Development of intelligent outdoor camera sabotage detection system for large-scale camera systems using Deep Learning

KAUR Marapana[1], SMKK Siriwardhana[1], MS Dunukewila[1], HKAYD Kodithuwakku[1#], and TL Weerawardane[1]

[1]*Faculty of Engineering, General Sir John Kotelawala Defence University, Sri Lanka*
[#] <37-eng-0044@kdu.ac.lk >

Abstract **-** *The study presents a newly created camera-tampering detection system for outdoor cameras, aiming to overcome the boundaries of human monitoring. It is intended to be implemented in large scale camera systems to identify frequent tampering events like defocus, occlusion and changes in orientation, and provide real time alerts and visual feedback through a user friendly web portal designed especially for this purpose. The system can effectively recognize and categorize tampering instances by utilizing deep learning algorithms, which reduces dependency on human operators and lowers the risk of human mistake. To detect and categorize tampering, three algorithms are utilized, and the features of each algorithm are listed. Security staff can take the necessary measures to stop potential security breaches or the loss of important surveillance footage by quickly identifying tampering occurrences. The suggested method strengthens the monitoring process's dependability and efficiency, which in turn strengthens the security of the outdoor surveillance infrastructure*

Keywords **- artificial intelligence, machine learning, deep learning, camera sabotage detection**

I. INTRODUCTION

The security environment has seen a revolution in recent years due to rapid growth in surveillance technology, which has resulted in the widespread deployment of outdoor CCTV camera systems. These devices serve as crucial tools for boosting security precautions and preventing criminal activity in a variety of outdoor settings. Outdoor CCTV camera systems offer significant insights and proof that can support investigations and maintain public safety by continuously monitoring and surveillance.

As camera networks expand, Maintaining the integrity and functionality of these systems raises different challenges. The effectiveness of outside surveillance can be greatly impacted by tampering events such as defocus, occlusion, and orientation shifts. For the camera network to be reliable and to guard against future security breaches, it is crucial to quickly identify and handle such tampering incidents. Traditional surveillance systems mainly rely on human monitoring and visual inspection to find any tampering efforts on outdoor cameras. However, since security staff are not always watching the cameras and because of having a large number of cameras, these systems are frequently unable to immediately identify and respond to tampering efforts.

This project attempts to create an AI-based outside camera tampering monitoring system to address this problem. The system will be able to recognize and notify security workers of tampering incidents in real-time by utilizing artificial intelligence and real-time analysis. The project aims to develop a novel solution that improves the overall security and dependability of outdoor surveillance infrastructure with an emphasis on implementing the system in a large-scale outdoor CCTV camera network.

II. LITERATURE REVIEW

The paper (Ribnick et al., 2006) describes a real-time technique for detecting camera tampering developed for use in surveillance and security applications. The technique detects large differences between older frames of video and more recent frames by comparing the frames using three different measures of image dissimilarity. The method then normalizes the frames and tests a set of conditions to determine if camera tampering has occurred (Ribnick et al., 2006).

So, in this method, a live video is received by the program, and it is stored in two different buffers. The first buffer is labelled as the short-term pool, and stores frames that are less than 10-50 seconds old. The second buffer is the long-term pool, which stores frames until they are an additional two minutes old. Each time a new frame is pushed into the short-term pool, the short-term and long-term pools must be compared in order to determine if camera tampering has occurred. Using a generic measure of image dissimilarity, every frame in the short-term pool is compared to every frame in the long-term pool. Histogram chromaticity difference, Histogram L1R difference and Histogram gradient direction difference are a variety of measures used here in order to extract the most useful information when Measuring image dissimilarity. When detecting the camera tampering, each time a new frame is pushed into the short-term pool, the three dissimilarity measures described above are computed. Based on these three values the decision is made as to whether tampering has occurred by evaluating a

set of conditions, each of which is given by a set of three thresholds. For each condition, if the three thresholds are exceeded, the condition is evaluated as true. If any of the conditions are true, the decision is made that tampering has occurred. Then this algorithm was tested on an extensive set of test videos totaling over 19 hours in duration. And, this method is insensitive to changes in illumination, small camera movements, and transient non-tampering events(Ribnick et al., 2006).

The paper "Automatic Control of Video Surveillance Camera Sabotage" describes the main kinds of sabotage that could be done to a video surveillance system. Such as, Image occlusion, Image defocus and change of the field of view. It also describes how these sabotages could be affected by a video surveillance system while triggering an alarm when a sabotage has been detected (Gil-Jiménez et al., 2007).

