6G Connected Vehicle Framework to Support Intelligent Road Maintenance using Deep Learning Data Fusion

Mohammad Hijji, Rahat Iqbal *Member, IEEE*, Anup Kumar Pandey, Faiyaz Doctor *Senior Member, IEEE*, Charalampos Karyotis, Wahid Rajeh, Ali Alshehri, Fahad Aradah

***Abstract*— The growth of IoT, edge and mobile Artificial Intelligence (AI) is supporting urban authorities exploit the wealth of information collected by Connected and Autonomous Vehicles (CAV), to drive the development of transformative intelligent transport applications for addressing smart city challenges. A critical challenge is timely and efficient road infrastructure maintenance. This paper proposes an intelligent hierarchical framework for road infrastructure maintenance that exploits the latest developments in 6G communication technologies, deep learning techniques, and mobile edge AI training approaches. The proposed framework abides with the stringent requirements of training efficient machine learning applications for CAV, and is able to exploit the vast numbers of CAVs forecasted to be present on future road networks. At the core of our framework is a novel Convolution Neural Networks (CNN) model which fuses imagery and sensory data to perform pothole detection. Experiments show the proposed model can achieve state of the art performance in comparison to existing approaches while being simple, cost-effective and computationally efficient to deploy. The proposed system can form part of a federated learning framework for facilitating large scale real-time road surface condition monitoring and support adaptive resource allocation for road infrastructure maintenance.**

***Index Terms*— 6G, Deep Learning, Mobile Edge Intelligence, Pothole Detection, Federated Learning, Intelligent Transportation Systems.**

I. Introduction

In the near future we are set to see far greater numbers of CAVs on our roads. Additionally, IoT and Edge based AI technologies provide the promise of delivering the next generation of proactive Intelligent Transport System applications. These applications will provide real-time decision-making support and optimized resource allocation. At the same time, 6G communication technologies will form the necessary network backend to accommodate the needs arising from the application of these technologies namely hyper-fast data rates; ultralow latency; high reliability; and more secure communication. One of the most important smart city tasks is securing the presence of a well-maintained road network. A critical aspect towards achieving this goal is the effective detection and timely repair of potholes. In order to provide an intelligent and cost-efficient solution, a large deployment of sensors, the corresponding communication protocols and data processing mechanisms are required to be developed and deployed.

Potholes are a well-documented nuisance for road users and a challenge for authorities. A reader complained to the New York Times in 1910 that “a steady succession “of potholes was rendering travel “a burden rather than a joy” [1]. Potholes endanger road users and may cause significant harm to vehicles and drivers. The expense of repairing potholes is high and need special budgeting. According to an AA survey, almost 61 million drivers had their cars affected in some way from potholes. The expense of repairing these damages was nearly $684 million in a single year [2]. The number of potholes filled in England every year, according to the report, is nearly 16000 with 4099 just in London alone. It costs nearly 100 million and 11 million GBP, in England and London respectively, to repair these potholes. According to the United Kingdom’s Department of Transport, a third of all local roads in England are now in need of urgent maintenance and repair based on data collected between 2020 and 2022. As stated in a report commissioned by the House of Commons, the pothole repair work in the British road network would take nearly 14 years to complete [3]. Road users want pothole-free roads to commute, and maintenance of roads and highways is vital for effective traffic management. Maintaining such a vast network of roads necessitates both expertise and funds. Over the years, specialists surveyed roads and the accountable authorities then repaired these roads. These procedures were labour-intensive, expensive, and time-consuming [4]. This method is also unable to meet the continued demand to keep roads in good condition. There can be severe delays to repair road damages which authorities blame on the shortage of human resources and funding to maintain an ever-expanding road network and growing road users. The inspection carried out by experts depend on their visual perception and may lack consistency. The effects of climate change are also only expected to worsen this situation due to the results of extreme weather phenomena such as extreme temperatures and floods to the road surface condition.

Mohammad Hijji is with the University of Tabuk, Kingdom of Saudi Arabia, (email: m.hijji@ut.edu.sa).
Rahat Iqbal is with the College of Engineering and IT, University of Dubai, UAE. (e-mail: riqbal@ud.ac.ae).
Anup Kumar Pandey is with Interactive Coventry Ltd., Coventry CV1 2TT, U.K (e-mail: a.kumar@interactivecoventry.com).
Faiyaz Doctor is with the Centre for Computational Intelligence, School of Computer Science and Electronic Engineering, University of Essex, U.K. (email: fdocto@essex.ac.uk).
Charalampos Karyotis is with Interactive Coventry Ltd., Coventry CV1 2TT, U.K (e-mail: charalampos.karyotis@interactivecoventry.com).
Wahid Rajeh is with the University of Tabuk, Kingdom of Saudi Arabia, (email: wahid.ra@ut.edu.sa).
Ali Alshehri is with the University of Tabuk, Kingdom of Saudi Arabia, (email: a.alshehri@ut.edu.sa).
Fahad Aradah is with the University of Tabuk, Kingdom of Saudi Arabia, (email: 411010386@sut.ut.edu.sa).

To keep pace with these demands, it is necessary to start using advanced sensors and automated solutions. Several sensory inputs have been proposed towards this purpose. However, utilizing specific input signals, such as image or accelerometer data alone, bears significant limitations. Several studies propose using artificial neural networks to exploit these inputs in order to inspect and identify damage to the road surface [5-7]. However, effective Machine Learning (ML) models need datasets of adequate volume and quality for training [58]. This data collection process would require numerous dedicated recording devices. Nowadays, smartphones are becoming more ubiquitous and can record greater quality imagery, sensory and GPS data. 6G technologies can exploit this opportunity and support data collection from a multitude of road-user smartphones. In addition to the challenge of collecting adequate data for training reliable ML solutions, recent review studies have also pointed out privacy in data analytics as a major challenge for smart city applications. The smart devices used for data collection continuously capture sensitive data such as faces, license plates and others [56]. This is especially true for pothole detection applications that can use a variety of data including image, video and sensory data.

As demonstrated by recent research efforts Deep Learning (DL) methods that rely on diverse input data from multiple devices are subject to several challenges such as increased latency, energy consumption, network congestion and data privacy issues requiring costly computational and network resources [8]. Nevertheless, modern distributed optimization algorithms such as Federated learning, 6G technologies, and multiple access schemes such as AirComp can accommodate huge numbers of edge devices and dynamically integrate the data coming from these devices [8]. Training and inference using DL methods can be done locally on the edge to maintain data privacy while being updated by the aggregation of global models to improve model accuracy and generalisability. Incorporating these technologies into smart city applications will lead to the development of next generation edge solutions that will be more accurate, efficient, secure and cost effective.

