Distributed acoustic sensor systems for vehicle detection and classification

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 10 March 2023, accepted 19 March 2023, date of publication 22 March 2023, date of current version 31 March 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3260780

RESEARCH ARTICLE

Distributed Acoustic Sensor Systems for Vehicle Detection and Classification

CHIA-YEN CHIANG[1], MONA JABER[1], (Senior Member, IEEE), KOK KEONG CHAI[1], AND JONATHAN LOO[2], (Member, IEEE)

[1]School of Electronic Engineering and Computer Science, Queen Mary University of London, E1 4NS London, U.K.
[2]School of Computing and Engineering, University of West London, W5 5RF London, U.K.

Corresponding author: Chia-Yen Chiang (c.chiang@qmul.ac.uk)

This work was supported by the Joint Program between Queen Mary University of London (QMUL) and Beijing University of Posts and Telecommunications (BUPT).

ABSTRACT Intelligent transport systems (ITS) are pivotal in the development of sustainable and green urban living. ITS is data-driven and enabled by the profusion of sensors ranging from pneumatic tubes to smart cameras which are used to detect and categorise passing vehicles. Simple sensors, such as pneumatic tubes, are successfully deployed for counting passing vehicles but are not useful for vehicle tracking or re-identification. Smart cameras, on the other hand, collect comprehensive information but suffer from occlusion, patchy coverage, and compromised vision in adverse weather and visibility. This work explores a novel ITS data source based on an optical fibre which acts as an uninterrupted length of virtual sensors using a distributed acoustic sensor (DAS) system. Based on real DAS data collected in the field, we first present a study of latent DAS features that uniquely identify a given vehicle, otherwise referred to as the vehicle signature. We formulate a classification problem that examines incoming DAS data to extract vehicle signatures and identify the different types of vehicle. To this end, we implement different classification methods and present a comparative performance analysis that reveals novel insights into the potential role of DAS for ITS applications. This work is a pilot study of DAS for vehicle classification that is driven by real DAS data and validated by promising results where a vehicle's type is correctly identified with 94% accuracy and the size of a vehicle with 95% accuracy.

INDEX TERMS Intelligent transport system (ITS), distributed acoustic sensors (DAS), classification, vehicle type.

I. INTRODUCTION

Intelligent transport systems (ITS) are concerned with applying technology to enable informed traffic management and environment-friendly traffic flow. Vehicle detection is an essential ITS objective with a rising interest in vehicle classification. The vehicles' characteristics of interest to ITS include the size, the occupancy, and the engine type, as each would have a different impact on traffic planning and environmental pollution. In recent years, numerous types of traffic detector have been deployed to monitor traffic flow, such as remote traffic microwave sensors [1] and magnetic sensors [2], in addition to more traditional sensors such as pneumatic road tubes [3]. However, the data generated by these sensors is mostly limited to counting of vehicles to apply in tasks such as optimising traffic flow by controlling traffic lights. More recently, video sensors have been deployed for traffic flow monitoring with the ability to classify the type of passing vehicles and to track their movements. For instance, the authors of [4] employ advanced deep learning techniques in the analysis of video footage to detect and track vehicle movement. It is, however, extremely challenging and costly to cover and analyse kilometers of road networks with video cameras and there are inevitable limitations due to blind areas and poor performance in adverse light and weather conditions [5]. Furthermore, privacy concerns due to omnipresent cameras in urban spaces and related regulations, such as the General Data Protection Regulation (GDPR),

The associate editor coordinating the review of this manuscript and approving it for publication was Maurice J. Khabbaz.

have resulted in public reluctance about the spread of video sensors [6]. The authors in [7] propose a deep neural network model to mine microphone data and classify different transportation modes, including On-Road and four related sub-categories: 'Auto Rickshaw', 'Bus', 'Car', 'Pedestrian'. Despite the encouraging results obtained, such a method does not support continuous and uninterrupted sensing and is not suitable for the problem defined in this work: detecting different types and/or sizes of passing vehicles. A network of wireless acoustic sensors in proposed in [8] to classify two types of military vehicles. A classification system is designed that can mitigate the impact of sensor faults and environmental noise. In this case, 23 sensors are used to cover 0.27 square kilometers, hence the proposed classifier is not scalable for ITS solutions.

In this work, we propose to use fibre-based distributed acoustic sensors (DAS) to collect uninterrupted data along the length of the road network. Indeed, fibre optic communication is available across the globe and is based on optical fibres deployed underground (or underwater). Unused optical fibre cables, sometimes referred to as dark fibre, are often purposely laid along used fibre for future use; these can be re-used to enable a DAS system. By connecting a DAS interrogator to a fibre, each point along the fibre becomes a sensing unit and, therefore, yields a system that achieves continuous detection along the length of the road. The DAS system continuously monitors fluctuations of the reflected probe signal that are caused by external vibrations; these are referred to as acoustic events. Current advances in DAS technology allow the localisation of acoustic events along the fibre and the categorisation of the event (e.g. a passing car or drilling). Commercial DAS systems are mostly based on optical backscattering technology and are widely used in monitoring of oil and gas pipelines [9], peripheral safety [10], structural health [11], and submarine power cables [12]. Recently, DAS systems are considered for traffic flow detection with promising results for vehicle detection and speed estimation [13]. In this context, DAS is a relatively new technology that promises to overcome the issues with blind areas and reduce the deployment cost of equipment and data analysis. More importantly, DAS is not affected by adverse weather or luminosity conditions and does not capture personal data such as faces and clothing.

A. RELATED RESEARCH

Advanced feature extraction techniques are currently applied to DAS data to get more information such as the signature of the detected event and its identification. In line with acoustic signals classification, manual feature extraction techniques were first adopted in DAS event classification. Such techniques include wavelet packet transform [14], [15], spectral substitution [16], Mel-spectrograms [17], and empirical mode decomposition [18]. The classification task is then conducted using conventional classifiers such as support vector machines (SVM such as [14]) and relevant vector machines [15]. In order to improve the classification results, some works propose to use advanced classifiers instead of conventional classifiers. For example, the authors in [18] use XGboost, an ensemble algorithm, and the authors in [16] and [17] use convolution neural networks (CNN), achieving a higher success rate than conventional classifiers. Instead of manual feature extraction, the authors in [19] apply a one-dimensional CNN (1D-CNN) directly to the raw signal followed by an SVM-based classifier, and outperform previous works.

