and CO sensors based on the BBA is shown in Fig. 5. The mass variation of these sensors observations for the above and below ASL value depends upon how far the sensors observation is from its respective ASL. The mass for class H below ASL is 0, and as the observation reaches toward ASL, the mass of class N decreases. Beyond the ASL, the mass of class N becomes 0, and the mass of class H starts increasing. The slope of the mass for all sensors beyond the ASL is different because the full-scale range for all the sensors is different. The sensor whose hazard class does not match the combined hazard class is a faulty node, and its observation is filtered out from the rest. But this method cannot detect those malfunctioning sensors whose hazard class matches the combined hazard class, but the reading is significantly different from the healthy ones. Hence, a support vector is created that measures the similarity among the sensor nodes whose hazard class matches the combined hazard class. The variation of the support value between a node pair is given in Fig. 6. S_x is the support value for a test data sample with CH_4 test, CO_2 test, and CO test as 3500, 2000, and 16 ppm. The value of S_x is maximum when ppm matches those test values, i.e., the masses related to those concentrations are equal to masses of the test concentrations of gases. The node pairs with similar masses are identical. The analysis of various samples indicated that the node pairs with S_x values more than 0.97 are identical pairs. The pairs whose S_x value is less than 0.97 are considered faulty nodes because their mass does not match with the majority of the nodes. Afterward, the mean of the observations of healthy nodes is transferred to the 1DCNN controller.

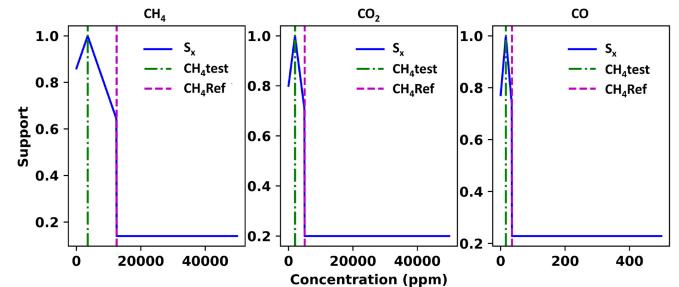


Fig. 6. Variation of support value for a node pair.

TABLE IV
COMPARISON OF CLASSIFIERS

Model	Accuracy	Precision	Recall	F1-Score	Number of Trainable Parameters
ANN	0.991	0.991	0.991	0.991	22,688
SVM	0.985	0.986	0.985	0.985	22,994
RF	0.989	0.989	0.989	0.989	-
1DCNN	0.996	0.996	0.996	0.996	1,518

TABLE V
ROBUSTNESS EVALUATION OF CLASSIFIERS

Model	Classification accuracy at		
	10% anomaly	20% anomaly	30% anomaly
ANN	83.2	68.2	65.6
SVM	78.0	67.9	65.3
RF	81.4	68.5	65.9
1DCNN	97.6	80	75.2

B. Performance Analysis of 1DCNN

To select a classifier for layer 2 of the EGHP, SVM, ANN, RF, and 1DCNN are compared based on their accuracy, precision, recall, and F1-Scores. Since the model is dedicated to edgel ML, the number of trainable parameters also plays an essential role due to the limited memory of the microcontrollers. The comparison report of these models is given in Table IV.

Although the difference between these models' accuracy, precision, recall, and F1-Score is minimal, 1DCNN has the highest scores among all the models and the lowest number of trainable parameters among ANN and SVM-based models. Additionally, to test the robustness of these models, the actual test samples are manipulated with some anomalous data points. Since each test sample consists of 60 data points, three categories are defined where in the first category, 10% of these 60 data points are replaced with anomalous data points. Similarly, in the second and the third categories, 20% and 30% of these 60 data points are replaced with anomalous data points. The classification performance of the proposed model is evaluated with the other classifiers using the above-mentioned test data set. The comparison of classification accuracy for the three categories is given in Table V.

The classification performance of the 1DCNN-based classifier is found to be superior to the other models. The additional benefit is its relatively very low number of trainable parameters than the other classifier which reduces the memory requirements of the model and suitably fits with the EML

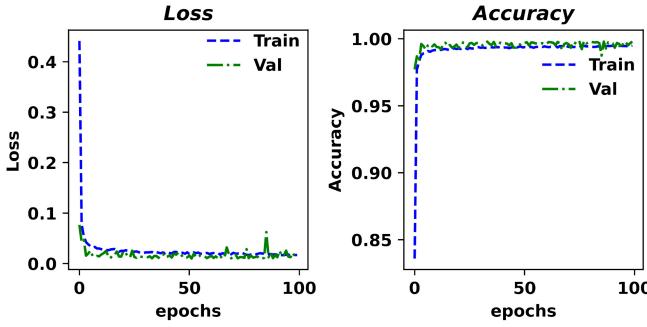


Fig. 7. Training and validation analysis of the 1DCNN model.

