The Visual Computer (2022) 38:4457–4468
https://doi.org/10.1007/s00371-021-02308-x

ORIGINAL ARTICLE

Online health status monitoring of high voltage insulators using deep learning model

Dipu Sarkar[1] · Sravan Kumar Gunturi[2]

Accepted: 3 September 2021 / Published online: 24 September 2021

Abstract

The inspection of electrical components has long been an important issue in the power distribution system. Unmanned drones are impressive surveillance systems with a powerful spatial and remote sensing capability. This paper proposes a system to monitor the health of the ceramic insulators that uses aerial images as a source of information and deep structured learning model for the data interpretation. The key drawbacks of existing monitoring systems are poor detection accuracy and lack of real-time execution, making it more complicated to obtain attributes from aerial photographs. The focus of this paper is to increase accuracy of detection while operating in real-time using You Only Look Once version 3 (YOLOv3). The novelty of the proposed system is that it combines deep learning and the Internet of Things using a single embedded device called Raspberry Pi. For the scientific investigation, we equipped Raspberry Pi with a test image as an input to detect an insulator's health status using YOLOv3. Many aerial images are not clear due to motion blur. Excluding such low-resolution training images will affect accuracy. So we used a super-resolution CNN to reconstruct a blurred image as high-resolution image. The efficiency of the proposed system has been tested using a private data set consisting of a variety of scenes containing high-voltage power line insulators. The results show that the suggested system is quick and accurate in the identification and classification of insulators.

Keywords Health monitoring · YOLOv3 · SRCNN

1 Introduction

To limit power interruption, utilities conduct visual examinations on their equipment to plan for the essential servicing or replacement. They carried these examinations by practising traditional patrolling methods which are often slow and need more manpower [1]. As the insulators play a major part in the safety of the power system [2], the power companies have given high preference to faulty insulator detection systems that are quick and less expensive for the inspection team. The insulators used in transmission and distribution systems are supposed to resist constant electrical, mechanical, and thermal forces under various environmental circumstances [3]. These forces may cause a depreciation of surface resistance, flash-over voltage, and puncture strength. This may produce a high leakage current, which degrades the insulation strength [4]. If an insulator string of transmission lines is exposed to contaminated atmospheres with high moisture, their insulating ability is decreased and dry areas on the insulator's plane may be conductive [5]. This leads to a partial flow of current and then to a line-to-ground discharge referred to as a flash over [6]. The partial discharge of the defective insulator may also release electro-magnetic waves and sound waves. The irregular dissipation of heat over the defective insulator plane ends in the decrease of insulation resistance and the rise of leakage current [7].

Over several years, researchers have been studying fast and effective models for the inspection of power components. In these, the Buzz method is the oldest method that performs a physical examination within each insulator in a string by trying to apply a large voltage across it to listen to a buzz-like sound [8]. The safety issues of this method have led to the investigation of the correlation between electromagnetic signals produced by partial discharge (PD) activity and

✉ Dipu Sarkar
dipusarkar5@rediffmail.com

1 Department of Electrical and Electronics Engineering, National Institute of Technology Nagaland, Dimapur 797103, India

2 Department of Electronics and Instrumentation Engineering, National Institute of Technology Nagaland, Dimapur 797103, India

defects, such as pollution and insulation breakdown. Wong et al. [9] uses VHF (30–300 MHz) signal processing methods to detect cracks on the insulators. The frequency spectrum generated by the fast Fourier transform was analysed by the fractal algorithm to detect anomalies in the insulator. Algae fungus [10] poses a potential hazard to the external insulation of the electrical system due to the distinctive characteristics of the organic foul. The impact of algae contamination was high when algae coverage was greater than 20%. Zhong et al. [11] demonstrated a method for detecting insulator anomalies using ultrasonic technology. Many researchers have used neural networks to simulate energy systems. The authors of [12] consider the height, the diameter, the overall leakage frequency, the surface conductivity, the number of sheds, and the number of chains on it. The efficiency and accuracy of the system deprives complex tasks like pre-processing of image models, segmentation techniques and mining in traditional machine-learning models [13]. To overcome the difficulties of the machine learning techniques [14,15], deep Learning methods were developed for extracting and classifying relevant data from new images. In recent times, the convolution neural network has made significant strides in biomedical imaging, such as mapping mitosis cell identification [16,17]. Nasr et al. [18] used 170 colour images of the MEDNODE dataset. They have augmented pictures by using different augmentation methods, such as cropping and rotation. They generated 35 augmented images from a single image. Lopez et al. [19] used the 1300 image dataset (ISBI 2016) in which they used 900 images for training and remaining for testing. They have achieved 81.33% accuracy with the VGG16 architecture of deep CNN. The authors of [20] investigate a new insulator detection technique for aerial photographs of UAVs in power transmission line inspection based on the single shot detector (SSD). On the aerial picture collection, the SSD is used to carry out the automated function learning process. Instead of inefficient and unguided hand-crafted feature extractors, SSD's model can extract high-level features and speed up detection. Li et al. [21] suggest a novel approach in which identification and segmentation networks are cascaded to classify the defect at the global and local levels. In the whole shot, the improved Faster RCNN is used to capture both defects and insulators. To completely extract features, ResNeXt101 is used as the feature extraction network, and the feature pyramid network is designed to improve the ability to detect small objects. The drone-based implementation of the YOLOv2 neural network model for insulator detection was introduced by [22].The technology was tested using real-time aerial photographs taken by drone. The current approaches to detect insulator damages are summarised in Table 1.

Only a few fully automatic inspection systems have been discussed in the literature. Therefore, a prototype model has been developed for testing the effectiveness of an embedded device for the deep learning object recognition algorithm. For this purpose, Raspberry Pi has been used as an embedded device that also operates both the Deep Learning and IoT functions. The YOLOv3 algorithm was used to classify the insulator, and a message is sent to the utility centre via the Blynk server if the bad insulator has been found. The rest of the article has been structured as follows: Sect. 2 discusses the theoretical background. Next, Sect. 3 provides a description of the methods used to implement the system proposed. Section 4 describes the results of the experiment. Section 5 gives the conclusions.

2 Theoretical background for deep learning model: YOLOv3

YOLO recognises objects by splitting the image into grid blocks rather than the regional proposal approach used in two-stage detectors. The function map of the YOLO output is intended to display bounding box coordinates, object score, and class scores. YOLO [23] also allows several objects to be recognised with one inference. Consequently, the speed of detection is much higher than that of traditional methods. Nevertheless, localisation errors are high due to grid unit processing, and the recognition precision is poor, making it unsuitable for object recognition applications.