This paper proposes a hierarchical mobile-edge intelligence framework for automated pothole detection that utilizes combinations of diverse input from a variety of sensors. The proposed framework includes a novel DL methodology for pothole detection that fuses together image and sensory data, it utilizes Federated edge AI training to support the continuous large-scale optimization of the developed model and ensure data privacy, and uses smart road signs, 6G and AirComp to secure efficient communications along a road network, support data exchange and warn drivers of hazardous potholes. Public authorities can use the proposed framework for optimising resource allocation for maintaining roads in good condition. It is important to point out that the proposed DL methodology does not require a specialised computer vision device or high computational capability making adoption and scalability very effective. The contributions of this paper are:

- A cost-effective method for collecting sensor data (accelerometer, imagery, and GPS) that uses standard smartphones or dashboard mounted cameras with generic mobile applications for collecting data to train new ML models.
- Development of a state-of-the-art pothole detection DL methodology that fuses sensory and imagery data.
- An intelligent hierarchical framework for automated pothole detection that utilizes federated edge AI training, smart road signs and 6G communication to: support large scale smart city deployment, adaptive resource allocation for maintenance purposes, and overcome training, data security and network communication issues and requirements.

The remainder of the paper is organised as follows. A literature review on state-of-the-art pothole detection methods using computer vision and accelerometer data is covered in section II. Section III discusses the DL pothole detection methodology used in this study. Section IV discusses the experimental results. Section V presents the proposed pothole detection framework that utilizes mobile edge AI, Federated Learning, 6G and Air comp technologies. Section VI concludes the paper with limitations and future work.

II. Literature Review

Road surface anomalies detection has been studied for many years. Overall, there are three approaches to monitor road surfaces: 3D reconstruction, vibration/sensory based and computer vision-based [9]. The 3D reconstruction requires a 3D laser scanner, which scans the surface and makes an accurate model compared to the baseline model to detect anomalies [10]. However, such laser scanners are very costly, and the methods are focused on the local accuracy of the 2D scan [11]. [12] and [13] proposed the polarization method to calculate the difference between horizontal and vertical polarisation. However, the polarisation filters may affect the quality of the images, hence reducing the detection accuracy. For this paper, the focus is on sensory based methods. More specifically on vision and vibration/sensor based methods which are both used by the proposed hybrid model. In the literature which follows we present several examples using these methods and focus on the application of DL, which is also a core component of the proposed methodology.

A. Vision based Methods

Vision-based methods need image processing algorithms to extract texture and then compare the extracted texture with the normal texture to find anomalies. In [7], the researchers used a charge-couple device (CCD) camera mounted on a vehicle to detect defects. The paper suggested using road surface gloss and calculating absolute deviation with reference to the low luminance level. The authors assumed that the low levels of luminance represent the road surface itself, while higher-levels represent the reflection. The paper calculated the deviation of luminance to predict road conditions. Bouilloud et al. used the Interactions between Soil, Biosphere, and Atmosphere (ISBA)-Route/ ‘‘Crocus’’ to predict road conditions. Their model depended on short term forecasts of meteorological conditions and long-term surface conditions simulation [14].

Recent developments in computer vision and computational processing power based on the latest GPUs have facilitated complex computation and DL to automate the detection of road surface defects. Deep Neural Networks (DNN) have gained popularity in the field. DNN are shown to deliver state of the art results in diverse industrial and smart city applications [59][60]. A large, labelled data set, preferably with different datapoints, is required to train deep neural networks-based models to achieve good accuracy [15]. GPU availability and parallel processing computational power have now made it affordable and accessible to develop and train such DNN models for real world industrial applications [16-21]. CNN has been successfully used to recognize images in several research efforts [22-26]. Steinkraus et al. discuss the importance of GPU's for DNN training and shows how they trained one of the largest CNN to date on subsets of the ImageNet dataset. [27].

Farnood et al. discuss the application of neural networks in automation road extraction and vectorized high-resolution images obtained from a satellite. The experiment used backpropagation to train the model with images of size 500x500 pixels [28]. Ren et al. has discussed how the region proposal algorithm and region-based CNNs (R-CNN) have contributed to object detection advances. However, such algorithms could be time-consuming and not economical. The paper presents an efficient method that shares convolutional layers with state-of-the-art object detection networks [29]. Kaiming et al. discussed how CNN's requirement to have a fixed size image may reduce the accuracy of the images or sub-images of an arbitrary size/scale. The paper proposes a spatial pyramid pooling strategy (SPP), to accommodate flexible size images in the networks [30]. In the work by Tsung-Wi et al. a feature pyramids network for object detection is presented [31]. The researchers have developed sliding window prospers (region proposal Network, RPN) [29] in combination with region-based detectors [32]. The proposed architecture leverages the pyramidal shape of convolution network features while creating a feature pyramid by combining low resolution, semantically robust features with high resolution, semantically weak features via top-down pathway lateral connections. Their method produces a feature pyramid with rich semantics at all levels and is built quickly from a single image scale. Handcrafted image features have been replaced with automatically computed features by CNNs. CNNs are robust to the variation in scale and can facilitate recognition from features computed on a single input scale. The method shows improvement over several baseline models [31]. As discussed above, computer vision and DL are being used to classify objects. CNNs have been applied in image classification [25], [33] and object detection [25], [34]. The image-based method is cost-effective in comparison to the 3D laser scan method, however can be sensitive to environmental factors such as light, shadow, water etc.

B. Vibration-based Methods

Vibration-based pothole detection methods use accelerometer data to detect potholes. This method is cost-effective, requires little storage and can be used in real-time [11]. However, the vibration method fails to differentiate between potholes and other forms of anomalies, such as hinges and joints on the road surface [11].

Mednis et al. discuss a vibration method to detect potholes. where data samples were collected using a customized application, and later detection algorithms -Z thresh, Z-peak were applied to find potholes [35]. In the research by Chao et al. the use of ML approaches for road pothole detection using smartphones is discussed. The paper reviews data processing and ML methods such as logistics regression, support vector machine, and random forest to detect potholes using features extracted from collected data [36]. Anguita et al. described how to make a standard human activity recognition dataset using smartphone captured data [37]. The dataset is further used in [38] to recognize human activities using a Support Vector Machine. Recently DL methods such as recurrent neural networks and one-dimension CNNs (1D-CNN) have been used to provide state-of-art results on activity recognition tasks. Huijuan and Jiping have used 1D CNN to extract features that were later fed into an SVM classifier [39]. Lee and Cho propose 1D-CNN for human activity recognition from accelerometer data. The model showed 92.71% accuracy and outperformed other approaches such as random forest [40].

Vibration-based methods can broadly be divided into three categories (1) Threshold-based methods, (2) dynamic time warping (DTW) and (3) ML methods. The threshold-based method detects anomalies when there is a change in amplitude or the signal's other properties across a specified value. Mohan et al. proposed two detectors to detect bumps and potholes. The proposed method was sensitive to speed and was conducted at 25 km/h [41]. The dynamic time rapping (DTM) measures similarities between two sequences which may vary in space and time [42]. Singh et al. have used accelerometer sensor data to detect anomalies using the DTM method. The method produced accuracy in the range of 88% and was not sensitive to speed [43]. Eriksson et al. collected data from a smartphone's sensor installed on a vehicle and used a wide range of filters to identify potholes and other severe road surface anomalies [44].

As demonstrated from the literature traditional methods for identifying potholes are costly and inefficient. Advanced ML methods such as DL are able to exploit large scale sensory data and are being widely used in road surface monitoring. However, computer vision methods are computationally intensive, rely on the availability of large amounts of data and fail in specific image recognition tasks e.g., differentiating between potholes and puddles. Vibration/sensory methods face difficulties to differentiate between potholes and other anomalies and rely on the driver to go over the pothole with the risk of damaging the vehicle and themselves in the process [11].

