

Smart and Fault-Tolerant Multisensor Fusion Model for UCM Methane Hazard Monitoring Based on Belief Divergence Backed DS Filter and Hybrid CNN-LSTM Classifier

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

Abstract—The underground coal mine (UCM) dynamic and complex environment impose various hazards that significantly affect the mining production and safety of the personnel. Flammable and poisonous gases significantly contribute to many fatal accidents. This study proposes a real-time-based reliable gas hazard monitoring system using multisensor data fusion. A hybrid of CNN-LSTM-based deep neural network (HCLM) is developed to serve the purpose. Due to the challenging environment of the UCM, sensor malfunctioning is inevitable and severely affects the performance of HCLM. A novel front-end filter (FEF) is developed based on Damper Shafer's (DS) theory and belief divergence-based weighted credibility metric to overcome the drawbacks of HCLM. In the laboratory trial, it is observed that the hazard classification accuracy of HCLM for the faulty node scenarios is 85%. In contrast, the accuracy of the HCLM integrated with FEF is maintained at 98%, even for multiple faulty node cases. Another novelty of this study is the tinyML implementation of the proposed model. Due to UCM's inherent complexities and challenges, traditional wireless communications face operational difficulties. Hence, a cloud-based machine learning operation is not a feasible option in UCM. Hence, using the concept of tinyML, the proposed model is directly deployed on a microcontroller near the data sources, thereby reducing network latency and security issues.

Index Terms—CNN-LSTM, Dempster–Shafer evidence theory (DSET), sensor fusion, TinyML, underground coal mine (UCM).

I. INTRODUCTION

THE PRESENCE of flammable and hazardous gases in underground coal mine (UCM) needs a reliable, efficient monitoring system for safe mining operations. Especially, in the case of the longwall mining operations, the higher capacity mining equipment resulted in an increased production rate and methane emission from the coal face [1]. Along with explosive gas like methane, other poisonous gases, e.g., carbon monoxide (CO), carbon dioxide (CO₂), and hydrogen sulphide (H₂S) can be found in the UCM environment [2]. If the concentration of methane gas in the coal mine environment reaches a flammable range [3], the consequent explosion may result in

Manuscript received 12 December 2022; accepted 12 July 2023. Date of publication 17 July 2023; date of current version 8 January 2024. (Corresponding author: Mayank Sharma.)

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

Digital Object Identifier 10.1109/JIOT.2023.3295823

fatal accidents. Apart from these gases, the temperature and humidity significantly affect the UCM environment.

The existing real-time gas monitoring solutions consist of portable gas monitors or a network of fixed sensor nodes mounted throughout the strategic locations inside the mines. Studies like those in [4] utilize fixed real-time gas monitors at the dedicated locations in the longwall coal mines operations for continuous methane observation. Apart from the operational simplicity, these monitors cannot recognize the spatiotemporal patterns associated with the hazardous and non-hazardous gas classes. Moreover, the studies in [4] and [5] reported an erroneous estimate of methane gas concentration due to the sensor malfunction.

The complex and dynamic environment does not support physical models like Kalman filters for gas hazard estimation and identification in the UCM. Therefore, most of the studies emphasize data-driven models. Fuzzy, Bayesian and other statistical models are most popular in this field [6], [7]. However, the dependency of these models on explicit rule base or conditional probability distribution may result in suboptimal outcomes if not perfectly designed. Moreover, these shortcomings can be avoided by using the auto-learning capabilities of a neural network (NN).

A study [8] proposed an artificial NN (ANN)-based methane (CH₄) dispersion in UCM. An ANN model with three hidden layers that can accurately predict methane concentration. Article [9] proposed an ANN-based approach for methane forecasting using the distributed sensor network. Similar studies can be observed in [10] and [11]. Both studies use variants of support vector machines (SVMs). The main motive of the study in [10] was to predict the severity of coal mine gas accidents, whereas in [11], the objective was to predict the hazard associated with continuous coal combustion (CSC). However, the ANN and SVM do not consider the time-series features, an essential quality of any sensor-based intelligent prediction system. This drawback can be efficiently handled using a recursive NN (RNN). Kumari et al. [12] utilized the advantages of a long short-term memory (LSTM)-based RNN network to forecast the gas concentrations. Studies in [13] and [14] give an idea to utilize the advantage of the hybrid of CNN and LSTM in the field of gas hazard prediction in UCM. Prince and Hati [13] predicted ventilation airflow using the CNN-LSTM hybrid model. In the study, comparing the

hybrid model with the existing model showed that the hybrid model outperformed the others. Similar work can also be seen in [14].

The models discussed in these articles are trained on the gas concentration data provided by some kind of sensor installed in the UCM. The real-time sensor observations are not always the same on which the NN model is trained. As already discussed, sensor malfunctioning is common in challenging and hazardous UCM environments. Therefore, any NN model trained on the healthy sensor node data would result in an erroneous outcome. Therefore, instead of using a single sensor, a fusion of multiple sensors may give the optimum estimation of the real-time gas concentration. Among the various sensor fusion methods, the Dempster–Shafer evidence theory (DSET)-based fusion is the most suitable option because of its sensor uncertainty handling capability and belief combination rule [15].

Apart from the sensor malfunctioning, real-time implementation of the NN models in the UCM is crucial. All the above-discussed studies emphasize the development of an NN-based model for hazard classification or forecast, but none have dealt with the real-time implementation of their proposed models. The reduced reliability of the wired communication system due to frequent cable damages prohibits the implementation of the NN models in the remotely installed systems. The dielectric property of coal and tunnel geometry increases the attenuation in the wireless signals, thereby limiting wireless communication to smaller distances. Therefore, the cloud-based NN model implementation is not suitable for the UCM.

Therefore, the real-time realization of the NN models for the gas hazard prediction in the coal mines is limited by the uncertainties and malfunctioning observations from the sensor nodes and inherent complexities and challenges toward the communication within a sensor network. This study proposes a real-time hybrid CNN-LSTM model (HCLM)-based methane hazard classifier for the UCM using multimodal time-series data fusion. The main attributes used in this study are CH₄ gas, temperature (Tem), and humidity (Hum). The reason for selecting Tem and Hum with CH₄ gas is because these two parameters significantly affect the ignition characteristic of the CH₄ gas [16], [17]. A DSET-motivated front-end filter (FEF) is designed to deal with uncertainties and malfunctioning sensors in real-time scenarios. For FEF, a prior study of this work in [18] and [19] is taken into account. The FEF, compared to the credibility in [19], is an evidential divergence-based credibility metric that can provide improved visibility between a faulty node and a healthy node.