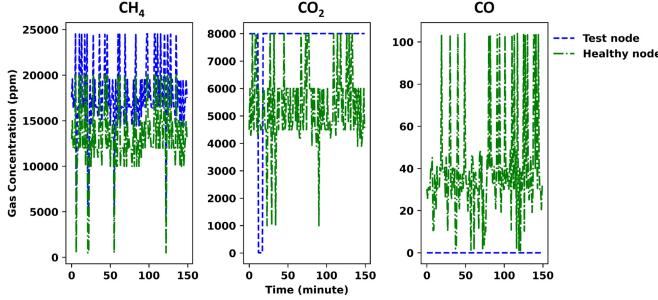


Fig. 8. Test node and healthy node data samples.

requirements. Hence, 1DCNN is selected as a classifier for the second layer of the EGHP. Before deploying the 1DCNN mode into the STM32F411, it is trained on UCM data samples for the hazard class mentioned in Table II. 1DCNN loss and accuracy for training and validation data samples are shown in Fig. 7. The training (Train), validation (Val), and test losses after 100 epochs are 0.015, 0.0148, and 0.0145, respectively. These losses are well below the allowable range of losses, indicating that the model is optimized for variance and bias-related issues for the hyperparameters presented in Table III. The proposed model's Train, Val, and Test accuracies are 0.994, 0.997, and 0.997, respectively. Then, the trained model is analyzed in the STM32 Cube MX application for its deployment in the controller. Although the 1DCNN classifier has a 0.99 accuracy score, it suffers from malfunctioning sensors observation. A detailed explanation of EGHP with DSF and without DSF for faulty node scenarios is presented in the next section.

C. Performance of EGHP on the Faulty Node and Healthy Node Samples

The proposed EGHP model is tested for different faulty node scenarios. The faulty node data are sampled using test nodes N1 and N2 shown in Fig. 4. The data samples for the analysis are shown in Fig. 8. There are five different cases of the faulty node considered in this analysis. Case 0 represents all five nodes are healthy. Cases 1–3 represent the CH_4 sensor malfunction in N1, the CO_2 sensor malfunction in N2, and the CO sensor malfunction in N1 nodes. Case 4 is for malfunction of the CH_4 sensor in N1 and the CO_2 sensor in N2. Case 5 indicates the malfunctioning of two methane sensors in N1 and N2. Accuracy, precision, recall, and F1-Score of the EGHP model without DSF are given in Table VI. A

TABLE VI
PERFORMANCE OF EGHP WITHOUT DSF

Case	Accuracy	Precision	Recall	F1-Score
Healthy	0.99	0.99	0.99	0.99
1	0.64	0.57	0.70	0.60
2	0.62	0.64	0.62	0.54
3	0.74	0.73	0.71	0.66
4	0.36	0.29	0.41	0.30
5	0.60	0.54	0.65	0.55

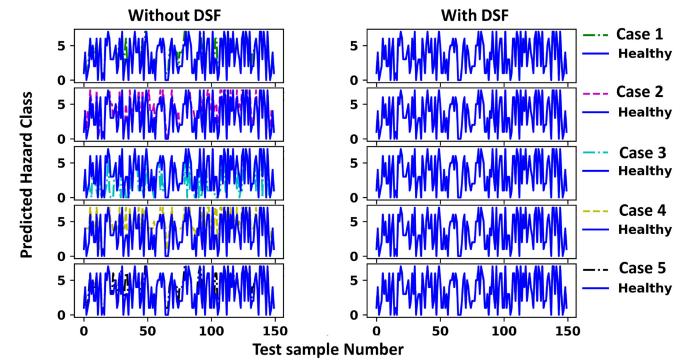


Fig. 9. Prediction analysis of EGHP.

total of 150 data samples for the cases mentioned in Table VI are tested on the EGHP model without DSF, i.e., only using 1DCNN as a classifier. The analysis report in Table VI has shown that for a faulty node scenario, the model's classification accuracy has degraded from 0.99 to a minimum of 0.36, which cannot be accepted in a practical system. The other scenario is when 1DCNN is used in conjunction with DSF. In that case, the accuracy, precision, recall, and F1-Score of all the faulty node cases match with the healthy node case because the observations of those faulty nodes are removed from the averaging process. The prediction outcome of all cases mentioned in Table VI using DSF and without using DSF is shown in Fig. 9. The figure shows that predictions of faulty node cases deviate significantly from the predictions of the healthy node when DSF is not present. But with the application of DSF, the deviations are negligible.