This paper proposes a unique pothole detection framework that overcomes current challenges concerning the development of effective large scale pothole detection applications. Our framework is able to address data availability issues and provide training datasets of adequate volume and quality by exploiting incoming data from numerous connected vehicles. In addition, it is able to support the network requirements of the bidirectional communication between the road users/IoT devices and the road infrastructure, by exploiting the benefits of 6G communication technologies. Moreover, by incorporating the federated learning approach, our framework supports edge training and secures data privacy. Finally, in order to overcome the weaknesses of computer vision and vibration-based systems, we propose a novel fusion DL model that uses both image and accelerometer data to detect road potholes. This model is a core part of our framework and is described in the following section.

III. Deep Learning Pothole Detection Methodology

Based on the review in section II, it was found that DL can deliver excellent results in pothole detection. In this paper, a fusion model of a two-dimensional Convolution Neural Network (2D-CNN) and a one-dimension Convolution Neural Network (1D-CNN) is used to identify road potholes based on image and accelerometer data. There are three stages in the proposed method: (1) data collection and processing, (2) designing, training and optimization of the hyperparameters and (3) a fusion CNN DL model as depicted in figures 1 and 2.

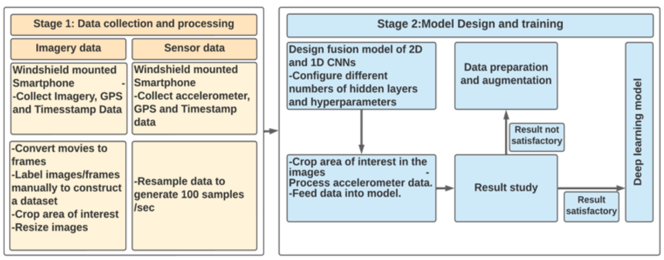


Fig. 1: Methodology Stages 1 and 2: Data collection, Model design, Training and Optimisation

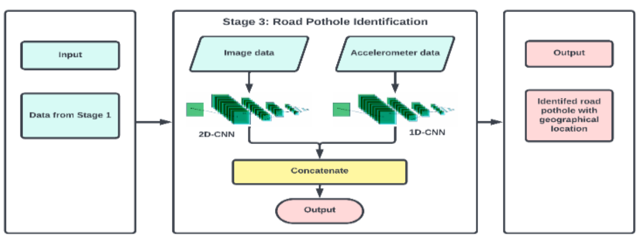


Fig. 2: Methodology Stage 3: Road Pothole Identification

For testing the proposed methodology, we conducted extensive data collection experiments. The collected and pre-processed data sets were used to train custom CNN-based models with a varying number of layers and hyperparameters. The derived models were tested to discover the best model that leads to the least prediction error. In this study, data was collected using an inbuilt camera and a generic sensor data collection application that is available on iOS smartphones.

A. Data Collection

The data has been collected using a smartphone camera and an app installed on the same device. The smartphone was securely placed on the windshield of a vehicle, as shown in Figure 3a. The camera of the smartphone was used to collect images. Simultaneously, an application installed on the same smartphone was used to record 3-Dimensional (X, Y and Z) accelerometer data, along with the corresponding timestamp and GPS information. The sampling rate of the accelerometer was set to 100 Hz. However, due to hardware limitations, the GPS sampling rate was set to 1 Hz. Due to the restriction in the GPS sampling rate, the image sampling rate was also set at 1 Hz. The data was stored on the iOS smartphone and later downloaded to a computer for further analysis. This study's data was collected while driving on motorways, A-roads, B-roads and in town. It was observed that motorways and A-roads were generally in good condition, and most potholes were noticed on B-roads and inside the town. The pothole-data used in this study were collected while driving at a constant speed of 30 miles per hour. The GPS data sampling rate and image sampling rate were both set at 1 Hz. The sensor data sampling rate was set at 100 Hz, which was later merged to create a datapoint for one second. At the mentioned sampling rate, there were 7200 data points recorded. Later, sensory data and imagery data were augmented to increase the size of the data set and achieve better accuracy [15]. Table 1 shows the classes and the number of data samples in each class after augmentation. In the data set, each of the sensory data samples has a corresponding image.

Table I: Road surface categories description and number of samples in the final dataset

Road Quality	Number of Images and sensory data sample
Normal Road Surface	22344
Pothole as seen from the dash-board camera	11016

1) Smartphone Placement: The positioning of the smartphone on the windshield is an essential factor for achieving high accuracy in pothole detection. For this study, the smartphone was securely placed in the middle of the windshield's width and at the highest point as shown in Figure 3a. To have a clear view of the road. Figure 3b shows how the smartphone camera is positioned outward facing, and Figure 3c shows an example of the pothole captured. The same smartphone was used to record the accelerometer data. There was no ambiguity or confusion about the smartphone's axes and the vehicle's axis. In most of the research studies that were reviewed, the smartphones were mounted on the windshield or the dashboard. A limited number of researchers have studied the performance when the smartphone was kept in the glove box or the driver's pocket. As demonstrated by the experimental results in these studies, the placement of the smartphone in the driver's pocket or the glovebox resulted in lower detection rates.

(a) (b) (c)

Fig. 3: a) Smartphone placement at the windshield b) The road as seen from the smartphone c) Sample of a pothole

2) The orientation of the smartphone: Images showing a pothole are two-dimensional. Therefore, pothole detection will not be impacted by the placement of the camera. However, Pothole detection using accelerometer data is sensitive to the orientation of sensors. [45] and [35] assume a fixed position for analysing data from the smartphone. In [36], the researchers applied Euler angles to align the orientations for their study

[36]. The smartphone in our study was placed securely upright on the windscreen making sure that accelerometer data for all three axes are in sync with the axes of the vehicle.

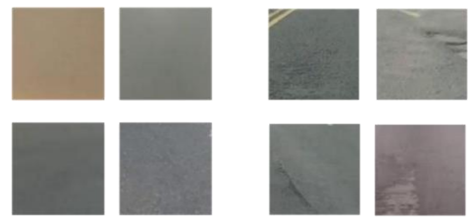
3) Speed Dependency: Inoue et al. discussed the problem of motion blur when recording the video of an object from a moving vehicle [46]. The degree of motion blur is related to camera exposure and the speed of the target. This study uses images of road potholes that are stationary. The only motion blur in the collected images could have been caused due to the speed and vibration of the vehicle. The degree of the motion blur can be reduced by keeping the exposure time low and the shutter speed fast. However, this will reduce the brightness of the videos and affect the road pothole detection rate. The paper by Douangphachanh & Oneyama discusses how the average speed of a vehicle plays an important role when measuring road roughness [47]. The speed of the vehicle influences road anomaly detection rate when using accelerometer data. The amplitude of the signal captured by a smartphone accelerometer when a vehicle passes over a pothole depends on the vehicle's speed. For this study, data was collected from all kinds of roads. However, most potholes (95%) were recorded on a B road and in town. The potholes on the motorway and A-road were not included because, at speeds of 60 miles/hour, a vehicle will cover over 26 meters per second. Practically it will be uncommon to find large potholes on these roads. Also, the current sampling rate is restricted to 100 Hz, which will not be sufficient to capture a pattern in accelerometer data at those speeds. These possible potholes will not have had a significant impact on this study. The smartphone (iPhone Xs Max, iOS) used in this study can record 4K video at 60 frames per second.