Additionally, a TinyML-enabled standalone gas hazard prediction system is proposed to overcome existing NN models' remote and cloud computation challenges. The whole fusion model is realized into a 32 bit, battery-operated microcontroller, capable of doing all FEF and NN-model computation in the actual site, thereby reducing the requirement of remote and cloud-based computation. This in-situ implementation of the proposed model reduces the data latency and security issues and cannot be affected by the wireless

transmission challenges in the UCM. The novelties of this study are as follows.

- 1) Integration of novel FEF unit with hybrid CNN-LSTM model to improve the system's reliability in uncertain and malicious sensor node scenarios.
- 2) An evidential divergence-based credibility metric is developed using DSET.
- 3) Conversion of the HCLM to the TFLite model using complete integer quantization technique to realize TinyML.

The real-time realization of the TinyML or edge machine learning integrated with FEF, capable of handling malfunctioning sensor observation, is the first of its kind in a UCM scenario.

II. METHODOLOGY

The distributed wireless sensor nodes for UCM environment monitoring are clustered into groups based on their distances from each other. Hence, each sensor node cluster observes a specific portion of the mine. The sensor nodes within a cluster are sensing the same environment portion of the UCM. Each cluster has its FEF and a central HCLM unit. The FEF filters the malfunctioning sensor node observation within a cluster and then transmits the mean of the healthy observations of that cluster to the HCLM unit. Then, the HCLM unit uses the fusion of the CH₄, Tem, and Hum-based time-series extracted features using the hybrid of time-distributed 1-D-CNN layers followed by LSTM layers. Then, The extracted features are fed to the fully connected (FC) layers for hazard prediction. The block diagram of the proposed model is shown in Fig. 1.

This study's experimental validation is done for a single cluster consisting of five sensor nodes. The description of the experimental prototype, FEF and HCLM are as follows.

A. Description of the Experimental Prototype

The prototype consists of five sensor nodes, FEF and HCLM units. Each sensor node consists of one methane sensor (TGS-6810-DOO-catalytic combustion type) and a digital temperature and humidity sensor (DHT-22). The integrated FEF-HCLM unit is deployed into an ARM Cortex M4-based microcontroller. The complete experimental setup is shown in Fig. 2. All the nodes, i.e., N1, N2, N3, N4, and N5, within the cluster are identical except for N1 and N5. Both the nodes contain faulty methane, temperature and humidity sensors. Fig. 2 also shows the STM32F411 microcontroller used for the FEF and HCLM.

B. Description of the Data

For the training and validation of the HCLM model, a publicly available data set mentioned in the Mendeley data repository [20] is used. The data set contained methane concentrations, temperature (Tem), and humidity (Hum) with their timestamps. The test set data samples are synthesized in a laboratory environment using the prototype mentioned in Fig. 2. Data frames, each of 60-s samples, are extracted from the data mentioned above. There is a total of 64 720 data frames extracted for CH₄, Tem, and Hum. Then, every five consecutive data frames are considered as time-distributed

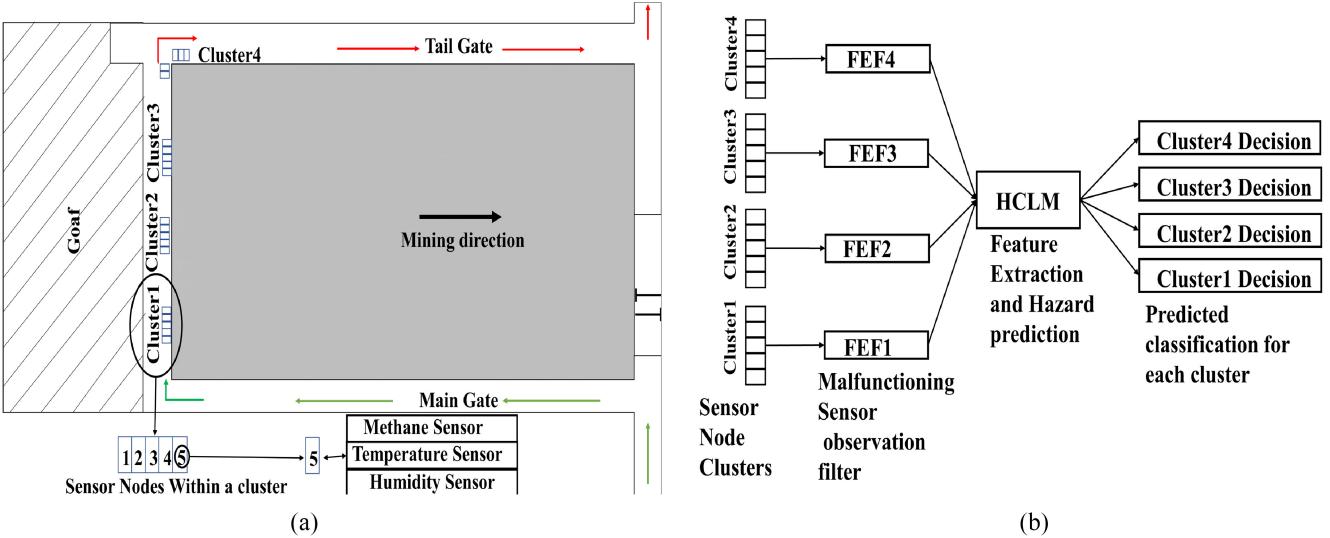


Fig. 1. Gas monitoring system. (a) Sensor cluster in UCM. (b) Proposed hazard prediction model.

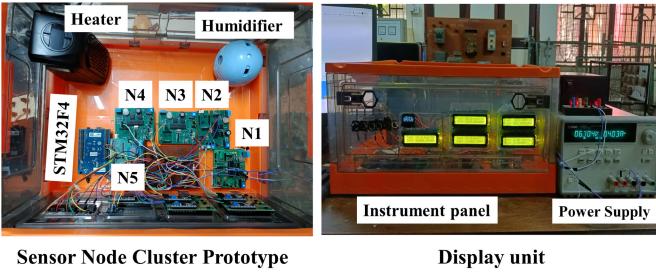


Fig. 2. Prototype model.

TABLE I
HAZARD CLASS

Class	Hazard level	Condition
1	Very Low	No hazard
2	Low	Hazardous temperature or Humidity-based event detected
3	High	Hazardous methane concentration-based event detected
4	Very High	Event of all abnormalities

input. Therefore, the input shape of the HCLM model is like (Batch_Size, Time-stamps, and Features). Four hazard classes are evaluated based on the input attributes mentioned in Table I.