YOLOv2 was suggested [24] to address the aforementioned issues. It improves detection efficiency by adopting a batch normalisation process for the convolution layers. It also incorporates an anchor box, multi-layered training, and fine-grained characteristics [25]. However, the accuracy of detection for small objects is indeed low. Therefore, YOLOv3 [26] was introduced to resolve the drawbacks of YOLOv2. As shown in Fig. 1, YOLOv3 consists of convolution layers and consists of a deep layered network for better accuracy [27]. YOLOv3 uses the residual skip relation to tackle the fading gradient problem of deep networks and a method of up-sampling and concatenation that provides fine-grained features for the recognition of small objects [28]. The most notable feature is the identification at three different levels that used in a pyramid network feature [29]. It helps YOLOv3 to track objects of different sizes. When an image is given as input with three channels (i.e. R, B, and G) into the YOLOv3 system, data regarding bounding box coordinates, scores of objects and classes are obtained as shown in Fig. 1. The outcomes from the three levels are mixed and analysed using non-maximum suppression. After this, the results of the final detection are determined. Hence, YOLOv3 is suitable for object recognition applications in aspects of precision and speed.

Table 1 State-of-the-art monitoring methods to detect insulator damages

Past research	Proposed method
[8]	The buzz method
[9]	VHF (30–300 MHz) signal processing methods
[10]	Algae foul method
[11]	Ultrasonic technology
[13]	Machine learning models
[15]	Deep learning models
[19]	VGG16 architecture of deep CNN
[20]	Single shot detector
[21]	Faster RCNN
[22]	YOLOv2

3 Proposed insulator health monitoring system

The principal technique for insulator condition monitoring is shown in Fig. 2. The pre-processing of the drone surveillance inspection image can be carried out in the following steps. First, the initial drone inspection image series must be split into two categories. One is a high-quality image array that can be transferred directly to the next stage; the other is a low-resolution blurred image package. Also, the image set is transformed into a high-resolution image set using the super-resolution reconstruction process described briefly in Sect. 3.2. The transformed images can be later combined with the initial high-resolution images and then used as new image sets for further processing. In addition, the original dimension of the source images has been resized to a new scale of 416×416 at the reference stage to accelerate the learning process. To extract insulator features, the resized image is tested as an input to the Darknet-53. The feature-pyramid-network (FPN) method yields predictions over three distinct stages through Darknet-53. Overall, the YOLOv3 predictions include bounding box variables, item ranking, and class prediction.

The suggested system has four components that may be explained:

1. Image pre-processing was performed in the first phase to distinguish low-level fuzzy photographs from good pictures.
2. In the second stage, an image super-resolution process was performed to transform low-level fuzzy photographs into super-resolution pictures using the SRCNN method.
3. In the third phase, the transformed pictures were sent through a deep learning object detector named YOLOV3.
4. Finally, if YOLOv3 discovered any bad insulators in the photographs, the information about the faulty insulators is forwarded to the utility team via the Blynk android application utilising IoT architecture.

3.1 Image pre-processing unit

Here, the Laplace distribution approach was used to distinguish the blurred aerial image from the clear aerial image. The Laplace operator has been chosen to calculate the second-degree image differential that increases the contrast between the neighbourhood image elements. In general, the Laplacian operator is used first to transform the image and then to determine the deviation. In sharp pictures, the boundary tends to be fairer, so that the difference rises considerably. Alternatively, the blurred edge detail of the images is comparatively less so that the difference becomes less.

Therefore, if the deviation is less than the threshold specified, the image will be considered as blurred. Conversely, the image is labelled as clear if the disparity is greater than the stated threshold value as shown in Fig. 3. As shown in Fig. 4, the initial surveillance images are subdivided into a dubbed image group and a normal image group after the above procedure. The identified blurred images are then used as inputs to the SRCNN model. Finally, the transformed SRCNN images and the clear images in the initial dataset are merged to form a new image dataset as shown in Fig. 4.

3.2 SRCNN reconstruction

3.2.1 SRCNN

Few aerial images show blurriness due to shuddering drone body and imagery exposure problems; this will severely impede the effective monitoring of insulator health from a drone inspection. Image pre-processing would be a required phase in the development of a system of deep learning practices. The SRCNN model proposed by [30], a network training algorithm for the pre-processing of image data. The SRCNN algorithm addresses the above-mentioned issues effectively by changing poor-resolution images to super-resolution images, respectively. The implementation of the SRCNN model typically consists of three stages: image patch