Basically, edges are used in this paper to detect whether camera sabotage occurs or not because edges are more robust against illumination changes than smooth parts of the image (Gil-Jiménez et al., 2007). When an object covers the lens of the camera, part of the image is no longer visible, and a number of pixels of the background model disappear as well. So, the entropy of the pixels belonging to the background model is computed. A sabotage will be detected when the current entropy for at least one of these blocks is lower than a given threshold (Gil-Jiménez et al., 2007). The defocus of the camera can be detected with the help of degradation of the edges. Therefore, the number of pixels in the background model can be easily counted for each new frame. If the camera has been defocused, this number for the current frame will be much lower than that of the background model (Gil-Jiménez et al., 2007). The detection of shifts on the field of view is computed using the zero-mean normalized cross-correlation (ZNCC). In this case, only the pixels of the background model is computed to speed up the algorithm (Gil-Jiménez et al., 2007).

Camera tampering is a serious issue in video surveillance, and it can be classified into three types: covered, defocused, and moved. These tampers can be detected by using deep neural networks for classification and recognition tasks, researchers have hypothesized that global image features can be used to classify normal and tampered images. Recent architectures, such as Alexnet, Resnet, and Densenet, have shown promise in scene classification, with Alexnet performing the best among the four models with an accuracy of 75%(Mantini and Shah, 2019).

Digital images are widely used and shared on various social media platforms, making it difficult to differentiate between real and manipulated images. To address this issue, this study proposes an approach based on ResNet50v2, a state-of-the-art deep learning architecture. The model takes image batches as input and utilizes the YOLO convolutional neural network's weights by using the architecture of ResNet50v2. The use of transfer learning enables more efficient training of the model, reducing its complexity and training time. Assigning meaningful weights to initialize the proposed model increases its efficiency(Qazi, Zia and Almorjan, 2022).

The paper "Automated camera sabotage detection for enhancing video surveillance systems" proposes a new method for detecting video camera tampering events like occlusion, Blurry images and displacement of camera. Camera occlusion is detected by making use of the foreground objects' area and position. Blurry images are detected using edge information and camera displacement is detected using frame count per panning sweep. The effectiveness of the proposed method is tested using a public dataset for camera sabotage detection on panning surveillance systems by comparing it with an existing method. The results show that the proposed method is superior to the existing method with reduced false alarms and has equal performance in terms of true positive detection. The performance of the proposed method is also evaluated on static surveillance systems by comparing it with six other existing methods on a public dataset and the results obtained are encouraging. Also, the proposed method automatically detects routine problems with cameras like dirt on the camera lens, fog and smoke(Sitara and Mehtre, 2019).

The study(Huang et al., 2014) proposes an automated method for rapidly detecting camera tampering and various abnormalities for a video surveillance system. The proposed method is based on the analyses of brightness, edge details, histogram distribution, and high-frequency information, making it computationally efficient. The proposed system runs at a frame rate of 20–30 frames/s, meeting the requirement of real-time operation. Experimental results show the superiority of the proposed method with an average of 4.4% of missed events compared to existing works(Huang et al., 2014).

III. PROBLEM STATEMENT

Video and photo cameras are frequently used for outdoor monitoring of objects or areas, but they are not always monitored regularly(Nicodeme, 2021). This means that feedback is not provided to monitoring personnel in real time. This lack of supervision can result in a variety of error cases, particularly in remote areas. Errors such as camera defocusing due to dust, raindrops, and fog, as well as occlusion caused by objects such as paint, bags, and dirt, and orientation changes due to animal activity, can all occur. As a result, there is currently no effective system available to detect these types of errors and inform authorities(Scholtz and Ngxande, 2022).

The aim of the project is to develop a CNN model capable of detecting commonly occurring camera sabotage errors, such as defocusing, occlusion, and orientation changes, and to provide real-time notifications to the appropriate personnel responsible for monitoring the outdoor areas and objects.

First, it is planned to detect tampering errors using OpenCV. Next, collecting data is required to train the model. Python is the expected programming language to code the program. To achieve the aforementioned objectives, it is proposed to build a model using convolutional neural networks. Tuning and training are done to minimize output error. After minimizing the error to the least the model is uploaded to the Jetson Nano.