B. Data Processing

The processing of data is a prerequisite for obtaining good results. The processing for the image and accelerometer data is broken down into three stages: resampling, labelling and augmentation. This study aimed to use CNN to process raw data as they are derived from the smartphone without applying many data processing methods. The images and video frames obtained from the smartphone camera had different sizes depending upon their makeup and model. In this paper, we used 128x128 size images for training the CNNs. The following steps were taken to prepare the training dataset.

1) Resampling: Resampling was applied to the accelerometer data as it was observed that the smartphone was unable to sample the accelerometer data at the fixed frequency uniformly. The sampling rate in the iOS app's accelerometer was set at 100Hz, however it was noted that the accelerometer was sampled in the range of 70-100Hz. The 1D CNN uses fixed-length data and, hence, to have a consistent sampling rate at 100 Hz, the data with lower sampling rates would normally have to be deleted. For this study, data was resampled at 100 Hz, and missing values were filled by interpolating the data uniformly.

2) Labelling: The proposed fusion model takes imagery and accelerometer data as inputs. The accuracy of a deep learning supervised model depends upon the accuracy of data labelling. Douangphachanh & Oneyama have discussed how the accuracy of object detection depends on how correctly the data were labelled [47]. The collected video (MOV) data was converted to a suitable image format, namely jpg using a small script in Python. The images obtained were examined manually and images with unwarranted features were discarded. The images were then divided into two classes: No pothole (DM00) and pothole (DM01) based on manual inspection and annotation. Table I shows the number of images in the final dataset. Figures 4a and 4b show the images of normal road surface and potholes, respectively, captured from the dashboard camera.



(a) (b)

Fig. 4: a) Normal Road surface (no pothole) b) Damaged Road surface (pothole)

A software script was used to match location and time from the photos and videos to the sensor data's location and time for labelling the sensor data. GPS location and timestamp was used to tag the dataset and produce tuples of 100 samples (sampling rate). The accelerometer data was recorded at the sampling rate of 100 Hz on all three axes (X, Y, Z). A one-second data sample has 100x3 timestamps. Figure 5 shows the accelerometer reading on the Y axis representing a significant movement compared to the reading on X and Z axes. Figure 5 shows a 4 second data sample with pothole detection in the 3rd sec (timestamps 200-300). It can be noticed that the accelerometer reading on Y axis had a significant dip during timestamps 200-300 when the pothole was detected. The whole 1-sec window was labelled as a pothole.

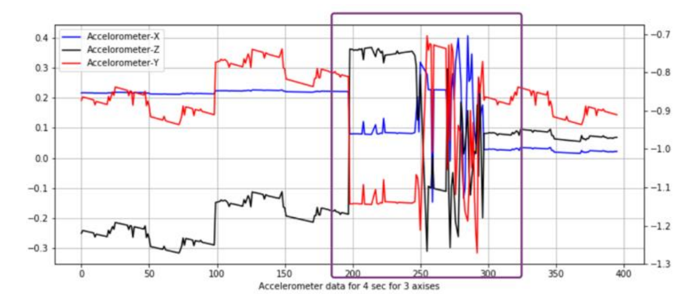
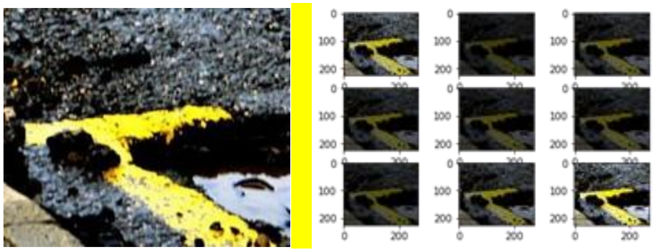


Fig. 5: Labelling of a pothole on accelerometer data

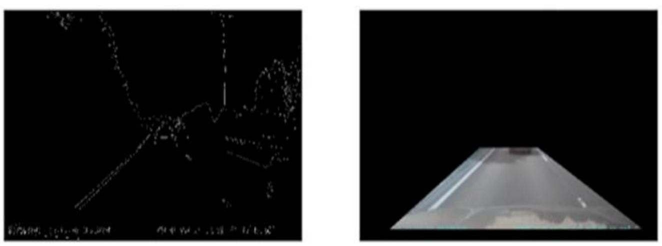
3) Data augmentation: Data augmentation methods, such as permutation and scaling, were applied to increase the accelerometer dataset size. The scaling factors +/-5% were applied to accommodate variation in the data collected from different types of vehicles. The permutation on the dataset was applied to increase the number of potholes detected samples. The data set of images was prepared carefully to include images from various weather conditions such as dry road and potholes filled with water and at different times to accommodate brightness variation. Due to the limitation of time and geographical reach, data augmentation was performed on the images to increase the number of images in the data set. The image processing library NumPy was used to flip the images horizontally and vertically. The brightness of the images in the dataset was adjusted in the range of 0.2 to 1 to replicate images with brightness variations to emulate differences in brightness conditions as shown in figure 6.



(a) (b)

Figure 6: a) A pothole image to be augmented b) Pothole images with adjusted brightness

A canny edge detector [48] was used to extract edges in an image when selecting the area of interest. Before applying this method, a gaussian blur was applied to smoothen the images. Figure 7(a) shows an image with a Canny edge detector, and Figure 7(b) shows an image with the area of interest, which is within the marked lane in which the vehicle was moving. As we want to record potholes on-road only, the black part of the image is masked in Figure 7(b) to mark the area of interest between the white lines delimiting the road clearer.



(a) (b)

Fig. 7: a) Image with canny edge detector b) area of interest within the canny edges (road lane)

C. CNN Model Training for Classification

Yamashita et al. discuss how CNN, a class of Artificial Neural Networks, is inspired by the visual cortex of animals [49]. CNNs were mainly designed to be used when performing the identification and classification of images. CNN can learn features of two-dimension images to produce efficient results. The CNN architecture comprises of three main layers: Convolution Layer, Pooling Layer and Fully Connected layer. CNNs reduce the model's learning complexity by sharing the weights during training [50]. The model's capacity and complexity can be changed by changing the number of CNN layers and their organization. This study used a fusion of two-dimensional Convolution Neural Networks (2D-CNN) and one-dimensional Convolution Neural Networks (1D-CNN). The 1D-CNN, takes accelerometer data as input and the 2D-CNN takes images as input. Further down the layer hierarchy of the Deep Neural Networks (DNN), outputs of 1D-CNN and 2D-CNN are concatenated, and dense layers are used to produce the final result. The 2D-CNN models with hidden layers of non-linear transformation ranging from three to five layers were used to obtain the classification results. 1D-CNN models with multiple non-linear transformation layers ranging from two to seven layers were used to get the classification results. It was observed that the 2D-CNN model with five hidden layers and 1D-CNN networks with two hidden layers were able to produce the most satisfactory results. The activation function used on hidden layers was ReLu (Rectified Linear Unit), and the SoftMax activation function was used on the output layer. The CNN used categorical cross-entropy loss function and Adam optimizer.

$$L(x, y; \theta) = -\frac{1}{N}\sum_{1}^{N}(-y_p \,.\, log\,(\hat{y_{i-p}}) - y_n.\,log(\hat{y_{i-n}})) \quad (1)$$

$$\theta = \theta - \eta.\nabla_\theta L(x^i, y^i, \theta) \quad (2)$$

Where θ is a weight parameter. The training aims to minimize loss (1) and get optimal value for the weight parameter θ (2). For the sample x_i, the predicted negative probability is denoted by $\hat{y_{i-n}}$, and positive probability by $\hat{y_{i-p}}$. The algorithm used to train the model is shown in algorithm 1.