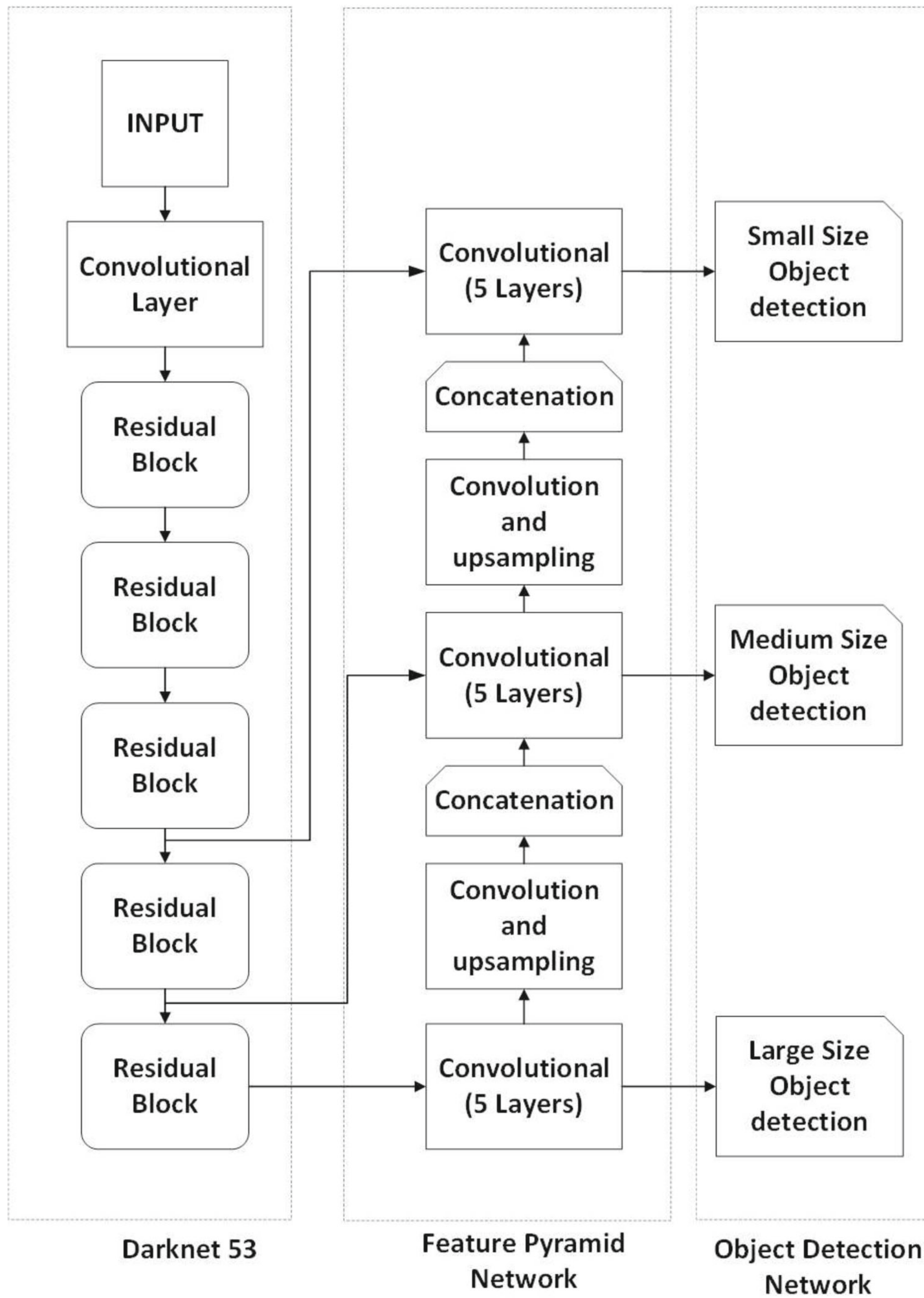

Fig. 1 Overview of YOLOv3 network architecture

extraction, nonlinear projection, and super-resolution reconstruction.

Image patches extraction This procedure selects (overlapping) maps from the Y-frame and shows each map as a large spatial vector. These variables comprise of a set of function maps, the number of which is proportional to the vector dimension. The standard procedure is to get the patches from the initial monitoring photographs and use a sequence of convolution filters. Each convolution filter can be considered as a basis, and simple development can be included in the network for optimisation. The process of the first layer can be defined as follows:

$$F_1(Y) = \text{maximum}(0, W_1 * Y + B_1) \tag{1}$$

where W_1 represents n_1 filters of the size ($c \times f_1 \times f_1$), 'c' is the number of raster bands of the input, f_1 is the map coverage of the filter, B_1 is an n_1-spatial segment, and '$*$' represents convolution. The result of the convolution cycle includes 'n_1' attribute features and the final performance of

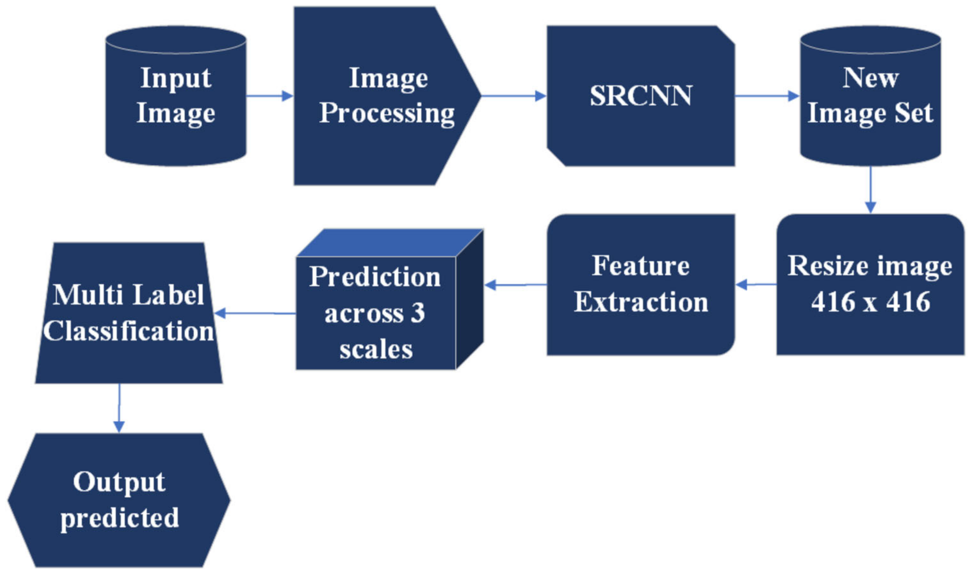

Fig. 2 Proposed insulator monitoring model using YOLOv3

the primary convolution stage is achieved by the rectified linear unit ReLU (max(0,x)) activation system [31].

Nonlinear projection The n_1-dimensional attribute parameter is derived from every input image of the primary convolution layer. The resulting n_1-dimensional attribute vector in the second convolution layer is not projected linearly to the n_2-dimensional feature segment. The second layer processing can be defined as:

$$F_2(Y) = \text{maximum}(0, W_2 * F_1(Y) + B_2) \tag{2}$$

where W_2 denotes n_2 filters of size $n_1 \times f_2 \times f_2$ and B_2 represents a n_2-spatial vector. The convolution product contains n_2 feature maps.

Reconstruction of super resolution image The last convolution sectional layer combines whole the super-resolution frames generated with the aid of upper layer to generate a super-resolution model, that is, the SRCNN network's final image production. Therefore, the third layer cycle may be seen as:

$$F_3(Y) = W_3 * F_2(Y) + B_3 \tag{3}$$

where W_3 denotes c filters of scale $n_2 \times f_3 \times f_3$ and B_3 represents a c-sized vector.

3.2.2 SRCNN structure

The SRCNN uses the two cubic interpolation models to expand the collected smeared monitoring image as necessary and to record the interconnected image as Y.