The primary goal is to create a system capable of detecting camera sabotage incidents. This includes detecting occlusion, defocusing, and camera orientation changes using a CNN(Zhan et al., 2017). When tampering occurs, the system will provide real-time feedback via a web portal accessible to the appropriate personnel. The system will also include a mountable enclosure to ensure that the camera remains in place and is not subject to vandalism or other forms of tampering. This system is an Intelligent device. It is capable of categorizing each sabotage into one of three categories (defocusing, occlusion and orientation changes). Notably, many potential tampering events can be included into one of these categories, ensuring that the system remains adaptable and responsive to various types of security breaches. By achieving these goals, the system will help to ensure that the camera remains operational and capable of reliably and effectively monitoring outdoor areas and objects. This can lead to better outcomes to improve surveillance practices in remote locations and provide a practical and efficient technique for camera operators.

IV. SCOPE / LIMITATIONS

The scope of this project is to develop a camera sabotage detection system for large-scale camera monitoring system, that can identify cameras even when the view is covered by a bag or paint and under various weather conditions. The system will also detect possible problems and give real-time feedback to the operator.

The system has limitations in detecting sabotage incidents at night due to reliance on visual information and inadequate illumination. Additionally, real-time feedback requires a Wi-Fi connection, which may restrict its usage in areas with limited connectivity.

V. METHODOLOGY

A. Framework and methodology of the study design

In a CCTV system, it is a must to have regular monitoring processes; otherwise, it is of no use. However, it may depend on the purpose of the user. A person is assigned to monitor the system. As mentioned above, there are some issues, so a proper solution to optimize the CCTV monitoring system is provided and it makes the person's task easier. This system is a development of the existing CCTV camera system, which functions normally. Here, the main focus is on outdoor cameras. The live footage is taken to the system, and that data is processed to predict errors. Then, the user can know if an error has occurred in the camera via the web portal. The data is stored in a cloud database, which can be accessed by the system and the web portal. This system sequentially checks each camera in the large camera system and provides real-time feedback to the user via the web app.

The live footage is fed into the proposed design, and the view can be monitored through a monitor, as is the general way for the user. A model is running in the system to detect and identify errors. Primarily, sabotage, blurry footage, and changes in camera orientation are detected by the system(Huang et al., 2014). An AI-trained model is deployed on a single-board computer to detect sabotage. Firstly, it detects whether there is tampering with the camera. Secondly, it identifies the type of error, and a snapshot of the error is pushed to the cloud database. This data is then sent to the web application. The user is able to see the errors related to the cameras via the web app. Moreover, the web app is user-friendly and built with an attractive user interface, making it easy for the user to check if the cameras are sabotaged. There is no need to waste time and money on CCTV monitoring. Additionally, it aids in the maintenance process of the camera system.

B. Block diagram of the proposed design

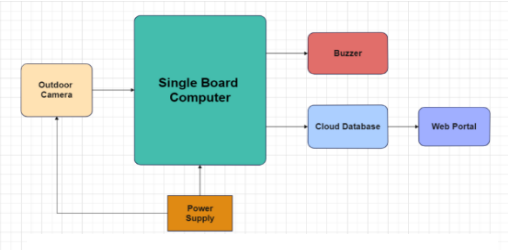

Figure 1. Block diagram of the proposed design
Source: Generated by the authors

Figure 1 shows the block diagram of the proposed design framework. The NVIDIA Jetson Nano board is used as the single-board computer to perform computational tasks and run our designed algorithm. One of the key strengths of the Jetson Nano is its AI computing capabilities. It supports well-known AI frameworks like TensorFlow, PyTorch, and Caffe, making it suitable for applications such as image

classification, object identification, and natural language processing. By accelerating the execution of certain AI workloads on the Jetson Nano, real-time performance is made possible even at the edge.

The camera feed is provided to the Jetson board through the DVR using an HDMI to USB converter as the input. The data is processed inside the Jetson board, and the output is sent to the cloud database (Firebase). Additionally, a buzzer is used as an output component to notify when tampering has occurred. The information about the cameras and the detected errors can be visualized through the web app. Only authorized users can log in to the system.

General case: The existing system for CCTV camera monitoring is improved by the proposed design. The proposed design is applicable to large-scale CCTV systems. This device can be connected to an existing camera system using necessary video converters. Authorized users are registered through the web app, and an admin is assigned to the system. Only the admin can create new users.

Prototype for KDU: To build the prototype of the proposed design framework, several outdoor cameras, a Jetson board, converters, DVR, and monitors will be used. A suitable location on the KDU premises is selected for the demonstration.