Algorithm 1: *To train fusion Convolutional Neural Networks model*

Input: Labelled dataset: {X, Y}
Output: Optimal θ^*;
Initial θ , epoch = 0, learning rate α
repeat

1. Sampling labelled data batch $\{x_i, y_i\}$ from {X, Y}
2. Performing forward propagation of the network and compute $[\hat{y_{i-n}}, \hat{y_{i-p}}]$
3. Compute loss L by
4. Compute adaptive gradient by SDG (Eq:2)
5. Update parameter $\theta \leftarrow \theta - \alpha\frac{\partial L}{\partial \theta}$
6. epoch = epoch + 1

until (epoch > Epochs)

IV. Results Discussion

The data set was split into 70% for training, 15% for testing and 15% for validation. To evaluate the performance of each model we used standard measures such as: classification accuracy, precision, recall and F1 score. Initially, the dataset was used to train and test the 2D-CNN model that used image data and the 1D-CNN model that used accelerometer data in isolation with various combinations of hyperparameters. For the 2-D CNN model that used image data the training was conducted with batch size 10 and over 50 epochs. During the experiments we varied the number of layers of the CNN. For a two hidden layer CNN the testing accuracy was in the range of 56% and 66%, with a median of 60.50%. This model achieved 62% precision, 51% recall and 56% F1-Score. For a 2D-CNN model with the number of hidden layers set to five the testing accuracy was in the range of 66% and 100%, with a median of 97%. The model achieved 84.80% precision, 92.40% recall and 88.44% F1-Score.

For the 1D-CNN model that used accelerometer data we tested several models with a varying number of layers from 1 to 7. Given the lower computational resources needed for the accelerometer data, the training of each model was conducted with batch size 100 and over 500 epochs. It was found that the

1D-CNN model with just two hidden layers was able to achieve the highest performance 95% testing accuracy, 93% precision, 98% recall and F1-Score of 91%. Figure 8 summarizes the results for the best configuration for each CNN in isolation.

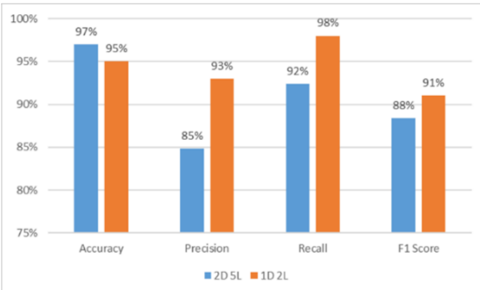


Fig. 8: Accuracies of the best 1D CNN (accelerometer data) and 2D CNN (image data) models in isolation

Based on the obtained results we used two combinations of hidden layers in the 1D and 2D CNNs respectively to test the performance of the fusion CNN model. The choice was made based on the best performance and the lower complexity for both CNNs. The first combination used was with two hidden layers in both 1D-CNN and 2D-CNN (1D-2L-2D-2L) and was chosen based on the lower complexity for both topologies; the second combination evaluated had two layers in 1D-CNN and five layers in 2D-CNN. The latter combination demonstrated the best performance during the previous phase. The accuracy and performance are shown in detail in Figure 9. The fusion model with two hidden layers in the 1D-CNN and five hidden layers in the 2D-CNN had the highest testing accuracy, 95.71% compared to a testing accuracy of 80.84% for the model topology that utilized two layers in both CNNs. This 1D 2L-2D 5L model had an average precision rate of 87.2%, average recall rate of 92.7% and F1-Score of 90%. It is important to highlight that while not the best fusion model the 1D 2L-2D 2L model had considerably better accuracy compared to the 2D-2L model that achieved 60.50% testing accuracy. While the accelerometer CNN model appears to have better performance in some of the measures it fails to provide a timely warning to the driver and only identifies the pothole after the vehicle goes over it.

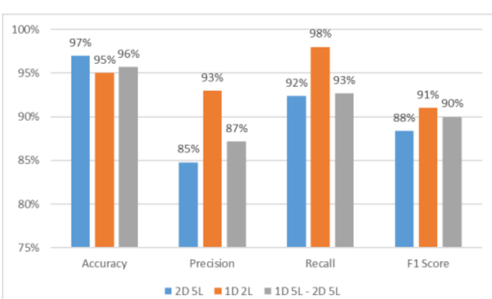


Fig. 9: Accuracies for all models.

As demonstrated by the results the proposed fusion model achieves state of the art performance in identifying potholes. Moreover, by utilizing both accelerometer and image data the fusion model can overcome the disadvantages of other modern techniques that utilize each data input separately. By using accelerometer data, the proposed fusion approach can overcome the difficulty of modern computer vision methods to differentiate between potholes and other similar visual patterns. By utilizing image data, the proposed approach goes beyond current vibration-based methods by providing a proactive solution that protects both the driver and the vehicle.

V. Smart City Pothole Detection IoT Framework

As per most DL models the accuracy of the one proposed in this paper is only expected to increase with the size of the data set and when the training data have potholes of different sizes, shapes and reflect diverse environmental conditions. In modern cities almost all vehicles on the road are equipped with a smartphone while CAV are expected to be a prominent technology in the near future [57]. These vehicles can potentially be data collection hubs for our methodology. Every vehicle on the road running the application can contribute to the creation of a more intelligent model. This is achieved through a hierarchical architecture where the end user in the vehicle constantly performs pothole identification using an edge-based DL model, while capturing new image and accelerometer data and updating the local model. The captured data used to train the local models should remain private so the individual drivers positions or routes are not stored centrally. The local model updates can then be aggregated in an intermediate step in a local embedded server existing on smart road signs creating new locally aggregated models which in turn are transmitted to a global server where all local models are aggregated to create a new smarter global model. The global model is then disseminated to the end users in a backward process in fixed time intervals. 6G supports this intensive multi bi-directional communication. The process can be seen in figure 10.

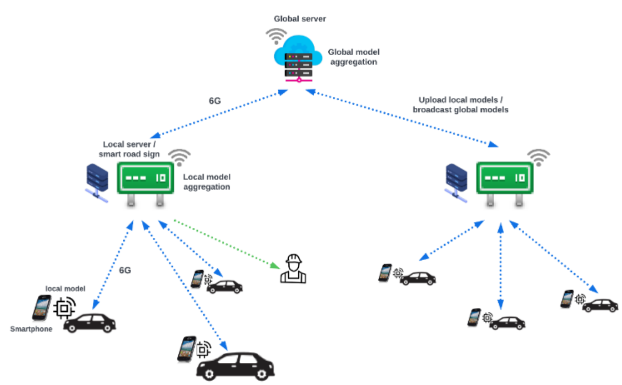


Figure 10: Hierarchical model architecture.