Fig. 3 Calculation of blurriness in the input images

Super-resolution restoration targets at restoring Y to a rasterised image H suitably high-resolution equivalent to the real X. The appropriate end mapping function F(Y) may be obtained by training. The basic arrangement for the SRCNN model is shown in Fig. 5. It can be observed that there is a three-stage CNN in the total structure. The primary convolution stage extracts image frames from Y, then identifies certain characteristics with low resolution. The next convolutional layer uses nonlinear mapping to produce high-resolution features. Finally, the reconstruction of super-resolution images is accomplished through the third convolution layer, which is analogous to creating images close to the actual image resolution. To achieve the mapping feature F between super-resolution images, the θ= $(W_1, W_2, W_3, B_1, B_2, B_3)$ system variable must be trained

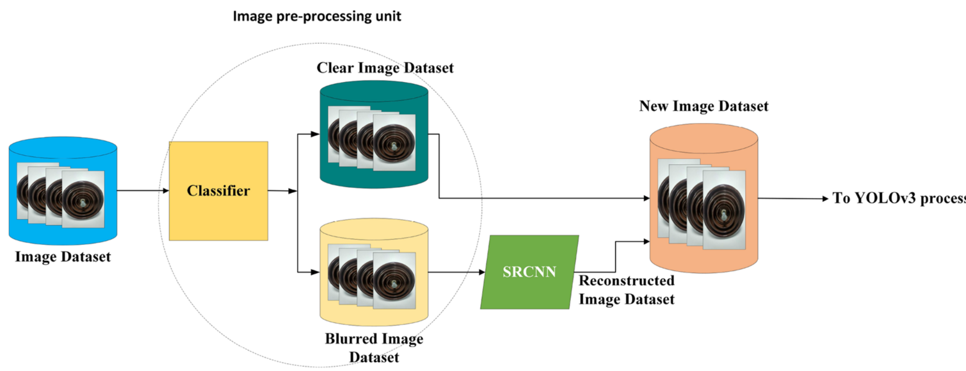

Fig. 4 Image pre-processing unit

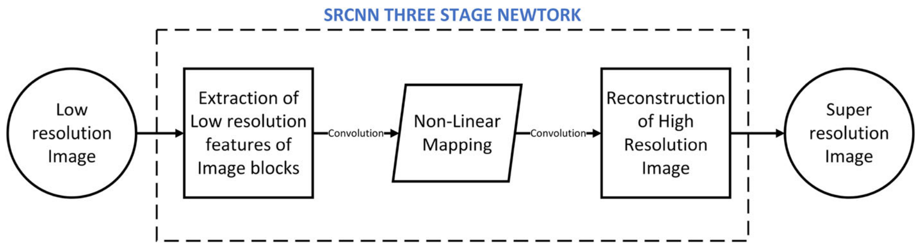

Fig. 5 SRCNN network structure

in the training process.

$$L(\theta) = \frac{1}{n}\sum_{i=1}^{n} ||F(Y_i; \theta) - X_i||^2 \tag{4}$$

where the n denotes number of training examples, X_i is the actual image, Y_i represents the poor-resolution reference, and $F(Y_i; \theta)$ is the clear one produced by SRCNN model.

3.3 Monitoring insulator health using YOLOv3

YOLOv3 makes use of the FPN concept to predict boxes at various scales. To complete the detection process, it uses a variety of convolution layers and extra layers (residual layers) and manages the attributes of the full image to predict any minimum bounding rectangle. Meanwhile, it forecasts top-to-bottom training across all bounding rectangles in all classes, which hold significant average accuracy and high efficiency in real time. YOLOv3 begins the process by splitting the input monitoring image as $N \times N$ blocks and adds a bounding rectangular anchor for every ground truth on the map. For each bounding box, the network finds 4 parameters (t_x, t_y, t_w, t_h), as shown in Fig. 6 then use a method to predict 4 related coordinates: mid co-ordinates: (b_x, b_y) of the bounding rectangle, the height b_h and the width b_w. The minimum prediction of bounded rectangles and the equations of Intersection over Union (IOU) are given as follows:

$$b_x = \sigma(t_x) + c_x \tag{5}$$

$$b_y = \sigma(t_y) + c_y \tag{6}$$

$$b_w = p_w e^{t_w} \tag{7}$$

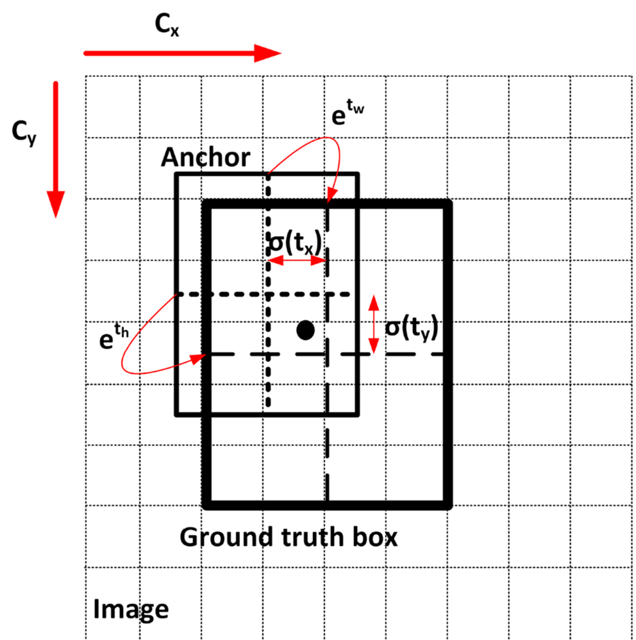

Fig. 6 Bounding box detection

$$b_h = p_h e^{t_h} \tag{8}$$

$$\text{IOU} = \frac{\text{area}(BB_{dt} \cap BB_{gt})}{\text{area}(BB_{dt} \cup BB_{gt})} \tag{9}$$

where the amount of variance between the detected bounding rectangle and the ground truth box is the IOU shown in expression 9. BB_{gt} is the rectangle of ground truth relying on learning label, BB_{dt} is the recognising bounding rectangle, and (.) shows the area of region.

3.3.1 YOLOv3 network structure

As stated in Sect. 2, Fig. 1 shows the YOLOv3 main device configuration, which acquires Darknet-53 framework. This network is a YOLOv2 fusion [32], Darknet-19 [33], and ResNet [34]. So, YOLOv3 mainly uses 1×1 and 3×3 convolution kernels, and some relevant shortcut structures. First, the input surveillance image is processed, then its size is changed to 416×416, and then YOLOv3 is used to process it.