VI. PROTOTYPE DESIGN AND SYSTEM ARCHITECTURE

A. Image Processing

Initially, the code was developed using image processing that can identify a generalized tampered error. which doesn't identify a particular error. In order to develop the code, an open computer vision library was used. The code uses a background subtraction algorithm to identify moving objects in the video. The bounding boxes of the moving objects are then calculated after being tracked. Then, each bounding box's surface area is determined and added to a running total(Sitara and Mehtre, 2016). The system will emit an alert indicating that tampering has been discovered if the running total goes beyond a predetermined threshold.

B. Transfer learning

Mainly, 2 pre-trained models were used where the weights are trained on the famous "ImageNet" dataset. MobileNetV2 and DenseNet121 were the pre-trained model architectures used for detecting tampering errors. The input layer of the MobilnetV2 was omitted and changed according to the size and channels of the input dataset. Moreover, the output layers were popped out and changed the nodes to 4 since 4 classes will be predicted. A flattened layer was constructed before the final nodes. In the training procedure, all the layers were frozen from training except for the final layer. Initial and intermediate layers have the trained weights and biases according to the ImageNet dataset. The same procedure was performed in DenseNet121 as well in order to compare which architecture performs better.

C. Hardware setup

In order to practically implement the model, Nvidia Jetson Nano was chosen. First, the operating system was booted using the Ubuntu OS file(Ivanov and Yudin, 2019). Then the required Python libraries were installed. A buzzer alarm was set up at the operator to alarm the user in case of a tampering incident. There will be two outputs from the 8-channel quad processor. One output is sequentially fed to the Jetson while the other output will be displayed on the TV monitor all cameras together.

D. Webpage development

The Webpage is used to give real-time feedback to the user. The website is accessed by entering the Uniform Resource Locator (URL). The following building blocks were used to create the website. As shown in Table 1, the following are the main building blocks of the web and their functions

Table 1. Building block of web and functions

Building block of the web	Function
Hyper Text Markup Language (HTML)	To structure the web page and its content
Cascading Style Sheets (CSS), Bootstrap	To style the webpage
Hyper Text Processor (PHP)	To handle server-side scripting
JavaScript, AJAX	To make the web page interactive on both client-side and server-side

Source: Generated by the authors

The web development section was divided into 2 phases. That is the front end and the back end. The term "front-end" refers to the user interface, while "back-end" means the server, application and database that work behind the scenes to deliver information to the user.

VII. DATA AND RESULTS ANALYSIS

After training both DenseNet and MobileNet, the training loss, validation loss, training accuracy and validation accuracy for each architecture were evaluated. Based on the graphs on epochs, decisions were taken on how to improve the hyperparameters of each architecture. Initially, the model was trained only for 2 classes: obstruction and normal conditions.

As shown in Figure 2, the loss function Binary Cross Entropy wasn't a good fit. 1970 of the total amount of data with a batch size of 32 was used in the training process. "Adam" was used as the optimizer with a learning rate of 0.0001. Insufficient model capacity and overfitting may be the issues of the below curves.

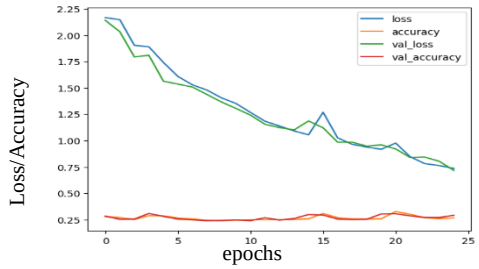

Figure 2. losses and accuracies of the Densenet architecture with Binary Cross Entropy

Source: Generated by the authors

It is seen in Figure 3 that by changing the loss function to Sparse Categorical Cross Entropy the accuracies and losses were in the range. The “Adam” optimizer was used with the learning rate of 0.0001. There were 1576 of training data and 394 of validation data with a batch size of 32.

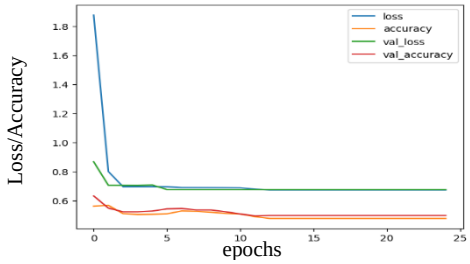

Figure 3. Losses and accuracies of the Densenet architecture with Sparse Categorical Cross Entropy

Source: Generated by the authors

As shown in Figure 4, huge fluctuations were observed when using the loss function, Binary Cross Entropy. “Adam” optimizer was used with a learning rate of 0.0001. 1970 total number of images were taken as data inputs.