The abnormal event is considered based on the statistical features of the attributes for 5-min period, i.e., based on 300-s samples. The temporal locations of the first (Q1), second (Q2), and third (Q3) quartile within a 60-s data sample can identify the hazardous and nonhazardous pattern. Those 60-s samples, in which either Q3 and Q2 or Q3 and Q1 surpass the threshold [21], are considered hazardous events. If any hazardous event occurs, it is found that the temporal location of Q3 and Q2 comes after Q1. Otherwise, in the case of Q3 and Q1, the temporal location of Q1 is before Q3 and Q1. Such hazardous events are observed for five such 60-s windows, and the mean of hazardous events over 5 min is calculated. The 5-min window of 300 samples is leveled

as hazardous if the mean observed hazardous event is more than 60%. As discussed above, the hazardous or nonhazardous events for methane, temperature, and humidity are identified. The leveling mentioned in Table I is based on identifying hazardous events related to the CH₄, temperature and humidity. Z-score normalization is used to normalize the data samples, and the normalized data set is split into train, test and validation sets. The 45 304 samples are used as the training set, and 9708 samples are used for each validation and test set.

C. FEF Unit

The FEF unit is responsible for filtering the malfunctioning sensor observations from the rest of the healthy observations, reducing uncertainty. It becomes significant when there is a high chance of sensor malfunctioning, like in UCM's harsh environment. In this study, Damper Shafer's (DS) is adopted to create a weighted credibility metric. The Euclidean distance is used to decide whether evidence from any sensor node is credible or not. Additional weight is applied to the evidence of nodes based on their divergence. The main components of DS are the frame of discernment (FOD), basic belief assignment (BBA) function, and evidence combination rule. The details of these terminologies are given in the works of [15], [18], and [22]. In this study, to design a dedicated FEF unit, a novel BBA function is created. The detailed explanation of all the components in the context of this study is as follows.

1) *FOD*: It is a set of all possible hypotheses in a sensor fusion operation. The hypothesis for CH₄ sensors in a cluster is defined as a hazardous (*H*) concentration level or nonhazardous (*N*). So, FOD for modality *x* is represented as follows:

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

where *x* represents CH₄, Tem, and Hum. The power set consists of all possible subsets of the elements in the FOD. It can be represented as follows:

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

The next step is to find the masses associated with each element of the POW set that can be obtained using the dedicated BBA function defined as follows.

2) *BBA*: It maps the sensor output to the masses of elements of the POW set. The mass value ranges from 0 to 1, i.e., if the sensor observation highly supports the “*H*” hypothesis, then its belief would be more toward 1 and vice-versa. Hence, the BBA function can be represented as follows:

$$m:e \rightarrow [0, 1] \forall e \in POW_{CH_4}. \quad (3)$$

In (3), “*m*” represents the BBA function and “*e*” represents the elements of the POW set. The BBA functions for methane, temperature, and humidity are as follows:

$$m_x\{H\} = \begin{cases} 0, & MI_x < TH^L \\ 0.25, & TH^L \leq MI_x \leq TH^R \\ a1_x MI_x + b1_x, & MI_x > TH^R \end{cases} \quad (4)$$

$$m_x\{N\} = \begin{cases} a2_x MI_x + b2_x, & MI_x < TH^L \\ 0.25, & TH^L \leq MI_x \leq TH^R \\ 0, & MI_x > TH^R \end{cases} \quad (5)$$

$$m_x\{H, N\} = \begin{cases} 0.5, & TH^L \leq MI_x \leq TH^R \\ 0.25, & \text{otherwise.} \end{cases} \quad (6)$$

In (4)–(6), TH^L is the lower limit of the threshold and TH^R is the upper limit of the threshold, MI_x is the scaled input of modality *x*, defined as follows:

$$MI_x = \frac{D_x}{TH_x} \quad (7)$$

where D_x is the actual value of modality *x*, TH_x is the threshold value of modality *x*. The temperature and humidity values of more than 36 °C, 90%, resulting in a higher physiological strain index (PSI). However, a reduction in humidity and an increase in temperature results in a reduction in the lower explosive limit (LEL) of CH₄ [23], [24]. Hence, the TH_x for Tem and Hum is set to 35 °C and 80%. The LEL for CH₄ gas is 5%, but with the mixture of other gases and increased ambient temperature, it can be reduced from this level, so as per the guideline of the director general of mines safety (DGMS), India, a value of 12 500 ppm is considered as the threshold value beyond which it can be considered hazardous. The TH^L and TH^R values are decided based on the uncertainty levels of the sensor. A maximum of 5% uncertainty in the sensor measurement is considered, so TH^L and TH^R are as follows:

$$TH^L = \hat{M}_x(1 - 0.05) \quad (8)$$

$$TH^R = \hat{M}_x(1 + 0.05) \quad (9)$$

where \hat{M}_x is the scaled input at $D_x = TH_x$, i.e., \hat{M}_x will be 1 for each modality *x*. The parameters $a1_x$, $b1_x$, $a2_x$, and $b2_x$ are derived as follows:

$$a1_x = \frac{m_{a1x}\{H\} - m_{a2x}\{H\}}{MI_{a1x} - MI_{a2x}} \quad (10)$$

$$b1_x = m_{a2x}\{H\} - a1_x MI_{a2x} \quad (11)$$

$$a2_x = \frac{m_{b1x}\{N\} - m_{b2x}\{N\}}{MI_{b1x} - MI_{b2x}} \quad (12)$$

$$b2_x = m_{b2x}\{N\} - a1_x MI_{b2x}. \quad (13)$$

The linear model parameters $a1_x$, $b1_x$, $a2_x$, and $b2_x$ in (10)–(13) are the slope and intercept of the BBA function

TABLE II
LINEAR MODEL PARAMETERS

Below TH^L				
Modality (x)	MI_{b1}	$m_{b1}\{N\}$	MI_{b2}	$m_{b2}\{N\}$
CH4	0.95	0.25	0	1
Tem	0.95	0.25	0	1
Hum	0.95	0.25	0	1
Above TH^R				
Modality (x)	MI_{a1}	$m_{a1}\{H\}$	MI_{a2}	$m_{a2}\{H\}$
CH4	1.05	0.25	4	1
Tem	1.05	0.25	2.85	1
Hum	1.05	0.25	1.25	1

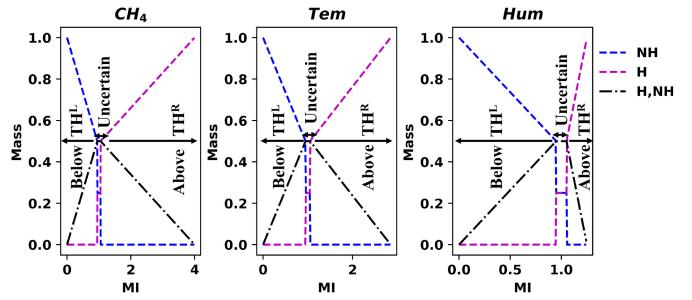


Fig. 3. BBA characteristics.

above TH^R and below TH^L . The required parameters in (10)–(13) are mentioned in Table II. In Table II, m_{b1} is an intercept of the BBA function at TH^L , i.e., 0.95, m_{b2} is an intercept at minimum sensor input value (0). Similarly, m_{a1} is intercepting at TH^R , whereas m_{a2} is intercepting at full scale of the modalities, i.e., for the case of CH₄, full scale is 100% LEL that is 50 000 ppm, so MI_{a2} at this point will be 4 and, hence, m_{a2} is intercepting at 4.