D. Hardware Emulation of the Proposed Model

The EGHP model is emulated on the hardware shown in Fig. 4. In this emulation, the study of a single section of the longwall panel is considered, as shown in Fig. 1. The sensor-node array consists of five sensor nodes having three sensors in each node for CH_4 , CO_2 , and CO gases. DSF and 1DCNN are implemented on separate microcontrollers. The DSF driver for the controller is developed using (1)–(17). The output of the DSF at any instant can be seen on the Live Expression panel of STM32Cube IDE, a proprietary application for STM32 microcontrollers. The real-time output of DSF at any moment is shown in Fig. 10.

Dat1, Dat2, Dat3, Dat4, and Dat5 are the IBs to store the reading of the nodes N1–N5. Ch4mean1, Co2mean1, and Comean1 store the output of DSF. All IBs are an array of size 3×60 stores CH_4 sensor data in the first index, CO_2 data in the second index, and CO sensor data in the last index.

Breakpoints Expressions Modules Registers Live Expressions SFRs		
Expression	Type	Value
↳ Dat1	int [3]	[3]
↳ Dat1[0]	int	13094
↳ Dat1[1]	int	3094
↳ Dat1[2]	int	1397
↳ Dat2	int [3]	[3]
↳ Dat2[0]	int	30001
↳ Dat2[1]	int	2000
↳ Dat2[2]	int	2
↳ Dat3	int [3]	[3]
↳ Dat3[0]	int	3000
↳ Dat3[1]	int	2000
↳ Dat3[2]	int	2000
↳ Dat4	int [3]	[3]
↳ Dat4[0]	int	3000
↳ Dat4[1]	int	2000
↳ Dat4[2]	int	2
↳ Dat5	int [3]	[3]
↳ Dat5[0]	int	3000
↳ Dat5[1]	int	2002
↳ Dat5[2]	int	4
> Fnoadch4	int [5]	[5]
> Fnoadco2	int [5]	[5]
> Fnoadco	int [5]	[5]
↳ Ch4mean1	float	3000
↳ Co2mean1	float	2000
↳ Comean1	float	2
Add new expression		

Fig. 10. DSF output.

Variables Breakpoints Expressions Modules Registers Live Expressions		
Expression	Type	Value
`#= gerr	int	9
`#= gcount	int	171
↳ Dat1	float [180]	[180]
↳ decision	float [8]	[8]
`#= decision[0]	float	0.999912024
`#= decision[1]	float	6.88351065e-005
`#= decision[2]	float	1.4041608e-006
`#= decision[3]	float	1.77863094e-005
`#= decision[4]	float	4.69353498e-019
`#= decision[5]	float	1.29711394e-010
`#= decision[6]	float	3.61624286e-009
`#= decision[7]	float	5.91936232e-022

Fig. 13. Prediction outcome of the 1DCNN controller.

Breakpoints Expressions Modules Registers Live Expressions SFRs		
Expression	Type	Value
`#= Dat4[2]	int	2
↳ Dat5	int [3]	[3]
`#= Dat5[0]	int	3000
`#= Dat5[1]	int	2002
`#= Dat5[2]	int	4
↳ Fnoadch4	int [5]	[5]
`#= Fnoadch4[0]	int	3
`#= Fnoadch4[1]	int	4
`#= Fnoadch4[2]	int	5
`#= Fnoadch4[3]	int	0
`#= Fnoadch4[4]	int	0
↳ Fnoadco2	int [5]	[5]
`#= Fnoadco2[0]	int	2
`#= Fnoadco2[1]	int	3
`#= Fnoadco2[2]	int	4
`#= Fnoadco2[3]	int	5
`#= Fnoadco2[4]	int	0
↳ Fnoadco	int [5]	[5]
`#= Fnoadco[0]	int	2
`#= Fnoadco[1]	int	4
`#= Fnoadco[2]	int	5
`#= Fnoadco[3]	int	0
`#= Fnoadco[4]	int	0
`#= Ch4mean1	float	3000
`#= Co2mean1	float	2000
`#= Comean1	float	2
Add new expression		

Fig. 11. List of healthy nodes identified by DSF.