Most of these works aim to detect and classify DAS signals generated by distinguishable events such as road works [16], [17], [19]. Others are concerned with distinguishable man-generated disturbances [15], [18]. However, none of these works attempts to distinguish similar events (e.g. type of moving vehicle) based on DAS-salient features. In other words, the listed works are not designed to classify the type or size of a passing vehicle using DAS. Recently, the authors of [14] presented an SVM-based classification method to identify types of vehicles, namely: cars, sport utility vehicles (SUV), and trucks with 71.69% accuracy. Another paper, [20], uses DAS signals to identify heavy vehicles by associating the high number of axles detected visually with a DAS disturbance threshold. It follows that this method does not distinguish vehicle types nor the different size/weight of two vehicles with the same number of axles.

B. CONTRIBUTIONS

In this work, we posit that DAS data contains representative features that uniquely characterise a given vehicle type and size. This is an important ITS problem that applies to different use cases. One such case represents controlled access areas, such as airports, ports, or manufacturing sites, where only authorised and known vehicles are allowed to roam. In this case, the problem we address can serve as a security measure to detect the movement of an unauthorised vehicle. Another use case relates to urban areas with restricted access to large vehicles. In this context, detecting the size of any vehicle approaching the restricted areas can be informed by the detected DAS signal. The code and data of this work are available here: https://github.com/Chiayen0503/Distributed-Acoustic-Sensor-Systems-for-Vehicle-Detection-and-Classification.git

In our investigation, we first conduct a DAS feature exploration study that aims to identify unique signatures that represent specific vehicles based on real collected data. For this purpose, we employ empirical mode decomposition (EMD) energy analysis [18] to extract a feature vector from the original normalized DAS data and a differential signal. Our study suggests that EMD is not a suitable method for feature extraction since it assumes that the vehicle signature is embedded in the energy carried by each time-domain Intrinsic Mode Functions. Another common approach for processing acoustic signals relies on Mel filter banks which mimic the human ear's acoustic perception, with better

discrimination for lower frequencies than higher frequencies. Recently, Google published a new model that replaces the traditional fixed features of Mel banks by a learnable frontend for audio classification (LEAF) [21]. Despite the learnable side of LEAF, it nonetheless focuses the feature extraction on frequency components of the acoustic signal; it follows that it behaves poorly in DAS signal classification.

Based on these findings, we adopt an automatic feature extraction approach using supervised learning and find that DAS data indeed contains unique vehicle signatures. This is confirmed by using raw DAS data as input signal to different classifiers commonly used for identifying acoustic signals, such as VGG-16 [22] which achieves 67% accuracy. Next, we propose the first 1D-CNN approach for the extraction of distinguishable latent features of similar DAS events created by moving cars. In particular, our method successfully classifies five categories of vehicles moving at different speeds. We argue that automatic feature extraction using a 1D-CNN is the most successful approach to capturing the vehicle class signature when followed by a softmax classifier. We present a performance analysis of the proposed method and we investigate the role of car size on the imprinted DAS signature. Our work outperforms prior art in [14] and reaches an average accuracy of 94% for car type and 95% for car size.

This is a pioneering feasibility study of using DAS in vehicle type/size classification based on a dataset acquired under controlled conditions. In the hope of motivating more research in this area, the dataset and code are made available to the public. Moreover, we posit that the methods proposed in this manuscript are not specific to this dataset but that the models obtained may benefit from re-training when presented with DAS data that was acquired in different conditions. To this end, We discuss the applications and limitations of the proposed methods for different conditions including traffic scenarios, road types, fibre characteristics, and weather conditions. Two use cases are examined: (a) Areas with vehicle-controlled access and (b) Urban areas with restricted access to large vehicles. In the first case, the list of permitted vehicles, road, and fibre conditions are known and remain mostly unchanged. In this case, the model would benefit from retraining based on collected data specific to the target area.

The second case is defined based on common concerns in transportation planning[1] where the dimensions and weight of vehicles on the roads are required in ITS planning. This work does not suggest to provide a complete answer to these questions; instead, we offer an alternative means to detect large vehicles when modelling the traffic flow. In urban areas, vehicles and traffic conditions are not predefined, but the covered road and the conditions of the fibre attached to each DAS interrogator and fixed. In this case, we posit that the model would benefit from retraining to better represent these characteristics, and propose to address the challenge of labelling by employing alternative data sources (e.g., camera-based) in a fixed location of the road. The advantage of the DAS-method in this case is that it offers a reliable and efficient detection of large vehicles for tens of kilometers with a single interrogator. The paper is organised as follows. The DAS technology, DAS system, and DAS dataset are first described in Section II. This is followed by a problem formulation in Section III and DAS data exploration in Section IV. Next, we present the proposed 1D-CNN approach in Section V. In Section VI, we present the obtained results and we discuss the findings in Section VII. We finally conclude the article in Section VIII.

[1] https://www.hse.gov.uk/workplacetransport/vehicles.htm

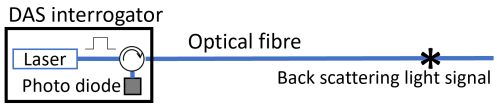


FIGURE 1. DAS system.

II. DISTRIBUTED ACOUSTIC SENSING: SYSTEM AND DATA

In this section we first present the general theory of DAS followed by the data acquisition method adopted in the collection of the dataset used for this research.