Based on the previous discussions a mass matrix for the five nodes is developed in

$$M_x = \begin{bmatrix} m_1\{H\}_N & m_1\{N\}_N & m_1\{H, N\}_N \\ m_2\{H\}_N & m_2\{N\}_N & m_2\{H, N\}_N \\ m_3\{H\}_N & m_3\{N\}_N & m_3\{H, N\}_N \\ m_4\{H\}_N & m_4\{N\}_N & m_4\{H, N\}_N \\ m_5\{H\}_N & m_5\{N\}_N & m_5\{H, N\}_N \end{bmatrix} \quad (14)$$

where M_x is mass matrices, m_1 – m_5 represents the masses of nodes 1–5 and *H* and *N* indicate hazardous and nonhazardous class at the individual sensor level. Similarly, the M_{Tem} and M_{Hum} are also created. The characteristics of the BBA for all three modalities are shown in Fig. 3. In Fig. 3, the BBA function for all modalities indicates unique characteristics for three different regions, i.e., below TH^L , within TH^L and TH^R , and above TH^R . Below TH^L , the mass of hypothesis *N* is always greater than the uncertainty ($\{H, N\}$), and the mass of hypothesis *H* remains zero. In the region beyond TH^R , the mass of hypothesis *H* becomes greater than the uncertainty, and the mass of hypothesis *N* remains zero. However, within TH^L and TH^R , both hypothesis mass becomes smaller than the uncertainty.

3) *Malfunctioning Node Detection*: The CH₄, Tem, and Hum mass matrices are then used to calculate the credibility metric for each sensor along each node. The steps involved are as follows.

a) *Belief and plausibility metric determination:* The belief of an element of the POW set is calculated as follows:

$$bel_x\{A\} = \sum_{B \subseteq A \neq \emptyset} m_x\{B\} \quad (15)$$

where A is the hypothesis, i.e., $\{H\}$, $\{N\}$, or $\{H, N\}$, and the plausibility for A is mentioned in the following:

$$pl_x\{A\} = \sum_{B \cap A \neq \emptyset} m_x\{B\} = 1 - bel_x\{\bar{A}\}. \quad (16)$$

Based on (15) and (16) Bel_x and Pl_x metrics are developed, mentioned in the following:

$$\begin{bmatrix} bel_{1x}\{H\} & bel_{1x}\{N\} \\ bel_{2x}\{H\} & bel_{2x}\{N\} \\ bel_{3x}\{H\} & bel_{3x}\{N\} \\ bel_{4x}\{H\} & bel_{4x}\{N\} \\ bel_{5x}\{H\} & bel_{5x}\{N\} \end{bmatrix} \quad (17)$$

$$\begin{bmatrix} pl_{1x}\{H\} & pl_{1x}\{N\} \\ pl_{2x}\{H\} & pl_{2x}\{N\} \\ pl_{3x}\{H\} & pl_{3x}\{N\} \\ pl_{4x}\{H\} & pl_{4x}\{N\} \\ pl_{5x}\{H\} & pl_{5x}\{N\} \end{bmatrix}. \quad (18)$$

Then, dominating hypothesis is identified by comparing the mean of the plausibility for the H and N hypothesis in (18). Using (19) the index of dominating hypotheses is computed

$$Ind_x = \operatorname{argmax} \left\{ \frac{1}{5} \sum_{i=1}^5 pl_{ix}\{H\} \geq 0.75, \frac{1}{5} \sum_{i=1}^5 pl_{ix}\{N\} \geq 0.75 \right\}. \quad (19)$$

In (19), Ind_x is either 0 or 1, indicating whether the masses of H or N hypotheses are dominating. Based on the index value, the belief column of the dominating hypothesis in (17) is selected to design the credibility metric.

b) *Belief distance measurement:* The Bel_x vector represents the belief column of dominating hypothesis as a result of plausibility comparison in the previous step, mentioned in the following:

$$\vec{Bel}_x = \begin{bmatrix} bel_{1x}\{A\} \\ bel_{2x}\{A\} \\ bel_{3x}\{A\} \\ bel_{4x}\{A\} \\ bel_{5x}\{A\} \end{bmatrix} \quad (20)$$

where A is the dominating hypothesis. In developing a credibility metric, one assumption is made, i.e., at least 60% of the nodes within a cluster are healthy. Therefore, the median of \vec{Bel}_x vector elements will always be associated with a healthy node. The deviation of each element of the Bel_x from its median is calculated, as mentioned in the following:

$$Dev_x = |\operatorname{Median}(\vec{Bel}_x) - \vec{Bel}_x|. \quad (21)$$

Any sensor measurement beyond 5% of the actual measurement shows a significantly larger deflection than others in the deviation vector (21). So, the threshold deviations are chosen at the 5% uncertainty level for all modalities and are represented as D_{xTH} .

However, the deviation measurement in (21) may interpret wrong if the observations are near the threshold because of the slope difference between these regions. So, a corrected belief vector is introduced in the following:

$$\vec{Bel}_{ixc} = \begin{cases} a2_xMI_x + b2_x, & Bel_{ix} = 0 \\ a1_xMI_x + b1_x, & Bel_{ix} = 0.25, Ind_x = 1 \\ Bel_{ix}, & \text{otherwise.} \end{cases} \quad (22)$$

Now, the corrected belief vector in (22) is used to calculate the deviation vector in (21).

c) *Belief vector updation:* Plausibility is the measure of the maximum belief of any hypothesis, including its associated uncertainty, so higher uncertainty results in higher plausibility but lesser belief. So, in this step, the beliefs in the Bel_x having a deviation of more than the D_{xTH} are penalized based on the divergence factor between the belief and plausibility, mentioned in the following:

$$W_i = \begin{cases} -\log(1 - \bar{E}_i), & Dev_{ix} > D_{xTH} \\ 1, & \text{otherwise} \end{cases} \quad (23)$$

where \bar{E}_i is the normalized divergence parameter between the plausibility and the belief, calculated as follows:

$$E_i = \frac{1}{2} bel_{ix}\{A\} \ln_2 \left(\frac{bel_{ix}\{A\}}{pl_{ix}\{A\}} \right) + \frac{1}{2} pl_{ix}\{A\} \ln_2 \left(\frac{pl_{ix}\{A\}}{bel_{ix}\{A\}} \right). \quad (24)$$