To exploit the potential of this framework, tackle the training challenges and develop a model that is constantly improving using immense amounts of new data captured by the road users, we suggest Federated Learning (FL) as a method for retraining and optimizing the global pothole detection algorithm. FL is a collaborative ML framework that can support the training of a global ML model without accessing edge devices' privately held raw data. The FL approach features align perfectly with the needs of our framework. As argued in [54], FL is designed for multiple users that collaborate to solve a complicated ML problem [54]. In our case, numerous smart devices within the vehicles roaming the roads, will contribute in developing an optimal classification algorithm. Additionally, by incorporating the local computing and model transmission concepts, the federated learning approach contributes to the privacy and security of the data captured by these devices. Strengthening of data privacy and security is a key feature of FL as demonstrated by recent research efforts [55]. Understandably, this is very important for our framework where sensitive location and image information are collected and used to develop the model.

Here dedicated edge servers are responsible for aggregating local learning model updates and disseminating global learning model updates [51]. The proposed FL approach involves learning an optimal DL model for pothole detection that will be continuously optimised from the data captured by the end users' smart phones. The models used by these devices will be updated based on the optimised global model to provide pothole identification. The training and updating process will involve a bi-directional 6G communication where local model updates will be communicated and aggregated to smart road sign servers, which in turn will be used to update the model existing in the global cloud-based server. At the same time model updates will be communicated periodically from this central server to the smart road sign servers which in turn will communicate the updates to the user smart phone or integrated infotainment systems (IIS). In our framework, the end user/driver is running the local edge-based DL model, through a smartphone. By using this algorithm, the end device identifies the pothole on the road and the information concerning the identified pothole is securely transmitted to the corresponding local smart road sign server. The proposed infrastructure may also include second generation traffic sign technology such as LED sign boards where the pothole warning can be displayed. In our framework the local server will wirelessly transmit the warning back to receiver/smart phone or IISs residing inside the vehicles approaching an intersection monitored by a smart road sign or notify available maintenance teams to proceed with the necessary actions as shown in figure 11.

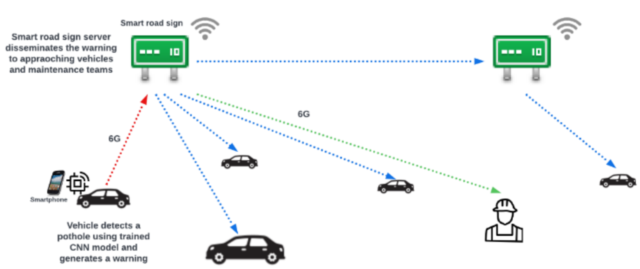


Figure 11: Pothole identification and warning process

The proposed edge training task faces considerable challenges concerning limited bandwidth and resources in the available wireless networks. Large scale deployment of the proposed framework results in a huge number of edge devices uploading the local model updates for global aggregation. In turn the updated model updates and potential generated warnings will need to be transmitted efficiently and in real time to a large number of vehicles on the road. An approach able to accommodate the needs of the proposed framework is Aircomp [52]. AirComp is a promising multiple access approach that is excellent for low-latency model aggregation [8]. As suggested by recent review studies, by concurrently transmitting the locally updated models, AirComp is able to harness interference to decrease communication bandwidth consumptions. Aircomp can support simultaneous transmission so that a dedicated access point is able to receive and estimate a summation-form function of the distributed data by exploiting the waveform-superposition property of multi-access channels [53]. This process can directly support the proposed FL approach.

VI. Conclusions and Future work

In this paper, we presented a unique hierarchical framework that exploits the latest communication, DL and edge AI training technologies providing a paradigm architecture for developing intelligent 6G enabled Intelligent Autonomous Transport applications that can use input from multiple CAV. This framework is a scalable large-scale solution where every vehicle utilizing the road can efficiently contribute to data collection and model optimization for automated pothole detection systems. Our experimental work demonstrated that the proposed CNN fusion model produced state of the art results, with 87.2% precision, 92.7% recall and an F1 score of 89.9% and can be effectively used to detect potholes of varying sizes, and shapes, under different environmental and lighting conditions. While the model's accuracy is already high, the FL architecture of the presented framework provides a promising base for continuously updating and optimizing the automated pothole detection model in a secure, efficient, and scalable way.

The framework proposed in this paper will enable the corresponding authorities to be notified in real time concerning potholes on the road network and prioritise essential maintenance work. This will tackle the challenge of carrying out necessary maintenance work in a timely and cost-effective manner and optimise the distribution and scheduling of services and resources. The companies that provide route planning services can also use the collected road-pothole data to optimise suggested traffic routes. Insurance companies may also utilize the application to support legal claims for damages caused by potholes to their customers' vehicles. Our future work will augment the database with more images /videos and sensory data. We will also use optimization algorithms to determine optimal parameters for the model and refine the proposed DL approach to improve its accuracy and explainability. Due to the usage of low-cost equipment, the sampling rates used in our experiments are not considered optimal. Future work will consider having better hardware to increase these rates.

Acknowledgements

The authors extend their appreciation to the Deanship of Scientific Research at the University of Tabuk for funding this work through Research no. 0195-1442-S.

References

[1] "Reality Check: Is our pothole problem getting worse?," BBC News, Mar. 27, 2019. Accessed: Apr. 26, 2022. [Online]. Available: https://www.bbc.co.uk/news/uk-46444109

[2] "Reported road accidents, vehicles and casualties tables for Great Britain," GOV.UK. https://www.gov.uk/government/statistical-data-sets/ras50-contributory-factors. (accessed Apr. 26, 2022).

[3] A. Haylen, "Local roads maintenance in England," 8383. [Online]. Available: https://researchbriefings.files.parliament.uk/documents/CBP-8383/CBP-8383.pdf

[4] A. Jahangiri, "Adopting machine learning methods to predict red-light running violations," in Proceedings of IEEE International Conference on Intelligent Transportation Systems, 2015, pp. 650–655.

[5] Pandey, A. K., Iqbal, R., Maniak, T., Karyotis, C., Akuma, S., & Palade, V. (2022). Convolution neural networks for pothole detection of critical road infrastructure. Computers and Electrical Engineering, 99, 107725.

[6] Pandey, A. K., Iqbal, R., Amin, S., Maniak, T., Palade, V., & Karyotis, C. (2021, November). Deep Neural Networks Based Approach for Pothole Detection. In 2021 4th International Conference on Signal Processing and Information Security (ICSPIS) (pp. 1-4). IEEE.

[7] L. Huidrom, L. K. Das, and S. K. Sud, "Method for Automated Assessment of Potholes, Cracks and Patches from Road Surface Video Clips," Procedia - Social and Behavioral Sciences, vol. 104, pp. 312–321, Dec. 2013, doi: 10.1016/j.sbspro.2013.11.124.