A. DAS SYSTEM

A DAS system is an opto-electronic device sensitive to the strain distributed over an optical fibre of the length of up to 40 to 50 km [23]. The technology is rooted in Optical Time Domain Reflectometry (OTDR) where a pulse of coherent light is periodically injected into a fibre and a fraction of the light reflected back via Rayleigh (elastic) scattering mechanism is captured by a photodetector at the launching end (see Figure 1. Each probing pulse results in a continuous time series of back-scatter intensity, commonly referred to as a *fibre shot*, with the time being proportional to the distance that the pulse has traveled along the fibre. In conventional OTDR the back-scatter intensity is a smooth exponentially decreasing function of fibre distance unless there is a sudden variation in the local fibre reflectivity coefficient due to a faulty splice or a fibre break. In a DAS system, in contrast to OTDR, the intensity is a random function of fibre position and the cumulative phase of the interference of light scattered back by the fibre within the interrogating pulse. This fibre interval giving rise to the back scatter interference is called a *resolution cell*. Although the back-scatter phase from each resolution cell is inherently random due to the random molecular structure of the fibre glass, it stays constant as long as the state of fibre within the corresponding resolution cell remains unchanged. If, however, a fibre is subjected to dynamic strain, the strain variation results in a variation of the back-scatter phase and hence back-scatter intensity. Therefore a series of back-scatter measurements at a given fibre distance carries information about the evolution of the strain applied at the corresponding fibre position.

The size of the resolution cell defines the spatial resolution of a DAS system and is determined by the width of the

probing pulse: a 100 ns pulse corresponds to the spatial resolution of approximately 10 m. The spatial resolution is different from the spatial sampling period. The latter is related to the rate of the analog-to-digital converter (ADC) that samples the fibre shots for the purpose of data logging and digital signal processing. Usually, the ADC rate is configured so that the sampling period is substantially smaller than the spatial resolution. The temporal sampling resolution is given by the interrogator's *pulse repetition frequency* (PRF) which is limited by the length of the sensing fibre. Maximum PRF varies from 2.5 kHz to 100 kHz for 40 km and 1 km long fibres, respectively. The output of a DAS system is thus a collection of digitized fibre shots acquired at a given PRF. Given that a fibre shot constitutes a set of measurements sampled along the fibre length, DAS data can be viewed as a 2D spatio-temporal array of scalar measurements.

Implementations of a DAS system can be grouped into two categories with respect to their interrogation method: intensity-based and phase-sensitive systems. In the former, it is the back-scatter power (or intensity) that is measured while in the latter the temporal changes in back-scatter phase constitute the acquired signal. Intensity-based systems are simpler in design. However, their *transfer function*, i.e., the strain-intensity relationship, is a non-linear and not predictable function of strain. Phase-sensitive DAS systems, being more complex in design, feature linear transfer (strain-phase) function which makes the acquired data more amenable to interpretation. For this research, we used a phase-sensitive Helios DAS [24]. The system utilizes a proprietary phase interrogation technique to recover temporal evolution of the back-scatter phase along the fibre. The system was operated with 100 ns pulse width corresponding to spatial resolution of 10 m. The temporal phase sampling frequency was 6 kHz. The ADC rate was set to 150×10^6 samples per second resulting in the spatial sampling period of $\approx$0.68 m.

B. DAS DATA

We set up a controlled experiment in which DAS data is collected along a 4.8 km stretch of an inter-city road leading into a town. The selection of this particular section of the road aimed to capture traffic at different speeds: faster outside the town and slower as the vehicles approach the town centre. The fibre used was purposely laid along the road stretch at 20 cm underground depth in a micro-trench. Fotech's Helios DAS interrogator (see Section II-A) was used to capture the DAS signals.

In this research, the DAS signals obtained from controlled movement of five different vehicles are studied. There are five conducted experiments in which the five vehicles moved at a given fixed speed (30 km/h, 40 km/h, 50 km/h, 60 km/h, 70 km/h) and with predefined order and inter-vehicle distance in both road directions. One side of the road was blocked and the experiments were carried at night to reduce the chance of external vehicles driving along the controlled road segment. The five vehicles are listed in Table 1.

Figure 2 shows a section of the collected data for the 50 km/h speed experiment. In this figure, the x-axis indicates the time in fibre *shots*, s, where one shot is equivalent to $1/1000.04$ seconds and the y-axis refers to the position along the laid fibre in fibre *bins*, b, where one bin is equivalent to 0.68 metres. The lightness of each pixel at bin b and shot s represents how strong was the signal displacement $\rho(b, s)$, as a result of the vibration generated by the moving vehicle. The dominant five lines, as highlighted in Figure 2, represent the track of the five vehicles moving at the same speed and respecting the same inter-vehicle distance. Two other lines can be seen in Figure 2. The first represents an out-of-experiment vehicle which was moving in the same direction and was positioned between Car 2 and Car 3. The other one represents another external vehicle which was moving in the opposite direction and passed both Car 4 and Car 5.

Each of the five experiments (for the different five car speeds) generated a two dimensional dataset where the first dimension is the fibre bin such as $1 \leq b \leq B$ and B is the last fibre bin, and the second dimension is the fibre shot $1 \leq s \leq M$, where M is the last acquired shot in the dataset. The reported value is that of the signal displacement $\rho(b, s)$ in radians for each position determined by s and b.

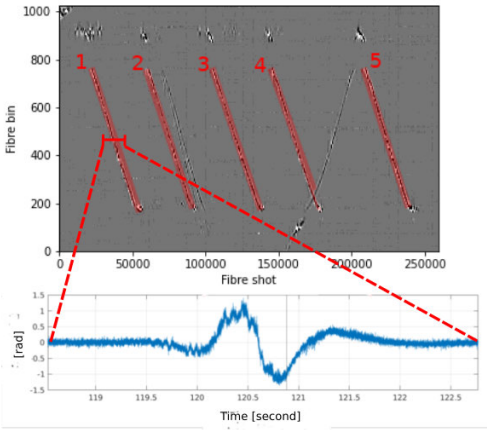


FIGURE 2. Section of DAS collected data (top) for the 50 km/h experiment showing the five target vehicles in addition to two unwanted ones. In addition, a sample DAS signal is taken at bin 450 for Car 1 as shown in the bottom, where 1000.04 fibre shots are presented as one second and the y-axis shows the disturbance level in radians.