The normalized E_i is mentioned in the following:

$$\bar{E}_i = \frac{E_i}{\sum_{i=1}^N E_i} \quad (25)$$

In (24), the first term is the divergence between belief and plausibility, and the second term represents the divergence between plausibility and belief. So, the lesser will be the uncertainty, the lesser will be the penalizing weight W_i . Using (20)–(25), a weighted Bel_x is calculated in the following:

$$\vec{Bel}_x = \begin{bmatrix} W_1 bel_{1x}\{A\} \\ W_2 bel_{2x}\{A\} \\ W_3 bel_{3x}\{A\} \\ W_4 bel_{4x}\{A\} \\ W_5 bel_{5x}\{A\} \end{bmatrix}. \quad (26)$$

d) *Credibility metric:* Now, the deviation of penalized Bel_x vector in (27) from its median is calculated as follows:

$$Dev_x = |\operatorname{Median}(\vec{Bel}_x) - \vec{Bel}_x|. \quad (27)$$

Using the deviations in (27), the credibility metric is calculated as follows:

$$Crd = 1 - Dev_x. \quad (28)$$

For more than 5% deviation, the credibility in (28) is found to be less than 0.95, i.e., any sensor observation below this threshold is considered faulty.

The observations of the faulty sensors are removed based on the above discussion, and the mean of the healthy node observations is fed to the HCLM for further classification.

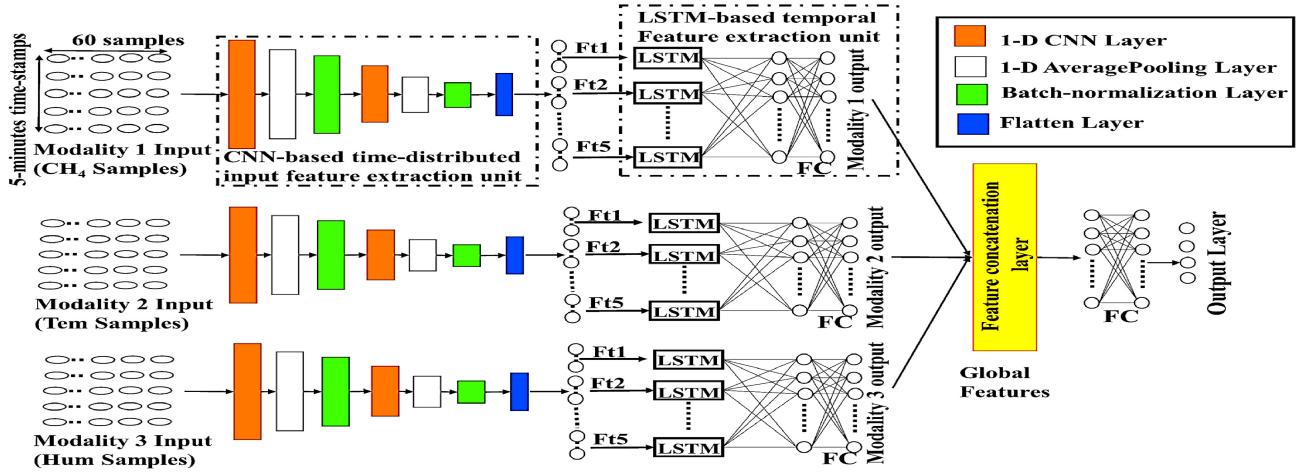


Fig. 4. HCLM architecture.

Additionally, it was observed that a minimum of three sensor nodes within a cluster are required for the optimum sensor fusion process. However, increasing the number of sensor nodes helps to reduce the associated uncertainties by averaging the observations of healthy nodes, thereby increasing accuracy; the maximum number of the sensor node depends on the mining constraints like the type of mining operation, mine geometry, and nature of service.

D. HCLM Unit

In this study, three modalities, i.e., CH₄, Tem, and Hum, are considered to classify the methane-based hazards in the UCM. Different subnetworks are used to extract the time-series features for these modalities. The extracted features from these subnetworks are concatenated and fed to the FC layer for hazard prediction. The architecture of the complete HCLM model is presented in Fig. 4.

In Fig. 4, the input side consists of three modalities, i.e., Modalities 1, 2, and 3 represent CH₄, Tem, and Hum input sequences, respectively. The shape of inputs for these modalities are (m, t, n) , where m is the batch size, t is the number of time stamps, and n is the number of input features. One CNN-based time-distributed feature extraction unit is used for each modality. The CNN units consist of a 1-D CNN layer, 1-D Maxpooling layer, batch normalization layer, and flatten layers. The output of these feature extraction layers is of the shape of (m, t, o) , where o is the output of the flatten layer for each batch and time stamp. The next unit is the LSTM layer follows the FC layers which extract the temporal features from the inputs provided by the previous CNN unit. The outputs of this unit, expressed as Modality 1 output, Modality 2 output, and Modality 3 output, are concatenated and fed to another FC. The final FC helps to classify the hazard class using the temporal features of the three modalities. Various hyperparameters associated with the HCLM are obtained using the random search method. ADAM is chosen as an optimization algorithm over the other optimizers like SGD and RMSprop because a comparatively lower training loss and higher accuracy are observed during various trials.

E. Tiny ML

TinyML or edge ML is a concept that implements the various machine learning models directly into energy-efficient edge devices like microcontrollers. The benefit of TinyML is that it reduces the latency offered by cloud-based ML models to almost negligible. Since the ML algorithms are directly running onto the microcontrollers, the data security, or privacy issues are almost negligible in this case [25], [26]. The motivation for Tiny ML in this study arises from the challenges the UCM offers. The inherent underground environment characteristics limit the usage of high and ultrahigh-frequency communication protocols, making cloud-based ML unfavorable for UCM cases. Hence, the proposed HCLM model backed by the FEF filter is implemented on the STM32F411 microcontroller. HCLM was initially trained and optimized using TensorFlow. Since the TensorFlow and Keras models usually have a larger size to be implemented on a 512-kB memory of microcontrollers, their size needs to be reduced while keeping the accuracy within acceptable boundaries. In this study, the focus is on the post-training integer quantization (PTIQ) [27] method for model compression because all other methods are applied during training, which involves an extra burden in achieving the baseline performance when the model is trained alone. However, the PTIQ method is applied to the trained model and is, hence, the most simple among all. Although if the required model compression level cannot be achieved by PTIQ alone, then one should consider using a combination, like, PTIQ, along with pruning and clustering. For PTIQ, Tensorflow lite library is used that converts a Keras model into a quantized tflite model.