1. The primary section contains two convolution layers. The image input size is $416 \times 416 \times 3$ and the size of kernel is $3 \times 3 \times 64$ and $3 \times 3 \times 32$, respectively. After the convolution cycle, the size of the output match function is reduced to $208 \times 208 \times 64$.
2. The second section consists of three convolution stages, accompanied by a residual layer. The size of kernel is $3 \times 3 \times 128$, $3 \times 3 \times 64$, and $1 \times 1 \times 32$ and after the convolution process, the output map is reduced to $104 \times 104 \times 128$.
3. The third section consists of 5 convolution layers with two layers of residual form layers. The convolution kernel size is $3 \times 3 \times 256$, $3 \times 3 \times 128$, and $1 \times 1 \times 64$ and after the convolution process, the output function map is reduced to $52 \times 52 \times 256$.
4. The fourth section comprises 17 layers of convolution accompanied by 8 layers of residual. The size of the convolution kernel is $1 \times 1 \times 128$, $3 \times 3 \times 256$, and $3 \times 3 \times 512$, and the map of the output function is reduced to $26 \times 26 \times 512$ after the convolution process has been completed.
5. The fifth section comprises 17 layers of convolution and also 8 layers of residual. The dimensions of the convolution kernel are $1 \times 1 \times 256$, $3 \times 3 \times 1024$, and $3 \times 3 \times 512$ and after the convolution cycle the map of the output function is reduced to $13 \times 13 \times 1024$.
6. The sixth section comprises eight layers of convolution form and four layers of residual form. The dimension of the convolution kernel is $1 \times 1 \times 512$ and $3 \times 3 \times 1024$, and after the convolution process, the map of the output function will remain the same.
7. The last segment consists of three networks for predictions. YOLOv3 predicts rectangular boxes on three stages and then selects certain scales attributes. The prediction of network is a tensor of $10 \times 10 \times (3 \times (4 + 1 + 4))$ for 4 minimum rectangular box corrections, 1 projection of objects, and 4 classifiers.

3.3.2 Training

The training of YOLOv3 network is split up into following 3 activities:

Step 1 As the size of the aerial image captured by surveillance drone is 5280×2970, that would be too high to be the input of the network. The real scale of the picture is changed to 416×416 to accelerate the process of training.

Step 2 The VOC2007 [35] dataset pattern is used to label the exterior form of the leading edge (LE) erosion, the vortex generator (VG) panel, the VG with vanished teeth, and the lightning receptor shows up in particular image.

Step 3 Initiate system variables of the YOLOv3 algorithm and train the network to get variables for the identification of specified objects.

3.3.3 Essential parameters

This paper presents an additional study of the choice of three essential variables

Batch size Theoretically, the greater the volume, the easier the preparation will be. Nonetheless, owing to the constraint of hardware constraints, we cannot increase the value forever, so authors attempted four different batch sizes of 8, 16, 64, and 128, respectively. There would be no loss of power when the batch size were chosen as 64, 16, 8 during preparation, so we chose 64 as the batch size based on the above-mentioned argument.

Weight decay To prevent over-fitting, first fix the correct rate of learning and then modify the decay measure from the constant value (0.01) to the final measure (0.0005).

Ignore thresh Ignore thresh is the IOU threshold value, which measures the amount of IOUs used in the loss calculation. If the pre-defined threshold is less, which will take to under-fitting. On the other side, if the threshold limit is high it is obvious to cause over-fitting. Thus, the ignore thresh value is set as 0.65 on the basis of above argument and the case at hand.

Selection of parameters The accuracy of the detection is influenced by the choice of the three variables mentioned above. To avoid under-fitting and over-fitting during planning, it is therefore necessary to change these parameters. To boost the precision of the typical detection, YOLOv3 adopts a multi-label categorisation that is distinct against the old interpretations that use a contradictory label. The logistic classifier is used to determine the objectness value for

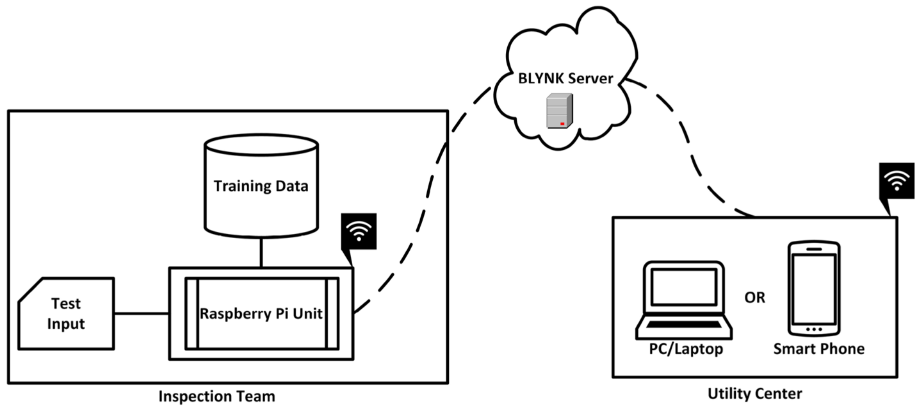

Fig. 7 Block diagram of proposed system

each bounding rectangle. For classification loss while testing, YOLOv3 uses a discrete cross-entropy loss for each number, which eliminates the MSE commonly exploited in past versions. The loss factor used in the YOLOv3 feature training is shown as follows:

$$L(s_n) = \begin{cases} -\log(s_n), & \text{if } g_n = 1 \\ -\log(1 - s_n), & \text{if } g_n = 0 \end{cases} \tag{10}$$

where n represents the sample count, $s_n \in [0, 1]$ reflects the objectness value anticipated by the system, which calculates the expected likelihood that the nth sample is insulator damage and g_n shows the ground truth. It needs to be observed that $g_n \in (0, 1)$ implies when the nth observation relates to the object class. Network variables are trained by reducing the loss to all samples, i.e. $\sum_n L(s_n)$.

In this article, the adaptive moment estimation, shortly referred to as the Adam optimisation procedure [36], is used to change network parameters. Adam would be the first-order optimisation mechanism that can substitute the standard stochastic gradient descent method and adjust network weights recursively depends on a training data. The model calculates the appropriate adaptive training score for various values by determining the gradient estimate of primary and secondary moments. It integrates two optimisation models, along with the benefits of root-mean-square propagation and adaptive gradient algorithm [37], that are useful for improving the effectiveness of scattered gradients and efficiency of training.