III. PROBLEM FORMULATION

Consider a raw DAS dataset of fibre length B (bins) collected over a duration of M (shots) fibre shots. Each fibre bin $1 \leq b \leq B$ represents a virtual acoustic sensor that measures the displacement of the optical signal in radians at any time $1 \leq s \leq M$. Let $\varrho \in \mathbb{R}^{B \times M}$ be this complete raw dataset over the entire length of the fibre B and over the full period of recording M. Let $\mathbf{x}_b^\omega \in \mathbb{R}^{d_b^\omega}$ be a data sample of index $1 \leq b \leq B$ that is collected from bin/sensor b during the duration $d_b^\omega < M$. The superscript $\omega \in \{0, 1, 2, \ldots T\}$ is the index of an event where $\omega = 0$ refers to an unwanted event (noise) and $\omega > 0$ indicates the type of the detected car

TABLE 1. The list of vehicles included in the trial with their respective kerb weights and lengths and categorisation as small or large.

Car number	Make and Model	Length (m)	kerb weight (kg)	Small/Large
Car 1	Jeep Compass - SUV category	4.4	1935	Large
Car 2	Renault Clio - Subcompact cars category	4	1205	Small
Car 3	Toyota CHR (Hybrid) - Compact cars category	4.3	1450	Small
Car 4	Fiat Doblo - Multi-purpose vehicle	4.4	1320	Small
Car 5	Renault Trafic - Light commercial vehicle	5	1810	Large

(e.g., $\omega = 1$ indicates Car 1, $\omega = 2$ indicates Car 2, etc). Based on the controlled experiment setup described in Section II-B, there is a unique $d_b^\omega > 0$ for each value of $1 \leq \omega \leq T$ at each bin b, given by $d_b^\omega = s_{b,e}^\omega - s_{b,s}^\omega$, where for all bins $1 \leq b \leq B$, $s_{b,s}^\omega$ is the starting shot of event ω at bin b and $s_{b,e}^\omega$ is the end shot. Let $d^\omega = \max(d_b^\omega)$ be the largest duration for an event $1 \leq \omega \leq T$ for all values of $1 \leq b \leq B$. Following the order of the cars, it can be said that for any bin b, $s_{b,s}^1 < s_{b,e}^1 < s_{b,s}^2 < s_{b,e}^2 < \cdots < s_{b,s}^T < s_{b,e}^T$ and all other events represent noise (i.e., $\omega = 0$).

We can then express the set of all data samples $\mathbf{x}_b^\omega$ in the dataset of type ω as $X^\omega \in \mathbb{R}^{B \times d^\omega}$ as the sequence of B data samples describing event ω that are collected by B sensors for a duration of d^ω seconds. The starting shot and end shot of each data sample in X^ω differ, and they each span a different duration; in this work d^ω is the largest duration of an event ω within a sequence X^ω. Similarly, we can express $\mathbf{X} \in \mathbb{R}^{B \times d^\omega \times T}$ as the union of all such sequences X^ω in the data set for all $1 \leq \omega \leq T$ where each data sample $\mathbf{x}_b$ in X^ω (we drop the superscript ω since all samples in X^ω have the same ω) is associated with a label $y_b = \omega$ and $b = 1 \ldots B$.

In this paper, we examine five different datasets $\{\mathbf{X}_{30} \ldots \mathbf{X}_\nu \ldots \mathbf{X}_{70}\}$ where $\nu \in \{30, 40, 50, 60, 70\}$ is the speed of the moving cars in each experiment and $\mathbf{X}_{JS}$ is the joint dataset that combines joint speeds. For simplicity of mathematical notation and unless more than one dataset is used at the same time, we use the notation $\mathbf{X}$ to indicate a given dataset without the subscript ν. The problem can thus be formulated as a classification predictive modelling; a process of predicting car types that are categorised into T classes/labels (e.g., Car 1, Car 2, etc. or Large/Small) by approximating a mapping function from input data samples in $\mathbf{X} = \{X^1, \cdots, X^T\}$ into discrete output labels Y. To this end, we reorganise the set of input data samples $\mathbf{X}$ into a training set $\mathcal{D}^t$ and a testing set $\mathcal{D}^v$. $\mathcal{D}^t \in \mathbb{R}^{D^t \times d^m}$ is a two dimensional matrix where D^t is a design parameter that indicates the number of samples included in the training set and $d^m = \max(d^\omega)$ for all values of ω in ρ. Each row in $\mathcal{D}^t$ is a data sample $\mathbf{x}_i$ from any X^ω, taken at bin i, with the corresponding label y_i. In case $d^\omega < d^m$, zero padding is applied to edge shots, i.e. $\rho(i, j)$ is set to 0 for all shots j where $j < s_{b,s}^\omega - (d^m - -d^\omega)/2$ and $j > s_{b,e}^\omega + (d^m - -d^\omega)/2$.

The training dataset $\mathcal{D}^t$ is shown in Eq(1), at the bottom of the page, where $\rho(i, j)$ is a single measurement value (in radians) for a particular sensor (bin i) at a certain time (shot j) within the data sample $\mathbf{x}_i$. In this case, the bin i represents the index of the data sample and j is shot index relative to the start shot of $\mathbf{x}_i$. For example, for a data sample $\mathbf{x}_i$ taken at bin i with label $y_i = \omega$, a value $a_{i,j}$ is equal to $\rho(i, s_{b,s}^\omega + j)$ and $a_{i,d^\omega} = \rho(i, s_{b,s}^\omega + d^\omega)$.

Similarly, $\mathcal{D}^v \in \mathbb{R}^{D^v \times d^m}$ is a two dimensional matrix where D^v is a design parameter that determines the number of samples used in the testing phase. Each sample $\mathbf{x}_i$ in $\mathcal{D}^v$ is associated with a label y_i as in Eq(1). The sum of $D^t + D^v \leq T \times B$, since each vehicle within the controlled experiment passes by each bin within the controlled stretch of the fibre and passes only one time through each bin during the recording of the full dataset ϱ.