III. RESULT

A. Performance Evaluation of FEF Unit

The FEF unit is evaluated for multiple sensor node faults. Four cases are considered to evaluate the performance of the FEF unit. The first case is for single node single sensor fault, the second comprises single node multiple sensor fault, the third is multiple node single sensor fault, and the fourth is multiple nodes multiple sensor faults. The parameter values

TABLE III
EVALUATION OF FEF UNIT

Node	Case 1			Case 2			Case 3			Case 4		
	CH_4 (ppm)	Tem (°C)	Hum (%)	CH_4 (ppm)	Tem (°C)	Hum (%)	CH_4 (ppm)	Tem (°C)	Hum (%)	CH_4 (ppm)	Tem (°C)	Hum (%)
1	570	47.5	39.9	11260	51.0	38.5	950	20.9	42.5	4750	42.2	44.0
2	579	48	40	11310	51.5	39.0	890	21.0	43.0	4850	41.5	43.0
3	630	49	42	11310	51.5	39.0	910	22.1	44.0	4910	42.1	42.0
4	200	51	43	13100	52.5	30.1	0	20.5	41.9	0	32.1	0.0
5	600	52.5	44.1	11320	53.0	39.2	1000	9	43.5	15000	0	100

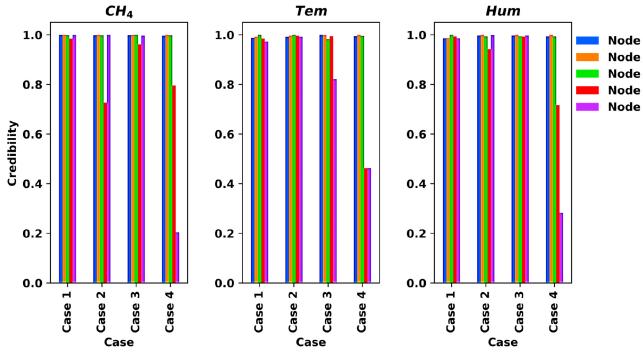


Fig. 5. Distance-based credibility.

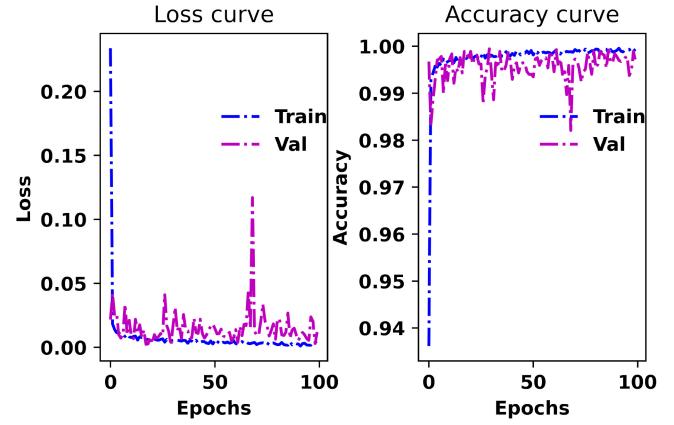


Fig. 7. HCLM learning curves.

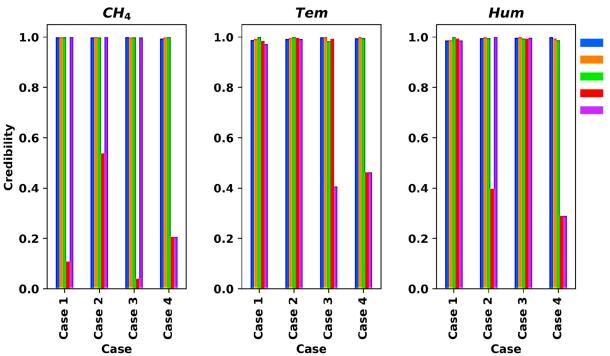


Fig. 6. Divergence-based weighted credibility.

for all four cases are shown in Table III. The parameters contributing to the faulty nodes are highlighted in bold.

The distance-based credibility and the divergence-based weighted credibility metrics for all four cases are shown in Figs. 5 and 6.

Considering case 1, where Node 4 has a faulty CH_4 sensor, it can be seen that distance-based credibility cannot segregate Node 4 because the credibility value is almost similar for all nodes, but the divergence-based weighted credibility of Node 4 significantly drops from the rest of the healthy nodes. In case 2, Node 4 has faulty CH_4 and humidity sensors. The distance-based credibility performs worst in the Humidity case but can segregate Node 4 for the CH_4 . Nevertheless, the proposed weighted credibility metric efficiently discriminates Node 4 for both of the sensors. Similarly, for case 3 and case 4, the weighted credibility metric outperforms the normal distance-based credibility metric. Hence, it provides more clear view toward faulty and healthy nodes.

B. Performance Evaluation of HCLM

1) *Training and Validation Performance:* The performance of the HCLM model is evaluated on the training and the validation set, and the learning curve is shown in Fig. 7.

From the learning curve, it is evident that after the 100 epochs, the categorical loss for training and validation sets are 0.0043 and 0.0063, which is pretty close. Hence, the model is optimized for the variance and bias problem.

2) *Effect of FEF on HCLM:* Any data-driven model can significantly suffer from data uncertainties. The uncertainty in the data can arise from various sources like noises in the environment, sensor malfunctioning, and errors in sensor readings. The effect of malfunctioning sensors on the classification accuracy of HCLM is evaluated in this section. As the FEF unit acts as a filter for faulty node data from the healthy ones within a cluster, it plays an essential role in the hazard classification model. The classification accuracy is observed to deteriorate significantly on the fault occurrence in any of the nodes. A single cluster consisting of five sensor nodes is considered, shown in Fig. 2, to evaluate this scenario. In the cluster, N1, N2, and N3 are kept healthy. However, offset, gain, and abrupt errors are introduced to the N4 and N5 during their firmware development stage. The consequence of such fault injection is that the faulty nodes' data significantly deviates from the healthy ones. The data sets for both cases are visualized in Fig. 8.

The performance of HCLM is evaluated for both cases, with and without FEF unit, on classification accuracy, F1-score, precision, and recall. The summary of these scores for both cases in the faulty node scenario is compared with the HCLM performance for nonfaulty scenarios in Table IV.

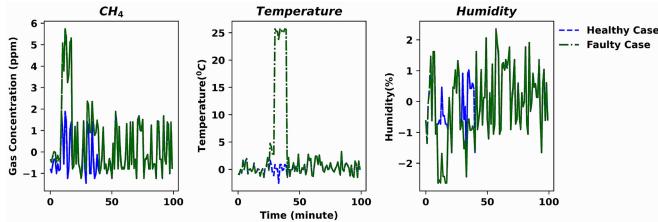


Fig. 8. Faulty and healthy nodes data.