3.4 Inspection team and utility team

The maintenance department carried out a physical inspection of the power line units. They collect different power component images. As shown in Fig. 7, every image in the set would be used as a test input to the Raspberry unit. The Raspberry is running a python application to track the structural health of the insulator via the deep learning system. The YOLOv3 model, as discussed in the Sect. 3.3.1, is the best deep learning model for the recognition of objects with greater precision.

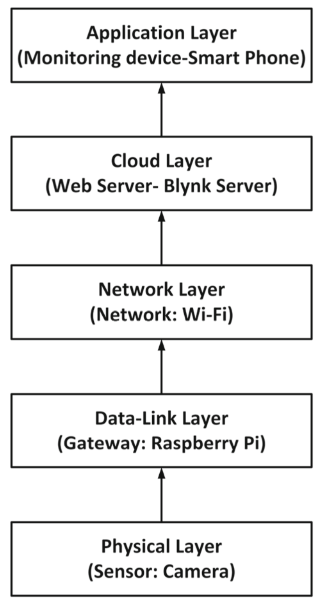

Fig. 8 The IoT architecture

The innovation of the proposed architecture is the integration by an embedded device of a deep learning interface into the Internet of Things (IoT). As seen in Fig. 7, the Raspberry board was used as an IoT device to monitor the state of the insulator remotely using an Android mobile phone or a machine. The IoT architecture of the new system is seen in Fig. 8. As seen in the diagram, the application layer maintains and monitors the user interface, and the cloud layer is linked to the web-server. The WiFi network is used by the network layer. Raspberry is used as a gateway device that is part of the data link layer of the architecture. In the end, the camera comes under the physical layer. After running a deep learning program, Raspberry can run another python module to transfer test data information to the Utility Centre via the Blynk Server wirelessly.The utility centre is fitted with an end module, i.e. a mobile device or a Blynk-enabled device. With

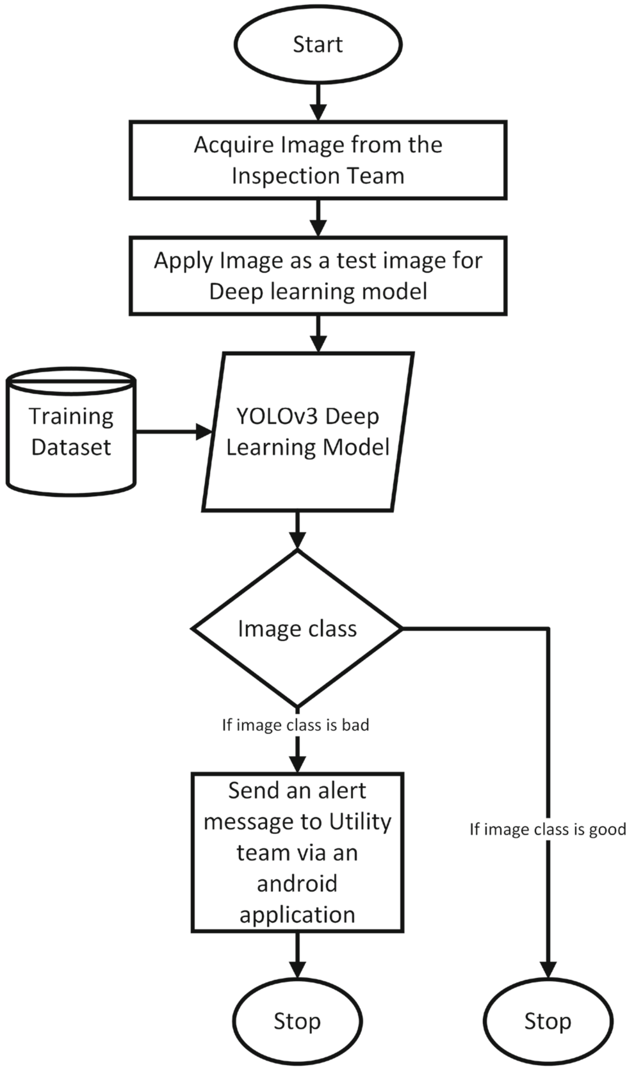

Fig. 9 Flow chart of proposed system

any bad insulator image as a test input, the power distribution end will be notified as a push notification to a smartphone. Figure 9 shows the clear design flow of the suggested system.

4 Results and discussion

4.1 Hardware configuration

The Raspberry pi 4 Model B, Quad-Core 64-bit Broad-com 2711, Cortex A72 Processor, Operating Power 5V@3A via USB Type-C Port is used in the system to check the effectiveness of the proposed high voltage insulator inspection system. The Raspbian OS was used as an operating software in the device. The Android mobile installed with Blynk application was used at the utility end.

Table 2 Insulator class distribution

Class	Good (%)	Bad (%)	Total (%)
Train	400 (50%)	400 (50%)	800 (80%)
Test	100 (50%)	100 (50%)	200 (20%)

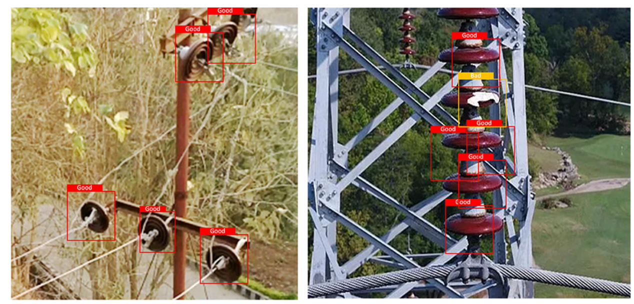

Fig. 10 Sample output of YOLOv3 model

4.2 Dataset

There are no public databases available for insulator groups. As a result, the authors experimented with a private dataset. The 1000 image dataset was used as an image source for the proposed model. For which 500 pictures have a good class of insulators. The remaining 500 pictures show a bad type of insulators. From the image dataset, 80% of the images have been utilised for training purposes and the remaining 20% have been employed for testing purposes. The class distribution of images is shown in Table 2.