IV. DAS FEATURE EXPLORATION

Similar to standard acoustic signal exploration, in this section, we examine DAS signals by studying the time domain and frequency domain representation. To this end, we extract the phase displacement $\rho(b, s)$ for one specific bin and speed over a window of 4,500 shots that spans across the visual marks left by Car 1 and Car 2, as shown in Figure 3 (Top) and (Middle), respectively. These two signals are fed into a cross-correlation algorithm which applies a lag ranging between -4500 and 4500 to one signal and, for each lag value, computes the dot product to determine the correlation. The result is show in Figure 3 (Bottom) and indicates high cross-correlation (close to 1) between these two different car types. Such a high correlation between the DAS signal temporal representation of these two significant different vehicles (Car 1:Large and Car 2:Small), indicate that temporal-based features are not sufficient to distinguish their respective signatures.

$$\mathcal{D}^t = \underbrace{\begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_{D^t} \end{bmatrix}}_{D^t/T \text{ random data samples from each } X^\omega} = \underbrace{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1d^m} \\ a_{21} & a_{22} & \cdots & a_{2d^m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{D^t1} & a_{D^t2} & \cdots & a_{D^td^m} \end{bmatrix}}_{\text{Each data sample has a duration of } d^m \text{ shots}} \quad Y = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{D^t} \end{bmatrix}}_{\text{labels } y_i = \{1, \ldots, T\}} \tag{1}$$

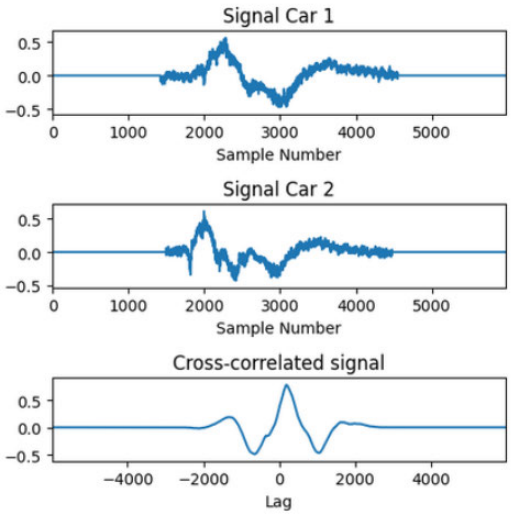


FIGURE 3. (Top) Car 1 DAS signal. (Middle) Car 2 DAS signal. (Bottom) A cross-correlation plot between both signals, when both cars are moving at the same speed across the same bin.

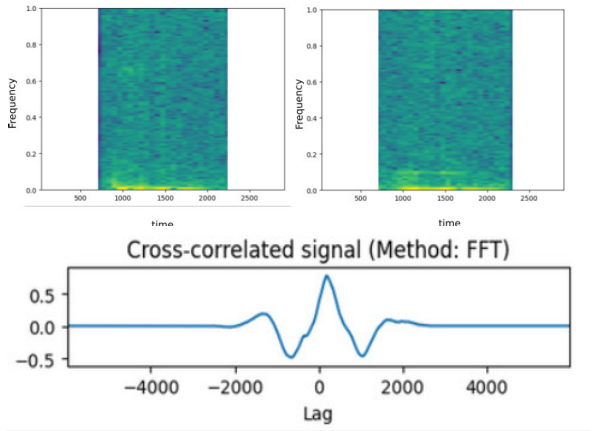


FIGURE 4. Top: Spectrograms of DAS signals generated by Car 1 (left) and Car 2 (right). The x-axis represents the time, the y-axis the frequency, and the colormap indicates the corresponding intensity in dB. Bottom: Both signals are processed with fast fourier transform and the resulting frequency components are cross-correlated using a convolution operation. The result is presented in the time domain.

Next, we examine the frequency domain representation by visually examining the spectrograms of both signals which are difficult to differentiate, as shown in Figure 4 (Top). Next, we feed the Fast Fourier Transform results of both signals in Figure 3 (Top) to the cross-correlation algorithm. The outcome is shown in Figure 4 (Bottom) which again shows high correlation and matches the finding from the time-domain cross-correlation shown in Figure 3 (Bottom).

These results suggest that, unlike traditional acoustic signals, DAS signals cannot be differentiated by examining the features corresponding to the frequency components. Next, we follow the EMD method in [18] to manually extract features, followed by an XGboost classifier. The resulting classification accuracy achieved is 26%, which is similar to random guessing between the five possible classes. This finding further confirms that, although such a method is suitable to classify distinguishable events (e.g. knocking, climbing) [18], it is incapable of extracting the latent features that represent the individual vehicle signatures.

Another successful approach in acoustic signal processing is Mel-filter banks; a method to decompose a signal into separate frequency bands in a scale that mimics the nonlinear human perception of sound. The problem with Mel-filter banks is that these need to be tuned for different acoustic signals (piano, speech, language, etc). A recently published model, LEAF, harnesses advances in 1D-CNN to create dynamic filter banks that can be learned for any acoustic signal including acoustic scenes, birdsong, music, speech, language classification [25]. We applied LEAF to our DAS signals and obtained an accuracy of 22%, again similar to random guessing.

The findings obtained through the data exploration indicate that DAS signals, although categorised as acoustic, cannot be represented with known acoustic features that mostly relate to frequency components. This leads us to the next step in our study that taps into the potential of deep convolution networks to automatically extract the features for successfully representing our data.

V. METHODOLOGY

Based on the problem formulation in Section III, we present here our methodology for two sub-problems: sample labelling and sample classification.

A. SAMPLE LABELLING

Data labeling requires the raw data to first be parsed for event detection, event framing, and finally event labelling. In this section, we apply linear regression to the dataset ϱ for the detection and labelling of events of interest.

Looking at Figure 2, the red lines show that there is a linear relationship between fibre bin (y-axis) and fibre shot (x-axis) for all cars. This indicates that all target cars drove at the same speed in the same direction with fixed distance as part of the controlled experiment. The path of each car is highlighted by the red lines in Figure 2. This finding allows us to use linear regression in the labelling task. In this work, we focus the analysis on a section of the road covered by bins $250 \leq b \leq 750$, where each spans 0.68 metres. A default car window of 50 bins is considered. By plotting the signal every 50 bins, we can record manually the start shot $s_{b,s}$ and the end shot $s_{b,e}$ for every bin b. We then feed the data in two linear models to find the start and end shot of each car signal at any bin $250 \leq b \leq 750$. We get two pairs of coefficients: (c_1, c_2) for the linear model representing the start shot $s_{b,s} = c_1 \times b + c_2$ of each bin b and (c_3, c_4) for the end shot $s_{b,e} = c_3 \times b + c_4$.