TABLE IV
EFFECT OF FEF ON THE MODEL PERFORMANCE

Case	Accuracy	Precision	Recall	F1-Score
HCLM on non-faulty case	1	1	1	1
HCLM without FEF on faulty case	0.85	0.83	0.85	0.87
HCLM with FEF on faulty case	0.98	0.96	0.96	0.97

TABLE V
COMPARISON OF BASELINE MODEL AND TFLITE MODEL

Parameter	Baseline Model	tflite Model
Model Size	190KB	86KB
Accuracy	1	0.98

TABLE VI
COMPARISON OF HCLM WITH PRIOR ARTS

Model	Case:No faulty nodes			
	Accuracy	Precision	Recall	F1-Score
ANN [8], [9]	0.94	0.87	0.94	0.84
SVM [10], [11]	0.95	0.95	0.95	0.95
LSTM [12]	0.96	0.96	0.96	0.97
CNN-LSTM [13], [14]	1	1	1	1
FEF-HCLM (Proposed)	1	1	1	1
Case:Faulty nodes as per Fig. 8				
ANN [8], [9]	0.80	0.73	0.84	0.72
SVM [10], [11]	0.83	0.79	0.82	0.77
LSTM [12]	0.83	0.80	0.85	0.83
CNN-LSTM [13], [14]	0.85	0.83	0.85	0.87
FEF-HCLM (Proposed)	0.98	0.96	0.96	0.97

From the comparison in Table IV, it is evident that the occurrence of faulty node data significantly affects the performance of the data-driven model, but with the integration of FEF as a front-end, the performance is significantly improved compared to the HCLM without FEF.

3) TinyML Implementation:

a) *Comparison of tflite model with baseline model:* The proposed system, i.e., the hybrid of FEF and HCLM (Baseline Model), is implemented in the STMM411 controller having 512 kB of Flash memory, 128-kB RAM, and 100-MHz clock speed. Initially, the HCLM is trained on a system with 512-GB memory, 8-GB RAM, and Windows 11 (Operating system). TensorFlow, Keras, tflite, and NumPy packages are used during the training of the HCLM. The trained HCLM is then optimized for “int-8” weights and activations using PTIQ and then converted into the tflite model. The model size and prediction accuracy for the trained HCLM (baseline model) and the quantized tflite model are mentioned in Table V.

TABLE VII
COMPARISON OF HCLM WITH PRIOR ARTS

Model	multi-sensor fusion	Faulty-sensor data treatment	Edge ML capabilities
ANN [8], [9]	No	No	No
SVM [10], [11]	No	No	No
LSTM [12]	No	No	No
CNN-LSTM [13], [14]	No	No	No
FEF-HCLM (Proposed)	Yes	Yes	Yes

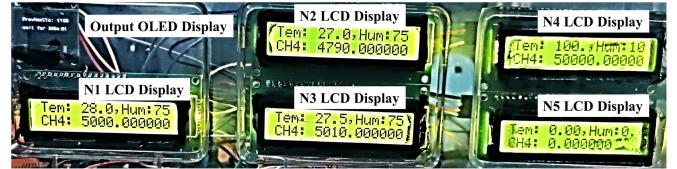


Fig. 9. Cluster observations.

Expression	Type	Value
refinedDat	Filtered Data	[3]
refinedDat[0]	float	Methane
refinedDat[1]	float	Temp.
refinedDat[2]	float	Hum.
unhealthy	Faulty Nodes	[3][5]
unhealthy[0]	int [5]	Methane
unhealthy[0][0]	int	Node1
unhealthy[0][1]	int	Node2
unhealthy[0][2]	int	Node3
unhealthy[0][3]	int	Node4
unhealthy[0][4]	int	Node5
unhealthy[1]	int [5]	Temp.
unhealthy[1][0]	int	Node1
unhealthy[1][1]	int	Node2
unhealthy[1][2]	int	Node3
unhealthy[1][3]	int	Node4
unhealthy[1][4]	int	Node5
unhealthy[2]	int [5]	Hum.
unhealthy[2][0]	int	Node1
unhealthy[2][1]	int	Node2
unhealthy[2][2]	int	Node3
unhealthy[2][3]	int	Node4
unhealthy[2][4]	int	Node5
HCLM_predicted_class	HCLM Prediction	0

Fig. 10. FEF output.

From the size comparison presented in Table V, a reduction of 104 kB in tflite version model as compared to the baseline model is observed. Moreover, due to the quantization process, no significant drop in model accuracy is observed in the tflite version model.

b) *Evaluation of the proposed model on lab-based prototype:* The proposed model is evaluated in a prototype developed in the laboratory. The evaluation of FEF and HCLM are as follows. The observations of all five nodes under a prototype cluster at some instant are shown in Fig. 9.

The FEF and HCLM output after filtering the malfunctioning nodes observations along with detected faulty nodes for all three modalities are observed using the STM32CubeIDE Live Expression panel, shown in Fig. 10.

IV. DISCUSSION

A. Comparison of the FEF-Integrated HCLM With Prior Arts

This section presents the performance comparison of the proposed model with the prior arts based on their classification accuracy and robustness. In the literature review, the preferred choices of models for mine gas hazard prediction are ANN, SVM, LSTM, and their hybrids like CNN-LSTM. Although most of the proposed models are used for regression tasks, in this study, those models are recreated with a softmax layer as the outer layer for multiclass classification. These models are trained and optimized for the data samples presented in this study, and their optimized instances are compared with the proposed model for the test samples, shown in Table VI.

From Table VI, with no faulty node case, it is evident that the hybrid CNN-LSTM and the proposed model outperform all other classifiers. This justifies that the hybrid of CNN-LSTM is a suitable classifier for the mining hazard case. Therefore, the architecture of HCLM is also based on hybrid CNN-LSTM. However, when these models are evaluated with faulty node data, a significant drop in performance is observed for all models except the proposed one. Hence, the proposed model can accurately classify the hazard even for faulty node scenarios.

B. Comparison of Existing Work for Real-Time UCM Scenarios

A brief comparison of the proposed model against the prior arts concerning the feature associated with data reliability and UCM applicability is shown in Table VII.

As discussed earlier, the multisensor fusion and faulty sensor data treatment are essential for realizing a reliable NN model in real time. As depicted in Table VII, none of the existing studies made any development in these aspects. Additionally, the existing studies do not emphasize the end-to-end implementation of their NN model; they are mostly validated in the laboratory. However, such a system requires a reliable communication with the sensor network in the UCM. In actual scenarios, wired or wireless communication networks face many challenges due to the inherent properties and complexities of the coal mines. Therefore, the currently available gas hazard monitoring solutions cannot be considered reliable to realize the real-time-based NN model. The solution is an edge ML-enabled standalone system that can perform the signal processing in the actual location. Therefore, the proposed model is realized by transforming the HCLM model into its TinyML version and deploying the integrated FEF and TinyML-based HCLM into a battery-operated microcontroller.