Complexity report per layer - macc=32,568 weights=6,072 act=2,720 ram_ic=752				
id	name	c_macc	c_rom	c_id
1	convld		18.5%	6.6% [0]
2	convld_1		59.0%	20.4% [1]
5	convld_2		19.7%	20.4% [2]
9	dense		2.2%	46.8% [3]
9	dense_nl		0.0%	0.0% [4]
10	dense_1		0.3%	5.8% [5]
10	dense_1_nl		0.4%	0.0% [6]

Fig. 12. 1DCNN model conversion summary.

The figure indicates that Dat1 and Dat2 readings for the CH_4 sensor deviate significantly from the other methane sensor's observations. The readings of the CO sensor for Dat1 and Dat3 differ from the rest. But the DSF output, which is the mean of healthy sensor reading, i.e., the values of Ch4mean1, Co2mean1, and Comean1 are almost similar to the healthy sensor readings. The list of the healthy nodes for each sensor is stored in Fnoadch4, Fnoadco2, and Fnoadco, as shown in Fig. 11. Then, the output of DSF is transferred to the 1DCNN model using serial communication between the DSF controller and the 1DCNN controller. The 1DCNN driver for the controller is developed and trained in Keras, and then the trained model is saved in an H5 format. Then, in Cube MX, the saved model is converted as a 1DCNN driver for the mentioned microcontroller. The model conversion summary is shown in Fig. 12. The prediction result of the 1DCNN controller is

shown in Fig. 13. The decision array stores the probability of all eight hazard classes. For the SA observation mentioned in Fig. 10, the predicted hazard class is 0 because the prediction probability of this class is highest among all other classes. Similarly, the other sections of the longwall can be analyzed using the proposed EGHP model with a dedicated DSF and sensor-node array for each section.

V. CONCLUSION

A multisensor fusion-based fault-tolerant gas hazard prediction model dedicated to the UG longwall operation is proposed. Among the two levels of EGHP, the first level incorporates DSF for each section of the longwall. The DSF is found to be an efficient model for filtering the malfunctioning sensors readings from the rest of the observations. For the second level of the fusion model, 1DCNN is selected, which classifies the hazard class for the entire face width. The comparison of the 1DCNN and other existing classifiers has shown that the performance of 1DCNN based on the various metrics scores is superior to existing classifiers. In the investigation, it is observed that the performance of 1DCNN deteriorates on the occurrence of faulty node scenarios. But, 1DCNN with DSF has predicted the hazard class with a classification accuracy of 99.6%. The proposed EGHP model is implemented on the ARM CORTEX M4 microcontroller to verify its application for EML. The model conversion summary generated by the Cube MX has shown that the macc and memory requirements of the proposed model are within the acceptable limit of the controller. The EML-based implementation of the proposed model ensures data security and reduced network latency offered by WSN.

Although the proposed model has various advantages over the traditional approaches for gas hazard prediction in the UCM longwall operations, this study is limited to the spatial analysis of the air of a longwall panel. The temporal analysis of the gas concentrations requires a more complex DNN network, increasing the computational cost. Therefore, the implementation of both spatial and temporal analysis of the gas concentrations, including DSET using edge devices requires additional research in this regard.