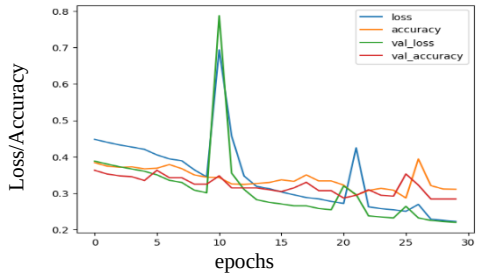

Figure 4. Losses and accuracies of the MobileNet architecture with Binary Cross Entropy

Source: Generated by the authors

A significant improvement is observed in Figure 5. There also a batch size of 32 with a total images of 1970 were taken as inputs. The optimizer was “Adam” with a learning rate of 0.0001.

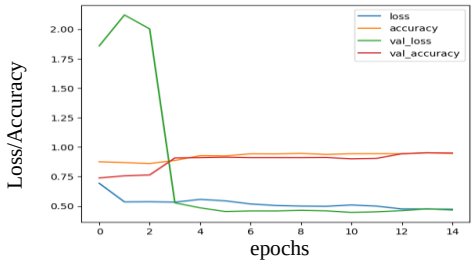

Figure 5. Losses and accuracies of the MobileNet architecture with Sparse Categorical Cross Entropy

Source: Generated by the authors

The above pictures show the tuning of hyperparameters related to the loss function. In Figures 2 and 4, binary cross entropy was used as the loss function and in Figures 3 and 5, Sparse Categorical Function was used as the loss function. Able to gain higher accuracy and lower loss by changing the loss to Sparse Categorical Cross Entropy.

The used pre-trained model here was trained on a multi-class classification task (ImageNet). Therefore, in order to get higher accuracy Sparse Categorical Cross Entropy needs to be used. Binary cross entropy loss may be more sensitive to noise in data. In the image processing code, the tampering was successfully detected without any errors. In some cases, the sparse categorical cross entropy loss function may work better than the binary cross entropy loss function. This is because the sparse categorical cross entropy loss function takes into account the fact that the ground truth labels are integers. This can be helpful when the model's predictions are not close to the ground truth labels, but they are still close to each other

Figure 6. Tampering detection

Source: Generated by the authors

To assess the effectiveness of the proposed model, the accuracies of the industry benchmarks were compared (Mantini and Shah, 2019). As presented in Table 2, the average of accuracies (training and validation) were considered.

Table 2. The comparison of Accuracies

	Accuracy for normal and occlusion conditions
Densenet161	68.5%
Proposed Densenet Model	69%
Alexnet	87%
Proposed Mobilenet Model	95%

Source: Generated by the authors

VIII. CONCLUSION

In conclusion, this study introduces a novel camera-tampering system designed specifically for outdoor cameras, with the aim of enhancing the capabilities of human monitoring. By leveraging deep learning algorithms, the system effectively identifies and categorizes various tampering events such as defocus, occlusion, and changes in orientation. This automated approach reduces reliance on human operators and minimizes the risk of human error. The system provides real-time alerts and visual feedback through a user-friendly web portal, enabling security staff to promptly respond to potential security breaches or the loss of surveillance footage. By enhancing the dependability and efficiency of the monitoring process, this proposed method strengthens the overall security of outdoor surveillance infrastructure.

IX. FUTURE WORK

It is advised to use data fusion techniques and make use of edge computing capabilities to improve the performance of the trained model installed on the Jetson board. Further insights can be gained by combining data from various sources, including sensors or cameras, which improves accuracy and resilience. The model's capabilities can also be improved by including different methods for tracking human actions and behaviours, which enables real-time analysis and pattern recognition. The system's overall efficiency can be considerably increased by these optimizations. Our upcoming study will carefully assess how well our model performs in challenging conditions including heavy rain, fog, and snow. To determine the accuracy and robustness of this assessment, controlled experiments will be used. The addition of specialized datasets and the use of fine-tuning techniques will be pursued to increase the model's applicability under difficult weather circumstances.

REFERENCES

Gil-Jiménez, P. *et al.* (2007) 'Automatic control of video surveillance camera sabotage', *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 4528 LNCS(PART 2), pp. 222–231. doi:10.1007/978-3-540-73055-2_24.