[8] K. B. Letaief, Y. Shi, J. Lu, and J. Lu, "Edge Artificial Intelligence for 6G: Vision, Enabling Technologies, and Applications," IEEE Journal on Selected Areas in Communications, vol. 40, no. 1, pp. 5–36, Jan. 2022, doi: 10.1109/jsac.2021.3126076.

[9] B. Emir, O. Samir, and H. Alvin, "Pothole detection with image processing and spectral clustering," in Proceedings of the 2nd International Conference on Information Technology and Computer Networks, Athens, Greece, 2013, vol. 810, pp. 48–53.

[10] K.T. Chang, J. R. Chang, and J. K. Liu, "Detection of pavement distresses using 3D laser scanning technology.", In Computing in civil engineering, 2005, pp. 1-11.

[11] T. Kim & S. K. Ryu, "Review and Analysis of Pothole Detection Methods", Journal of Emerging Trends in Computing and Information Sciences, 5(8), pp. 603-608, 2014.

[12] H. UedaK, "Detection of wet road surfaces by image processing", 40-45, 1995.

[13] Itakura, Y., Tsutsumi S., and Takehana T., "Polarization onboard sensor for detection of wet road surface, using polarization characteristics of reflection.", J IEE Jpn 66, 450-454, 1982.

[14] L. Bouilloud et al., "Road Surface Condition Forecasting in France," Journal of Applied Meteorology and Climatology, vol. 48, no. 12, pp. 2513–2527, Dec. 2009, doi: 10.575/2009jamc1900.1.

[15] L. Jing and Y. Tian, "Self-Supervised Visual Feature Learning With Deep Neural Networks: A Survey," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 43, no. 11, pp. 4037–4058, Nov. 2021, doi: 10.1109/TPAMI.2020.2992393.

[16] Q. Zou, Y. Cao, Q. Li, Q. Mao, and S. Wang, "CrackTree: Automatic crack detection from pavement images," Pattern Recognition Letters, vol. 33, no. 3, pp. 227–238, Feb. 2012, doi: 10.1016/j.patrec.2011.11.004.

[17] Oliveira H. and Correia P.L., "Crackit-an image processing toolbox for crack detection and characterization", in Proceedings of IEEE International Conference on Image Processing, 2014, 798-802.

[18] S. Mathavan, K. Kamal, and M. Rahman, "A Review of Three-Dimensional Imaging Technologies for Pavement Distress Detection and Measurements," IEEE Transactions on Intelligent Transportation Systems, vol. 16, no. 5, pp. 2353–2362, Oct. 2015, doi: 10.1109/tits.2015.2428655.

[19] R. Medina, J. Llamas, E. Zalama, J. Gómez-García-Bermejo, "Enhanced automatic detection of road surface cracks by combining 2D/3D image processing techniques", In 2014 IEEE International Conference on Image Processing (ICIP), 2014 Oct 27, pp. 778-782, IEEE.

[20] H. Yong, Z. Chun-Xia, "A Local Binary Pattern Based Methods for Pavement Crack Detection", Journal of Pattern Recognition Research, 5(1), 140-7, 2010.

[21] H. Oliveira and P. L. Correia, "Automatic Road Crack Detection and Characterization," IEEE Transactions on Intelligent Transportation Systems, vol. 14, no. 1, pp. 155–168, Mar. 2013, doi: 10.1109/tits.2012.2208630.

[22] C. Farabet, C. Couprie, L. Najman, and Y. LeCun, "Learning Hierarchical Features for Scene Labeling," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, no. 8, pp. 1915–1929, Aug. 2013, doi: 10.1109/tpami.2012.231.

[23] S. Zhan, QQ. Tao, XH. Li, "Face detection using representation learning", Neurocomputing,187,19-26, 2016 Apr 26.

[24] B. Hariharan, P. Arbeláez, R. Girshick, J. Malik, "Simultaneous detection and segmentation", In European conference on computer vision, 2014 Sep 6, pp. 297-312, Springer, Cham.

[25] A. Krizhevsky, I. Sutskever, GE. Hinton GE, "Imagenet classification with deep convolutional neural networks", Advances in Neural Information Processing Systems, 1097–1105, 2012.

[26] M. Wang, W. Li, D. Liu, B. Ni, J. Shen, S. Yan, "Facilitating Image Search with a Scalable and Compact Semantic Mapping", IEEE Transactions on Cybernetics, 2015.

[27] D. Steinkraus, I. Buck and P. Simard, "Using GPUs for machine learning algorithms", IEEE, 2006.

[28] FF. Ahmadia, MV Zoeja, H. Ebadia, M. Mokhtarzadea, "The application of neural networks, image processing and cad-based environments facilities in automatic road extraction and vectorization from high resolution satellite images", The international archives of the photogrammetry, remote sensing and spatial information sciences, 37, 585-92, 2008.

[29] S. Ren, K. He, R. Girshick, J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal net-works", In NIPS,2015.

[30] K. He, X. Zhang, "Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition", arXiv:1406.4729, 2014.

[31] TY. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, S. Belongie, "Feature pyramid networks for object detection". In Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 2117-2125.

[32] R. Girshick, "Fast r-cnn", In Proceedings of the IEEE international conference on computer vision. 2015, pp. 1440-1448.

[33] C. Luo, B. Ni, S. Yan, and M. Wang, "Image Classification by Selective Regularized Subspace Learning," IEEE Transactions on Multimedia, vol. 18, no. 1, pp. 40–50, Jan. 2016, doi: 10.1109/tmm.2015.2495248.

[34] M. Wang, X. Liu, X. Wu, "Visual classification by l1-hypergraph modeling". IEEE Transactions on Knowledge and Data Engineering, 27(9), 2564-74, Mar 20 2015.

[35] Mednis, A., Strazdins, G., Zviedris, R., Kanonirs, G., & Selavo, L. (2011). Real time pothole detection using Android smartphones with accelerometers. 2011 International Conference on Distributed Computing in Sensor Systems and Workshops (DCOSS). https://doi.org/10.1109/DCOSS.2011.5982206

[36] C. Wu et al., "An Automated Machine-Learning Approach for Road Pothole Detection Using Smartphone Sensor Data," Sensors, vol. 20, no. 19, p. 5564, Sep. 2020, doi: 10.3390/s20195564.

[37] D. Anguita, A. Ghio, L. Oneto, X. Parra Perez, JL. Reyes Ortiz, "A public domain dataset for human activity recognition using smartphones", In Proceedings of the 21th international European symposium on artificial neural networks, computational intelligence and machine learning 2013, pp. 437-442.

[38] D. Anguita, A. Ghio, L. Oneto, X. Parra Perez, JL. Reyes Ortiz, "Human activity recognition on smartphones using a multiclass hardware-friendly support vector machine", In International workshop on ambient assisted living, 2012 Dec 3, Springer, Berlin, Heidelberg, pp. 216-223.

[39] H. Wu et al., "One-Dimensional CNN-Based Intelligent Recognition of Vibrations in Pipeline Monitoring With DAS," Journal of Lightwave Technology, vol. 37, no. 17, pp. 4359–4366, Sep. 2019, doi: 10.1109/jlt.2019.2923839.

[40] SM. Lee, SM. Yoon, H. Cho H, "Human activity recognition from accelerometer data using Convolutional Neural Network", 2017 IEEE International Conference on Big Data and Smart Computing (BigComp), 131-134.