REFERENCES

- [1] M. Tutak, J. Brodny, D. Szurgacz, L. Sobik, and S. Zhironkin, "The impact of the ventilation system on the methane release hazard and spontaneous combustion of coal in the area of exploitation—A case study," *Energies*, vol. 13, no. 18, p. 4891, 2020. [Online]. Available: <https://www.mdpi.com/1996-1073/13/18/4891>
- [2] S. Schatzel, R. B. Krog, and H. Dougherty, "Methane emissions and airflow patterns on a longwall face: Potential influences from longwall gob permeability distributions on a bleederless longwall panel," *Trans. Soc. Min. Metallurgy Exploration*, vol. 342, no. 1, pp. 51–61, 2017.
- [3] M. Tutak and J. Brodny, "Predicting methane concentration in longwall regions using artificial neural networks," *Int. J. Environ. Res. Public Health*, vol. 16, no. 8, p. 1406, 2019.
- [4] F. Xiao, "Evidence combination based on prospect theory for multi-sensor data fusion," *ISA Trans.*, vol. 106, pp. 253–261, Nov. 2020.
- [5] L. A. Hunt, "Missing data imputation and its effect on the accuracy of classification," in *Data Science*. Cham, Switzerland: Springer, 2017, pp. 3–14.
- [6] J. Azar, A. Makhoul, M. Barhamgi, and R. Couturier, "An energy efficient IoT data compression approach for edge machine learning," *Future Gener. Comput. Syst.*, vol. 96, pp. 168–175, Jul. 2019.
- [7] B. B. Beamish and P. J. Croisdale, "Instantaneous outbursts in underground coal mines: An overview and association with coal type," *Int. J. Coal Geol.*, vol. 35, nos. 1–4, pp. 27–55, 1998.
- [8] D. J. Black, "Review of coal and gas outburst in Australian underground coal mines," *Int. J. Min. Sci. Technol.*, vol. 29, no. 6, pp. 815–824, 2019.
- [9] L. Muduli, P. K. Jana, and D. P. Mishra, "Wireless sensor network based fire monitoring in underground coal mines: A fuzzy logic approach," *Process Safety Environ. Protect.*, vol. 113, pp. 435–447, Jan. 2018.
- [10] E. Danish and M. Onder, "Application of fuzzy logic for predicting of mine fire in underground coal mine," *Safety Health Work*, vol. 11, no. 3, pp. 322–334, 2020.
- [11] B. W. Jo and R. M. A. Khan, "An event reporting and early-warning safety system based on the Internet of Things for underground coal mines: A case study," *Appl. Sci.*, vol. 7, no. 9, p. 925, 2017.
- [12] X. Wu, Z. Zhao, and L. Wang, "Deep belief network based coal mine methane sensor data classification," *J. Phys. Conf. Series*, vol. 1302, no. 3, 2019, Art. no. 32013.
- [13] S.-G. Cao, Y.-B. Liu, and Y.-P. Wang, "A forecasting and forewarning model for methane hazard in working face of coal mine based on LS-SVM," *J. China Univ. Min. Technol.*, vol. 18, no. 2, pp. 172–176, 2008.
- [14] P. Dey, S. K. Chaulya, and S. Kumar, "Hybrid CNN-LSTM and IoT-based coal mine hazards monitoring and prediction system," *Process Safety Environ. Protect.*, vol. 152, pp. 249–263, Aug. 2021.
- [15] N. Nesa and I. Banerjee, "IoT-based sensor data fusion for occupancy sensing using Dempster-Shafer evidence theory for smart buildings," *IEEE Internet Things J.*, vol. 4, no. 5, pp. 1563–1570, Oct. 2017.
- [16] I. Ullah, J. Youn, and Y.-H. Han, "Multisensor data fusion based on modified belief entropy in Dempster-Shafer theory for smart environment," *IEEE Access*, vol. 9, pp. 37813–37822, 2021.
- [17] N. Ghosh, R. Paul, S. Maity, K. Maity, and S. Saha, "Fault matters: Sensor data fusion for detection of faults using Dempster-Shafer theory of evidence in IoT-based applications," *Expert Syst. Appl.*, vol. 162, Dec. 2020, Art. no. 113887.
- [18] A. K. Singh and J. Kumar, "Fugitive methane emissions from Indian coal mining and handling activities: Estimates, mitigation and opportunities for its utilization to generate clean energy," *Energy Procedia*, vol. 90, pp. 336–348, Dec. 2016.
- [19] M. Kozielski, M. Sikora, and Ł. Wróbel, "Data on methane concentration collected by underground coal mine sensors," *Data Brief*, vol. 39, Dec. 2021, Art. no. 107457.
- [20] N. Szlazak, D. Obracaj, and J. Swolkien, "Enhancing safety in the Polish high-methane coal mines: An overview," *Min. Metallurgy Exploration*, vol. 37, no. 2, pp. 567–579, 2020.
- [21] A. Meshkov, O. Kazanin, and A. Sidorenko, "Methane emission control at the high-productive longwall panels of the Yavlevsky coal mine," in *Proc. E3S Web Conf.*, vol. 174, 2020, p. 1040.
- [22] P. Thakur, *Advanced Mine Ventilation: Respirable Coal Dust, Combustible Gas and Mine Fire Control*. Cambridge, MA, USA: Woodhead Publ., 2018.



Mayank Sharma (Member, IEEE) received the B.Tech. degree in electrical and electronics engineering from Lovely Professional University, Phagwara, India, in 2014, and the M.Tech. degree in mine electrical engineering from the Indian Institute of Technology, Dhanbad, India, in 2017, where he is currently pursuing the Ph.D. degree with the Department of Electrical Engineering.



Tanmoy Maity (Member, IEEE) received the B.E. and M.E. degrees in electrical engineering from Calcutta University, Kolkata, India, in 1990 and 1994, respectively, and the Ph.D. degree from Bengal Engineering & Science University at Shibpur, Howrah, India, in 2008.

He has six years of industrial and more than 22 years of academic experience. He is currently working as an Associate Professor with the Department of Electrical Engineering, Indian Institute of Technology, Dhanbad, India.