Huang, D.Y. *et al.* (2014) 'Rapid detection of camera tampering and abnormal disturbance for video surveillance system', *Journal of Visual Communication and Image Representation*, 25(8), pp. 1865–1877. doi:10.1016/j.jvcir.2014.09.007.

Ivanov, A. and Yudin, D. (2019) *Visibility loss detection for video camera using deep convolutional neural networks, Advances in Intelligent Systems and Computing*. Springer International Publishing. doi:10.1007/978-3-030-01818-4_43.

Mantini, P. and Shah, S.K. (2019) 'UHCTD: A comprehensive dataset for camera tampering detection', *2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance, AVSS 2019*, 1(d), pp. 1–8. doi:10.1109/AVSS.2019.8909856.

Nicodeme, C. (2021) 'Detection of Defective Videosurveillance Camera in Train Stations', *International Conference on Control, Automation and Systems*, 2021-Octob(Iccas), pp. 1185–1189. doi:10.23919/ICCAS52745.2021.9649890.

Qazi, E.U.H., Zia, T. and Almorjan, A. (2022) 'Deep Learning-Based Digital Image Forgery Detection System', *Applied Sciences (Switzerland)*, 12(6). doi:10.3390/app12062851.

Ribnick, E. *et al.* (2006) 'Real-time detection of camera tampering', *Proceedings - IEEE International Conference on Video and Signal Based Surveillance 2006, AVSS 2006*, pp. 10–15. doi:10.1109/AVSS.2006.94.

Scholtz, T. and Ngxande, M. (2022) 'Robust Anomaly Detection in CCTV Surveillance', *EPiC Series in Computing*, 85, pp. 104–128. doi:10.29007/dfjs.

Sitara, K. and Mehtre, B.M. (2016) 'Real-time automatic camera sabotage detection for surveillance systems', *Advances in Intelligent Systems and Computing*, 425, pp. 75–84. doi:10.1007/978-3-319-28658-7_7.

Sitara, K. and Mehtre, B.M. (2019) 'Automated camera sabotage detection for enhancing video surveillance systems', *Multimedia Tools and Applications*, 78(5), pp. 5819–5841. doi:10.1007/s11042-018-6165-4.

Zhan, Y. *et al.* (2017) 'Image forensics based on transfer learning and convolutional neural network', *IH and MMSec 2017 - Proceedings of the 2017 ACM Workshop on Information Hiding and Multimedia Security*, pp. 165–170. doi:10.1145/3082031.3083250.

AUTHOR BIOGRAPHIES

Roshan Marapana is an Electronic and Telecommunication engineering undergraduate in the Department of Electrical, Electronic and Telecommunication Engineering at General Sir John Kotelawala Defence University. His research interests include Machine learning, Deep learning and Data science. In the aforementioned fields, Roshan has completed initiatives. He can be contacted at udirosh1957@gmail.com for further collaborations and discussions.

Kasun Siriwardhana is an Electronic and Telecommunication engineering undergraduate in the Department of Electrical, Electronic and Telecommunication Engineering at General Sir John Kotelawala Defence University. His areas of interest in research include artificial intelligence, the internet of things, VLSI, photonics, and power electronics.

In the aforementioned fields, Kasun has completed initiatives. He can be contacted at kasunsiriwardhana22@gmail.com for further collaborations and discussions.

Malsha Dunukewila is an Electronic and Telecommunication engineering undergraduate in the Department of Electrical, Electronic and Telecommunication Engineering at General Sir John Kotelawala Defence University. Her areas of interest in research include Artificial Intelligence, Internet of things, Electronic and Telecommunication. Malsha has finished projects in the aforementioned fields. She can be reached at malshasewd@gmail.com for additional discussions and collaborations.

Yehan Kodithuwakku is an Electronic and Telecommunication engineering undergraduate in the Department of Electrical, Electronic and Telecommunication Engineering at General Sir John Kotelawala Defence University. His research interests include Artificial Intelligence, Electronics and Telecommunication. Yehan has published several papers in renowned conferences and journals. He can be contacted at yehankodithuwakku@gmail.com for further collaborations and discussions.

Prof. TL Weerawardane is a distinguished researcher at the General Sir John Kotelawala Defence University. With expertise spanning 3G/4G/5G Mobile Communication, Wireless Communication, and Telecommunication Networks, he also delves into cutting-edge fields like Industrial IoT, Data Science, Big Data Analytics, AI, and Cybersecurity. His pioneering publications have significantly advanced the understanding of modern communication systems and their intersection with emerging technologies. He can be contacted at tlw@kdu.ac.lk for further information.