[41] P. Mohan, V. Ramjee and R. Nericell, "Rich monitoring of road and traffic conditions using mobile smartphones", In Proceedings of the 6th ACM conference on Embedded network sensor systems, Raleigh, NC, USA, 2008, pp. 357-358.

[42] S. Sattar, S. Li, and M. Chapman, "Road Surface Monitoring Using Smartphone Sensors: A Review," Sensors, vol. 18, no. 11, p. 3845, Nov. 2018, doi: 10.3390/s18113845.

[43] G. Singh, D. Bansal, S. Sofat, and N. Aggarwal, "Smart patrolling: An efficient road surface monitoring using smartphone sensors and crowdsourcing," Pervasive and Mobile Computing, vol. 40, pp. 71–88, Sep. 2017, doi: 10.1016/j.pmcj.2017.06.002.

[44] J. Eriksson, L. Girod, B. Hull, R. Newton, S. Madden and H. Balakrishnan, "The pothole patrol: Using a mobile sensor network for road surface monitoring", In Proceedings of the 6th international conference on Mobile Systems, Applications, and Services, Breckenridge, 2008, 29-39.

[45] K. Yagi, "Extensional smartphone probe for road bump detection", In Proceedings of the 17th ITS World Congress (ITS Japan and ITS America ERTICO), 2010.

[46] M. Inoue, M. Jiang, Y. Matsumoto, T. Takaki, I. Ishii, "Motion-blur-free video shooting system based on frame-by-frame intermittent tracking", ROBOMECH Journal, 4(1),1-8, 2017 Dec.

[47] V. Douangphachanh and H. Oneyama, "Estimation of road roughness condition from smartphones under realistic settings", In Proceedings of the 201313th International Conference on ITS Telecommunications (ITST), Tampere, Finland, 2013.

[48] L. Ding and A. Goshtasby, "On the Canny edge detector", Pattern recognition, 34(3), 721-725, 2001.

[49] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, "Convolutional neural networks: an overview and application in radiology," Insights into Imaging, vol. 9, no. 4, pp. 611–629, Jun. 2018, doi: 10.1007/s13244-018-0639-9.

[50] AA. Al-Saffar, H. Tao, MA. Talab, "Review of deep convolution neural network in image classification", In 2017 International conference on radar, antenna, microwave, electronics, and telecommunications (ICRAMET), 2017 Oct 23, IEEE, pp. 26-31.

[51] T. Li, A.K. Sahu, A. Talwalkar and V. Smith, "Federated learning: Challenges, methods, and future directions", IEEE Signal Processing Magazine, 37(3), 50-60, 2020.

[52] C. Xu, S. Liu, Y. Huang, C. Huang and Z. Zhang, "Over-the-air learning rate optimization for federated learning", In 2021 IEEE International Conference on Communications Workshops (ICC Workshops), IEEE, June 2021, pp. 1-7.

[53] D. Wen, G. Zhu and K. Huang, "Reduced-dimension design of MIMO AirComp for data aggregation in clustered IoT networks", In 2019 IEEE Global Communications Conference (GLOBECOM), IEEE, 2019 December, pp. 1-6.

[54] Zhang, C., Xie, Y., Bai, H., Yu, B., Li, W., & Gao, Y. (2021). A survey on federated learning. Knowledge-Based Systems, 216, 106775.

[55] Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019). Federated machine learning: Concept and applications. ACM Transactions on Intelligent Systems and Technology (TIST), 10(2), 1-19.

[56] Iqbal, R., Maniak, T., & Karyotis, C. (2019). Intelligent remote monitoring of parking spaces using licensed and unlicensed wireless technologies. IEEE Network, 33(4), 23-29.

[57] Ahmed, M. L., Iqbal, R., Karyotis, C., Palade, V., & Amin, S. A. (2021). Predicting the public adoption of connected and autonomous vehicles. IEEE Transactions on Intelligent Transportation Systems, 23(2), 1680-1688.

[58] Iqbal, R., Doctor, F., More, B., Mahmud, S., Yousuf, U (2020): "Big Data analytics and Computational Intelligence for Cyber–Physical Systems: Recent trends and state of the art applications", Future Generation Computer Systems, Elsevier, Volume 105, pp. 766-778.

[59] Iqbal, R., Maniak, T., Doctor, F., Karyotis, C., (2019): "Fault detection and isolation in industrial processes using deep learning approaches", IEEE Transaction on Industrial Informatics, Volume 15, Issue 5, pp 3077-2084.

[60] Karyotis, C., Maniak, T., Doctor, F., Iqbal, R., Palade, V., & Tang, R. (2019, December). Deep learning for flood forecasting and monitoring in urban environments. In 2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA) (pp. 1392-1397). IEEE.

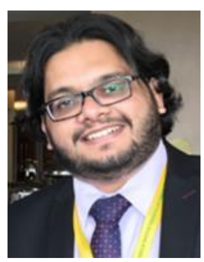
Mohammad Hijji (IEEE member) received his PhD degree in Computing from Coventry University, UK, in 2017. He was the Chairman of Computer Science Department, University of Tabuk, Saudi Arabia, from 2020 to 2022. He is currently the Vice Dean for Development and Quality, FCIT. His research interests include AI, Cyber Security, IoT, Smart City, Disaster and Emergency Management.

Rahat Iqbal (IEEE member) is a senior academic and inventor with 20 years' experience in industrial and academic roles. His research interests lie in Big Data Analytics, IoT, CAV, and AI. Prior to joining the University of Dubai, he was leading the Connected Mobility Research Group at Coventry University, UK.

Anup Kumar Pandey received the Ph.D. in DL and Computer vision from Coventry University in 2021. He brings over 15 years of Industrial experience. He has received three master degrees including MBA from Oxford university. He had worked on large scale IT project in USA and UK. He has authored few papers for journals and conference proceedings.

Faiyaz Doctor (IEEE senior member) is a Senior Lecturer at the Centre for Computational Intelligence, School of Computer Science and Electronic Engineering, University of Essex, with over 20 years' experience in the design and application of biological inspired techniques and intelligent agents approaches for real world systems.

Charalampos Karyotis is a senior data scientist at Interactive Coventry Ltd. He has a Ph.D. degree in computer science from Coventry University and over 8 years of experience in data-driven technologies and industrial projects. His research interests include ML, smart city systems and big data analytics



Wahid Rajeh received the PhD degree in Computer Applied Technology from Huazhong University of Science and Technology. He is an associate professor at College of Computer and IT at University of Tabuk and the head of the innovation lab. His research interest includes Cloud Computing, Big Data, Security Analysis, Grid Computing.

Ali Alshehri received his Ph.D. degree from Oakland University, Rochester, MI, USA, in 2019. He is currently an Assistant Professor of Computer Science at University of Tabuk, Saudi Arabia. His research interests include Information Security, Web Services, Privacy, Intrusion Detection and Prevention, Cryptography, IoT, Blockchain, and Digital Forensics.

Fahad Aradah Fahad Aradah received his Bachelor degree from the University of Tabuk. His research interests include Artificial Intelligence, IoT, Machine Learning and Deep Learning.