Next, each car window is labelled as either $\omega = 0$ if the window contains an unwanted signal, or $\omega \in \{1, 2, 3, 4, 5\}$ if it contains one of the controlled cars, in which case the value of ω is determined based on the predefined order of the

vehicles (see Section II-B). Let $x_b(s) = \rho(b, s)$ for a given bin b, it can be said that $x_b(s)$ is a signal displacement in a given shot s, and $\mathbf{x}_b \in \mathbb{R}^{d^\omega}$ for $s_{b,s} \leq s \leq (s_{b,s} + d^\omega)$ is a one-dimensional car signal, i.e. a data sample such as

$$\mathbf{x}_b = \{\rho(b, s_{b,s}), \rho(b, s_{b,s+1}), \ldots, \rho(b, s_{b,e})\} \quad (2)$$

B. PROPOSED 1D-CNN MODEL

We first build a 1D-CNN to extract the sample features followed by a softmax function that completes the role of classification. The proposed 1D-CNN is shown in Figure 5, where for each input signal $\mathbf{x}_i$ with index $1 \leq i \leq B$, the softmax classifier produces an output $\mathbf{p}_i = \{p_{i,1}, .., p_{i,\omega}, .., p_{i,T}\}$ such that $0 \leq p_{i,\omega} \leq 1$ is the confidence level of input sample $\mathbf{x}_i$ matching signals of type ω.

The proposed 1D-CNN is composed of thirteen layers, as shown in Figure 5. The model can be decomposed into three parts: feature extraction, high-level feature combination and classification. There are three CNN blocks in feature extraction. A CNN block is formed by convolution, batch normalisation and average pooling layers. The third block did not have a AVGpooling because its output was too small to be further compressed. The output of the high-level feature combination block (Fully connected layers) is a features vector $\mathbf{z}_i = \{z_{i,1}, \ldots, z_{i,Z}\}$ that corresponds to each input sample $\mathbf{x}_i$ and where $Z = 64$. This features vector $\mathbf{z}_i$ is then the input of the softmax function.

During the model training, kernels are optimised through backpropagation (BP). BP calculates a partial derivative of a loss function; for softmax the loss function is based on the cross-entropy $\text{CE}(i) = -\log(p_{i,\omega})$ or negative log likelihood. In this case, $\omega = y_i$ is the ground truth label of sample $\mathbf{x}_i$ and $p_{i,\omega}$ is the confidence level that is produced by the softmax function (τ), as shown in Figure 5.

In this work, we split the dataset $\mathbf{X}$ into a training dataset and testing dataset with a ratio $(D^t : D^v) = (8 : 2)$. The CNN is trained for 100 epochs where each epoch is conducted by multiple iterations when the whole data is exposed. In each iteration a batch size of data (32 $\mathbf{x}_i$) is processed before the CNN weights are updated.

C. EVALUATION

Given the pioneering nature of this research, we first build on the feature exploration in Section IV by examining the ability the model to extract representative features. This is done by calculating the extracted features cross-correlation between DAS signals of each two different car types in a given dataset $\mathbf{X}$. Next, we adopt accuracy as the main performance evaluation metric of the data model and the confusion matrix for class-based metric. The output of the data model is determined by the softmax function, i.e. the predicted class $\hat{y}_i \in \{1, \ldots, T\}$. In this case, $\hat{y}_i = \text{argmax}(\mathbf{p}_i) = \omega$ such that, the confidence level $p_{i,\omega}$ is the largest of all values in $\mathbf{p}_i$. A signal is correctly classified when its predicted class $\hat{y}_i$ is equal to its ground-truth y_i. To calculate the accuracy $\mathcal{A}$, we simply take number of correctly predicted samples divided by number of total samples in the validation dataset $\mathcal{D}^v$.

$$\mathcal{A} = \frac{\sum_{i=1}^{D^v} (\hat{y}_i = y_i)}{D^v} \quad (3)$$

We randomly split train/test set and validate our model for 10 times.

Ten different trained models are obtained in order to get a better picture of the general performance/accuracy rather than relying on a single model. Each model selects a number of batches of random samples from $\mathcal{D}^t$. The number of steps for each model is $D^t/32$, where 32 is the batch size. Therefore, when the model reaches the final step, the complete dataset $\mathcal{D}^t$ would have been exposed to the model for training. Similarly for the evaluation phase, $D^v/32$ batches of 32 samples each were selected until the complete validation dataset $\mathcal{D}^v$ is covered, where D^v is the number of data samples in $\mathcal{D}^v$.

VI. RESULTS AND ANALYSIS

In this paragraph, we present our results in which we measure the efficacy of the proposed method to identify a car type under different conditions. The 1D-CNN model used for vehicle type classification (see Figure 5) consists of around 40, 000 trainable parameters; when adjusted for vehicle size classification, the parameters drop just below 40, 000. The estimated computation time for two models is 37.36 and 46.77 seconds, respectively, training both for 100 epochs on a NVIDIA A100-PCIE-40GB GPU. The total size of our dataset is 563 MB. The labelling results are first summarised in Table 2.

TABLE 2. Number of samples for each dataset and for each car type.

Car Type\Dataset	$\mathbf{X}_{30}$	$\mathbf{X}_{40}$	$\mathbf{X}_{50}$	$\mathbf{X}_{60}$	$\mathbf{X}_{70}$
Car 1	0	451	711	451	551
Car 2	501	352	661	451	551
Car 3	501	451	810	451	551
Car 4	501	352	860	451	551
Car 5	501	352	761	451	551

In Section VI-A, we conduct a model representation study based on the features identified by each model. The cross correlation is conducted on dataset $\mathbf{X}_{60}$.

In Section VI-B, we present the results of classification for each individual data set $\mathbf{X}_v$ for $v \in \{30, 40, 50, 60, 70\}$ km/h and for the joint speed dataset $\mathbf{X}_{JS}$.

In Section VI-C, we regroup the five different vehicle types to two super-types: Large and Small (see Table 1). Hence, we reduce the number of classes from the existing $T = 5$ to $T' = 2$. This is motivated by the need to detect the size of vehicles rather than their make and model in urban ITS applications.