4.3 Evaluation of proposed system

The proposed system was evaluated by standard parameters such as accuracy, sensitivity, and specificity. Detection and classification process based on YOLOv3 results has been deployed accordingly. This shows the effectiveness of the YOLOv3 model; it was compared with Fast R-CNN model. During the training process, the accuracy of the two deep learning algorithms Fast R-CNN and YOLOv3 in recognising the training objects is measured while the number of iterations is 500, 1000, 2000, 3000, and 4000, respectively.

$$\text{Accuracy} = \frac{\text{True Positive} + \text{True Negative}}{\text{True Positive} + \text{False Positive} + \text{True Negative} + \text{FalseNegative}} \tag{11}$$

The test input and output images of the insulators are presented in Fig. 10 to demonstrate the efficacy of the proposed design. As shown in Figure, the test output images show that the proposed system is capable of identifying the insulator with the bounding box and the class also marked at the top of the rectangular box. The current work has classified

Fig. 11 Blynk application alert to utility team

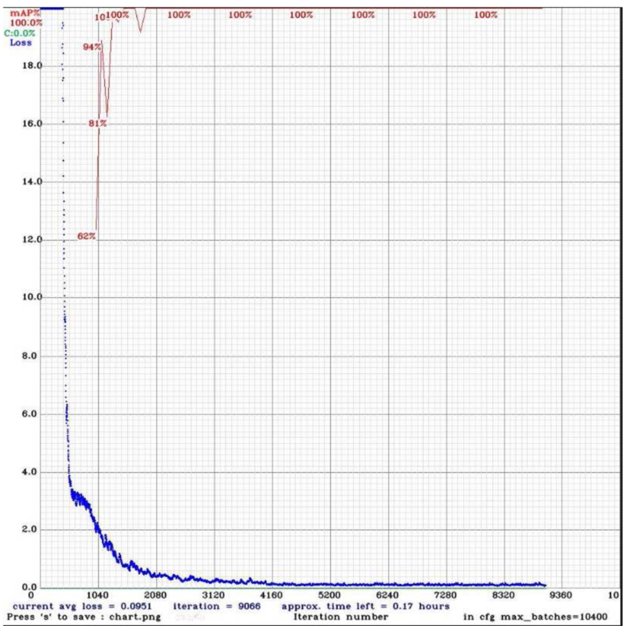

Fig. 12 Loss versus iterations

the insulator class with 95.6% accuracy. While testing bad insulator, the IoT part has alerted the Utility centre through Android application successfully as shown in Fig. 11. Further our model calculates sensitivity and specificity. Sensitivity shows the number of true classes, whereas specificity shows false classes. The metrics are determined as shown in the following equations.

$$\text{Sensitivity} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \tag{12}$$

$$\text{Specificity} = \frac{\text{True Negative}}{\text{True Negative} + \text{False Positive}} \tag{13}$$

$$\text{map} = \frac{1}{N}\sum_{i=1}^{N} \text{Average Precision}_i \tag{14}$$

The sensitivity and specificity of the proposed frameworks are shown in Table 3. In order to estimate the detection results, the authors measured the mean average precision (mAP) which is one of the common metrics used to calculate the accuracy of target detectors such as Fast R-CNN and YOLOv3. The mAP of two suggested models is also plotted in Table 3. The loss function has also been used to explain the discrepancy between the expected value and the real value of the network used to test the effectiveness of the forecasting model. Figure 12 demonstrates the loss function versus iterations for the suggested model indication map of 100%.

5 Conclusion

In this article, a deep structured learning model for data interpretation has been employed to monitor the health of the ceramic insulators. A vision-based arrangement was carried out to screen the health of the ceramic insulators that used camera pictures as a source of data. We have trained a YOLOv3 model to recognise and classify every image patch in the obtained images. During the training process, 1000 insulator images of size 256 × 256 were used. We obtained the results with 95.6% accuracy. The results show that the approach of deep learning model in the detection and classification of high voltage insulators was good. If any test image is applied at inspection unit, a warning message is sent to the utility centre through the Blynk server for every bad class of insulator. Further, the suggested model can be implemented with unmanned aerial vehicles (UAV) instead of cameras as an image source. The authors want to use generative adversarial networks to produce synthetic pictures in the future study. As a result, accuracy can be enhanced further.

Table 3 Comparison of state-of-the-art-models

Model	Sensitivity	Specificity	Prediction time in ms	Accuracy	mAP
Faster-RCNN [21]	90.3	92.7	554.25	90.5	81.2
SSD [20]	92.4	94.5	320.25	94.1	93.4
YOLOv2 [22]	95.6	97.1	280.50	96.4	96.6
YOLOv3 (**Proposed system**)	**99.4**	**99.8**	**254.2**	**98.9**	**100**

Bold values represent the suggested method’s dominance over other state-of-the-art models