V. CONCLUSION

This study proposes a reliable methane hazard monitoring model using the hybrid of FEF and HCLM. The laboratory investigation of the proposed system prevailed that the FEF-integrated HCLM outperformed the existing similar NN-based model in terms of accuracy, precision, recall, and F1-score. For the challenging and dynamic environment of the coal mine, the proposed system can be considered as a standalone

system for real-time-based signal processing and gas hazard estimation. The multisensor fusion approach can opt for a reliable solution for the UCM where sensor malfunctioning is inevitable because of various issues. The in-situ and edge ML realization of the complete setup can further reduce the overall wired or wireless network infrastructure, thereby increasing the flexibility and scalability toward the dynamic operating conditions.

However, the requirement of multiple sensor nodes compared to single sensor-based implementation increases the initial cost of the proposed system. The TinyML transformation is limited to the hardware requirement of the edge devices. Therefore, it reduces the NN-model design flexibility. The proposed work can be further extended to the hazard forecast of UCM using the transfer learning technique.

REFERENCES

- [1] S. Schatzel, R. 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 Explor. Inc.*, vol. 342, no. 1, pp. 51–61, 2017.
- [2] L. Muduli, D. P. Mishra, and P. K. Jana, "Application of wireless sensor network for environmental monitoring in underground coal mines: A systematic review," *J. Netw. Comput. Appl.*, vol. 106, pp. 48–67, Mar. 2018.
- [3] M. Tutak and J. Brodny, "Analysis of the impact of auxiliary ventilation equipment on the distribution and concentration of methane in the tailgate," *Energies*, vol. 11, no. 11, p. 3076, 2018.
- [4] S. Schatzel, R. 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 Explor. Inc.*, vol. 342, no. 1, pp. 51–61, 2017.
- [5] T. Howard, "University of Texas study underestimates national methane emissions at natural gas production sites due to instrument sensor failure," *Energy Sci. Eng.*, vol. 3, no. 5, pp. 443–455, 2015.
- [6] S. Basu, S. Pramanik, S. Dey, G. Panigrahi, and D. K. Jana, "Fire monitoring in coal mines using wireless underground sensor network and interval type-2 fuzzy logic controller," *Int. J. Coal Sci. Technol.*, vol. 6, no. 2, pp. 274–285, 2019.
- [7] X. Tong, W. Fang, S. Yuan, J. Ma, and Y. Bai, "Application of Bayesian approach to the assessment of mine gas explosion," *J. Loss Prevent. Process Ind.*, vol. 54, pp. 238–245, Jul. 2018.
- [8] D. P. Mishra, D. C. Panigrahi, P. Kumar, A. Kumar, and P. K. Sinha, "Assessment of relative impacts of various geo-mining factors on methane dispersion for safety in gassy underground coal mines: An artificial neural networks approach," *Neural Comput. Appl.*, vol. 33, no. 1, pp. 181–190, 2021.
- [9] 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.
- [10] M. You, S. Li, D. Li, and S. Xu, "Applications of artificial intelligence for coal mine gas risk assessment," *Safety Sci.*, vol. 143, Nov. 2021, Art. no. 105420.
- [11] J. Deng, W. Chen, C. Wang, and W. Wang, "Prediction model for coal spontaneous combustion based on SA-SVM," *ACS Omega*, vol. 6, no. 17, pp. 11307–11318, 2021.
- [12] K. Kumari et al., "UMAP and LSTM based fire status and explosibility prediction for sealed-off area in underground coal mine," *Process Safety Environ. Prot.*, vol. 146, pp. 837–852, Feb. 2021.
- [13] Prince and A. S. Hati, "Convolutional neural network-long short term memory optimization for accurate prediction of airflow in a ventilation system," *Expert Syst. Appl.*, vol. 195, Jun. 2022, Art. no. 116618.
- [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. Prot.*, vol. 152, pp. 249–263, Aug. 2021.
- [15] N. Ghosh, R. Paul, S. Maity, K. Maity, and S. Saha, "Faultmatters: 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.

- [16] M. Gieras, R. Klemens, G. Rarata, and P. Wolański, "Determination of explosion parameters of methane-air mixtures in the chamber of 40 dm³ at normal and elevated temperature," *J. Loss Prevent. Process Ind.*, vol. 19, nos. 2–3, pp. 263–270, 2006.
- [17] R. Li, X. Meng, and R. Si, "Experimental studies on the effect of humidity on methane explosion characteristic in confined spaces," in *Proc. Int. Conf. Civil Transp. Environ.*, 2016, pp. 1273–1278.
- [18] M. Sharma and T. Maity, "Multisensor data-fusion-based gas hazard prediction using DSET and 1DCNN for underground longwall coal mine," *IEEE Internet Things J.*, vol. 9, no. 21, pp. 21064–21072, Nov. 2022.
- [19] M. N. Khan and S. Anwar, "Paradox elimination in Dempster–Shafer combination rule with novel entropy function: Application in decision-level multi-sensor fusion," *Sensors*, vol. 19, no. 21, p. 4810, 2019.
- [20] M. Koziełski, M. Sikora, and Ł. Wróbel, "Data on methane concentration collected by underground coal mine sensors," *Data Brief*, vol. 39, Dec. 2021, Art. no. 107457.
- [21] P. Thakur, *Advanced Mine Ventilation: Respirable Coal Dust, Combustible Gas and Mine Fire Control*. Cambridge, U.K.: Woodhead Publ., 2018.
- [22] 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.
- [23] T. Wang et al., "Flammability limit behavior of methane with the addition of gaseous fuel at various relative humidities," *Process Safety Environ. Prot.*, vol. 140, pp. 178–189, Aug. 2020.
- [24] X. Shi, N. Zhu, and G. Zheng, "The combined effect of temperature, relative humidity and work intensity on human strain in hot and humid environments," *Build. Environ.*, vol. 69, pp. 72–80, Nov. 2013.
- [25] R. Sanchez-Iborra and A. F. Skarmeta, "TinyML-enabled frugal smart objects: Challenges and opportunities," *IEEE Circuits Syst. Mag.*, vol. 20, no. 3, pp. 4–18, 3rd Quart., 2020.
- [26] L. Dutta and S. Bharali, "TinyML meets IoT: A comprehensive survey," *Internet Things*, vol. 16, Dec. 2021, Art. no. 100461.
- [27] M. Sailesh, K. Selvakumar, and N. Prasanth, "A novel framework for deployment of CNN models using post-training quantization on microcontroller," *Microprocess. Microsyst.*, vol. 94, Oct. 2022, Art. no. 104634.



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



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, Shibpur, 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 (ISM), Dhanbad, India.