A. FEATURE REPRESENTATION

We first examine how well the proposed model represents the dataset through the extracted signature-specific features of each vehicle. To this end, we calculate the cross-correlation

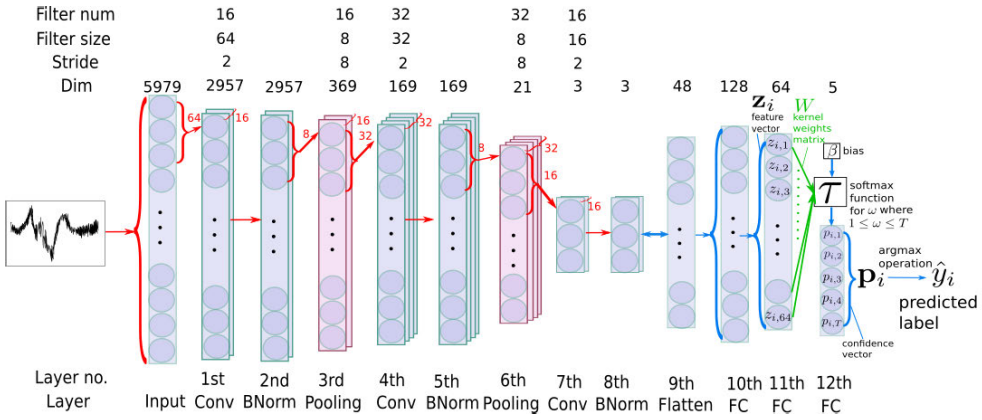


FIGURE 5. In this figure we show proposed 1D-CNN architecture for DAS signal classification. Number of filters, filter size, stride and dimension for each layer are presented.

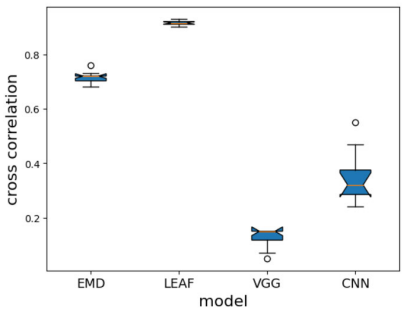


FIGURE 6. In this figure we show the cross-correlation between DAS signals corresponding to different car types for the four models: EMD, LEAF, 2D-VGG, and 1D-CNN.

of the model's features corresponding to each DAS signal of each of the five vehicles contrasted with the remaining four. In other words, we extract all DAS samples that correspond to each of the vehicles, based on the method in Section V-A in dataset $\mathbf{X}_{60}$ and calculate the correlation between $x_b^{\omega 1}$ and $x_b^{\omega 2}$ for all $b \in B$ and $\{\omega 1, \omega 2\} \in [1, ..T]$ such that $\omega 1 \neq \omega 2$. The results are shown in Figure 6 in which we present the correlation results of the four models: EMD, LEAF, 2G-VDD, and the proposed 1D-CNN. The results further confirm the inability of EMD (0.72 correlation) and LEAF (0.92 correlation) to represent the unique DAS signature of each vehicle. The 2D-VGG model seems to result in the least correlation between DAS signals generated by different vehicles with an average 0.13 whereas the proposed 1D-CNN results in an average 0.35. Both correlation outcomes of 2D-VGG and 1D-CNN are promising. It is anticipated that a two-dimensional DAS sample (as in the 2D-VGG model) would contain more information about the passing vehicle than a 1D-CNN since it captures the DAS data along the length of the vehicle [26]. However, the complexity of the 2D-VGG (i.e. the number of tunable parameters) is much higher than the number of data samples available to this research, thus is likely to result in overfitting. In the next section, we measure the classification of both models: 2D-VGG and 1D-CNN.

B. CLASSIFICATION RESULTS OF VEHICLE MODEL

The methodology described in Section V is applied to the five datasets in Table 2 and the hybrid dataset $\mathbf{X}_{JS}$. For this purpose, the dataset is split into 80% for training and 20% for testing; the testing results are summarised in Figure 7. The proposed method successfully classifies samples from any of the five vehicle types for both fixed speed and joint speed datasets with accuracy ranging from 82% to 99%. The worst accuracy of 82% is registered with $\mathbf{X}_{JS}$ and is a boxplot outlier (see Figure 7).

The 2D-VGG model is tested on the $\mathbf{X}_{JS}$ and, as expected, results in worse accuracy of 71% in the validation phase. This confirms that, due to its complexity, this model is not suitable for the given problem with limited data.

The average confusion matrices for each dataset are calculated as the average of the 10 models. Due to space limitations, only the confusion matrix resulting from $\mathbf{X}_{JS}$ is shown in Figure 8 (Left). It can be seen from the class-based accuracy figures that the model performs equally well (87–92%) for all car types except for the minority class Car 1 (83%). The Car 1 data in $\mathbf{X}_{30}$ was dropped manually because of an occlusion caused by an anonymous vehicle. This can be improved with further data collection or investigation on how to clean data when the signal of a target car overlaps with an anonymous car's. However, the results demonstrate that DAS signals include detailed latent information about the respective events. Such information can be used to track a particular vehicle in a wide area.

C. CLASSIFICATION RESULTS OF VEHICLE SIZE AND WEIGHT

We simplified the five-class classification problem to two-class classification. The first super group, *Large* consists of Car 1 and Car 5 as they are the top two heaviest vehicles in Table 1. The second super group, *Small*, includes the

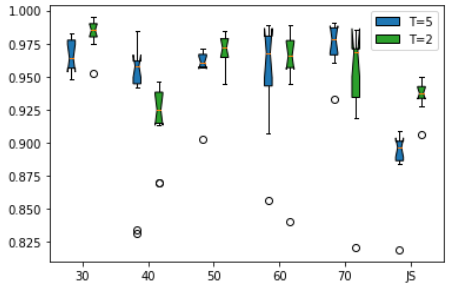


FIGURE 7. Classification results for $T = 5$ and $T = 2$. The boxes represent the inter-quartile range, the whiskers represent the most extreme, non-outlier data points, and the fliers represent the outliers. The horizontal lines of each box represent the median value.

remaining Car 2, Car 3, and Car 4. For this purpose, the dataset is split into 67% for training and 33% for testing; the testing results are summarised in Figure 7. As a result of re-organising the five existing labels into two groups where the first includes two labels and the second the remaining three, a data imbalance is created (Imbalance ratio is $1684/2656 = 0.63$). It follows that the data split of 80% : 20% adopted for the problem in Section VI-B gives under-par performance and the proposed 67% : 33% yields a better accuracy. The results are shown in Figure 7. A general trend of high accuracy shows that the DAS signal contains information about the car size (related to weight) in addition to that of the car type. The accuracy of classification in this case ranges from 81% to 98%.