References

1. Nguyen, V.N., Jenssen, R., Roverso, D.: Automatic autonomous vision-based power line inspection: a review of current status and the potential role of deep learning. Int. J. Electr. Power Energy Syst. **99**, 107–20 (2018)
2. Anjum, S., Jayaram, S., Jahromi, A.N., et al.: Detection and classification of defects in ceramic insulators using RF antenna. IEEE Trans. Dielectr. Electr. Insul. **24**(1), 183–90 (2017)
3. Marimuthu, K., Vynatheya, S., Vasudev, N., et al.: Quality analysis of ceramic insulators under electro thermal stresses. In: International Conference on High Voltage Engineering and Technology, pp. 1–6. IEEE (2019)
4. Gencoglu, M.T., Uyar, M.: Prediction of flashover voltage of insulators using least squares support vector machines. Expert Syst. Appl. **36**(7), 10789–98 (2009)
5. Yang, L., Zhang, F., Hao, Y., et al.: Effects of structure and material of polluted insulators on the wetting characteristics. IET Sci. Meas. Technol. **13**(2), 131–8 (2018)
6. Montoya, G., Ramirez, I., Montoya, J.I.: Correlation among ESDD, NSDD and leakage current in distribution insulators. IEE Proc. Gener. Transm. Distrib. **151**(3), 334–40 (2004)
7. Iwai, K., Hase, Y., Nakamura, E., et al.: Development of a new apparatus for contamination measurement of overhead transmission line insulators. IEEE Trans. Power Deliv. **13**(4), 1412–7 (1998)
8. Vaillancourt, G.H., Bellerive, J.P., Jean, C., et al.: New live line tester for porcelain suspension insulators on high-voltage power lines. IEEE Trans. Power Deliv. **9**(1), 208–19 (1994)
9. Wong, K.L.: Application of very-high-frequency (VHP) method to ceramic insulators. IEEE Trans. Dielectr. Electr. Insul. **11**(6), 1057–64 (2004)
10. Ouyang, X., Jia, Z., Yang, S., et al.: Influence of algae growth on the external insulation performance of HVDC insulators. IEEE Trans. Dielectr. Electr. Insul. **25**(1), 263–71 (2018)
11. Zhong, D., Yuan, F., Yao, J., et al.: Application analysis of ultrasonic technology in insulator contamination discharge detection. In: 2019 International Conference on Computer, Network, Communication and Information Systems. Atlantis Press (2019)
12. Gencoglu, M.T., Cebeci, M.: Investigation of pollution flashover on high voltage insulators using artificial neural network. Expert Syst. Appl. **36**(4), 7338–45 (2009)
13. Khan, S., Islam, N., Jan, Z., et al.: A novel deep learning based framework for the detection and classification of breast cancer using transfer learning. Pattern Recognit. Lett. **1**(125), 1–6 (2019)
14. LeCun, Y., Bengio, Y., Hinton, G., et al.: Deep learning. Nature **521**(7553), 436 (2015)
15. Bengio, Y., Courville, A., Vincent, P.: Representation learning: a review and new perspectives. IEEE Trans. Pattern Anal. Mach. Intell. **35**(8), 1798–828 (2013)
16. Malon, C.D., Cosatto, E.: Classification of mitotic figures with convolutional neural networks and seeded blob features. J. Pathol. Inform. **4** (2013)
17. Cruz-Roa, A., Basavanhally, A., Gonzalez, F., et al.: Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks. In: Medical Imaging 2014: Digital Pathology, vol. 9041, p. 904103 (2014)
18. Nasr, E., Samavi, S., Karimi, N., et al.: Melanoma detection by analysis of clinical images using convolutional neural network. In: 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 1373–1376. IEEE (2016)
19. Lopez, A.R., Burdick, J., Marques, O., et al.: Skin lesion classification from dermoscopic images using deep learning techniques. In: 2017 13th IASTED International Conference on Biomedical Engineering (BioMed), pp. 49–54. IEEE (2017)
20. Miao, X., Liu, X., Chen, J., Zhuang, S., et al.: Insulator detection in aerial images for transmission line inspection using single shot multibox detector. IEEE Access **7**, 9945–9956 (2019)
21. Xuefeng, L., Hansong, S., Liu, G.: Insulator defect recognition based on global detection and local segmentation. IEEE Access **8**, 59934–59946 (2020)
22. Diana, S., Pernebayeva, D., Bagheri, M., James, A.: IN-YOLO: real-time detection of outdoor high voltage insulators using UAV imaging. IEEE Trans. Power Deliv. **35**(3), 1599–1601 (2019)
23. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: unified, real-time object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 779–788 (2017)
24. Redmon, J., Farhadi, A.: YOLO9000: better, faster, stronger. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 7263–7271 (2017)
25. Pengcheng, X., Guan, H., Shu, C., Borgeat, L., Goubran, R.: An integrated approach for medical abnormality detection using deep patch convolutional neural networks. Vis. Comput. **36**(9), 1869–1882 (2020)
26. Redmon, J., Farhadi, A.: Yolov3 an incremental improvement. arXiv preprint arXiv:1804.02767 (2018)
27. Yicheng, L., Liao, L., Hao, W., Qin, J., He, L., et al.: Trajectory and image-based detection and identification of UAV. Vis. Comput. **37**(7), 1769–1780 (2021)
28. Chen, W., Huang, H., Peng, S., Zhou, C., Zhang, C.: YOLO-face a real-time face detector. Vis. Comput. **37**, 1–9 (2020)
29. Lin, T.Y., Dollár, P., Girshick, R., He, K., Hariharan, B., Belongie, S.: Feature pyramid networks for object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2117–2125 (2017)
30. Dong, C., Loy, C.C., He, K., Tang, X.: Learning a deep convolutional network for image super-resolution. In: European Conference on Computer Vision, pp. 184–199. Springer (2014)
31. Nair, V., Hinton, G.E.: Rectified linear units improve restricted Boltzmann machines. In: Proceedings of the 27th International Conference on Machine Learning (ICML-10), pp. 807–814 (2010)
32. Zhao, J., Qu, J.: Healthy and diseased tomatoes detection based on YOLOv2. In: International Conference on Human Centered Computing, pp. 347–353. Springer, Cham (2018)
33. Ghenescu, V., Mihaescu, R.E., Carata, S.V., Ghenescu, M.T., Barnoviciu, E., Chindea, M.: Face detection and recognition based on general purpose DNN object detector. In: 2018 International Symposium on Electronics and Telecommunications (ISETC), pp. 1–4. IEEE (2018)
34. Szegedy, C., Ioffe, S., Vanhoucke, V., Alemi, A.A.: Inception-v4, inception-resnet and the impact of residual connections on learning. In: Thirty-First AAAI Conference on Artificial Intelligence (2017)
35. Rao, C., Fan, Y., Su, K., Latecki, L.J.: Common object discovery as local search for maximum weight cliques in a global object similarity graph. In: International Conference on Discrete Geometry for Computer Imagery, pp. 219–233. Springer, Cham (2019)
36. Kingma, D.P., Ba, J.: Adam: a method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2019)
37. Ruder, S.: An overview of gradient descent optimization algorithms. arXiv preprint arXiv:1609.04747 (2016)

Dipu Sarkar received his B.Tech. in Electrical Engineering in 2003 from the University of Kalyani, W.B., India. He received his M.Tech. with specialization in electrical power systems from the University of Calcutta, India, in 2007 and his Ph.D. from the Department of Electrical Engineering, Bengal Engineering and Science University, Shibpur, India (presently known as Indian Institute of Engineering Science and Technology [IIEST] in 2013. Currently, he is an assistant professor in the Department of Electrical and Electronics Engineering of National Institute of Technology, Nagaland, India. His fields of interest are power systems operation and control, power systems stability, soft computational applications in power systems and smart grids, grid integrated renewable energy, service restoration and protective relay coordination.

Sravan Kumar Gunturi received his B.Tech. degree from Sasi Institute of Technology and Engineering Tadepalligudem, Andhra Pradesh, India, in Electronics and Communication Engineering in 2009. He received his M.Tech. degree with specialization in embedded systems from BVC Engineering College, Odalarevu, India, in 2012 and is pursuing his Ph.D. from the Department of Electronics and Instrumentation Engineering, National Institute of Technology, Nagaland, India. His areas of interest are machine learning, deep learning, and the Internet of Things.