Similarly to $T = 5$, the average confusion matrices for each fixed speed dataset and joint speed are calculated as the average of the ten $T = 2$ models. The confusion matrix resulting from $\mathbf{X}_{JS}$ is shown in Figure 8 (Right) and shows that the model performs slightly better in super group 2 (small car). Although an imbalanced data ratio between Large:Small (4759:8567) diminished the accuracy of minority class, the results are very encouraging and present a reliable means for the usage of DAS signal in ITS to detect and control the type of vehicles moving along different road types.

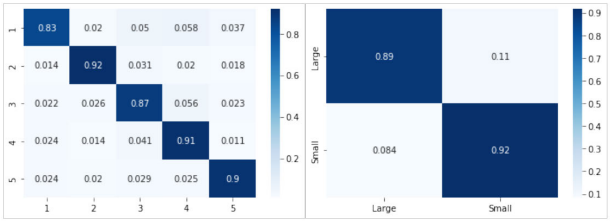


FIGURE 8. Average confusion matrix $\mathbf{X}_{JS}$: (Left) $T = 5$, (Right) $T = 2$.

VII. DISCUSSION

It should be noted that the data analysed in this work was collected under controlled conditions and therefore, the obtained results cannot be directly generalised without further examination. There are four aspects of this study that should be underlined in this context.

Firstly, the controlled experiment limited unwanted movements along the test road. Nonetheless, the proposed method is still successful in distinguishing the vehicle types despite external activities such as those created by other cars as shown in Figure 2. This was possible given the well-labeled data based on the a priori knowledge of the wanted car and speed. In the absence of such information, a distinct and uncontaminated DAS signature of a target car would be needed before hand. It would then be possible to re-identify this signature in the presence of other moving vehicles. Given the length of the DAS system spanning tens of kilometers, this is a realistic assumption to have in an urban environment.

Secondly, the dataset studied in this work is limited to five vehicle models only and cannot be readily generalised to distinguishing any vehicle model roaming the streets today. The applications that require identifying a specific vehicle, such as airports, manufacturing sites, energy plants, and similar entail a well determined list of authorised vehicles in contrast with undetermined list in open urban environments. It follows that this pioneering work can be expanded to represent the list of cars of interest in each application and would result in an effective, efficient, and reliable means of detecting any intruders and of tracking known vehicles. In the case of detecting the size of a vehicle for conditional access in urban applications, expanding the list of vehicle models in the dataset would likely result in a more representative model. Nonetheless, this study demonstrates the potential of DAS signals in estimating the size of a vehicle and offers a promising methodology for extracting this information given a larger list of vehicles.

Thirdly, given that the data was collected along the same fibre line and road, the obtained results cannot be generalised to any road and any DAS system. Indeed, the physical characteristics of the fibre used, the road type, and the depth of the fibre are known to impact the DAS signal. It follows that, for each deployed DAS system, the model should be re-trained to account for these physical characteristics and would yield a highly reliable method for detecting and tracking vehicles along the roads spanned by the system.

Lastly, it should be noted that the data used in Section VI were collected on a dry day, thus, it is not possible to fully analyse the effect of rainfall or other weather conditions. In [27] the authors suggest that the rainfall results in an acoustic signal of high intensity background noise and its energy spreads unevenly across frequency spectrum. It is expected that such spectral characteristics of noise could be filtered out using typical signal processing methods, e.g. low-band filters. The impact of rain and surface material on the DAS signal is examined in [28]. The work concludes that DAS signals collected under a dry surface have higher average signal-to-noise ratio than on a wet surface on a rainy day. Based on these two works, we anticipated that the effect of rain on the DAS signal is likely to increase the noise level but believe that

it could be treated, to a large extent, with signal processing methods that would target the suppression of noise generated by the rainfall. However, this can only be confirmed with representative data and further experiments which are part of our future work in this area.

In summary, the importance of this pioneering investigation is not limited by nor confined to the examined dataset. Instead, the field trial data has enabled this study and revealed the potential of DAS in addressing ITS problems related to moving vehicles. More importantly, DAS was shown as an alternative uninterrupted data source for ITS that covers tens of kilometers of road stretches and that is resilient to issues such as weather, visibility, and luminosity often hindering camera-based data. DAS is thus a robust data source that contains ITS-related information and is intrinsically GDPR-compliant since it does not include sensitive information (faces, clothing, etc). In this work, we shed light on the potential role of DAS in informing ITS problems as a standalone data source. Future work will examine how camera-based sensors could be used to label DAS signals in a given location and how DAS systems can then be used to track the movement of the identified vehicle over tens of kilometers.

Indeed, the method proposed presumes the possibility of data labelling, hence, harnesses the potential of supervised learning. The authors believe that data labelling for the dataset at hand is facilitated by the conditions of the controlled experiment but may be more challenging in real-traffic scenarios. To this end, we propose to employ an alternative data source to assist with the labelling in order to fully leverage the potential of DAS in ITS applications (e.g. in urban scenarios for restricted vehicle assess). Such a joint solution would benefit from the rich information provided by a single camera-based data source with controlled lighting for labelling and from the efficacy of DAS in monitoring tens of kilometers of roads.

VIII. CONCLUSION

This work examines the potential role of distributed acoustic sensors (DAS) as an alternative data source for enabling intelligent transport systems (ITS) applications. In particular, the problem of recognising the model and the size of the vehicle is investigated based on field trials that involved five specific vehicles. We first conduct a feature analysis of the incurred DAS signals and demonstrate that, unlike traditional acoustic signals, these cannot be represented by time-domain and frequency domain manually engineered features. Instead, we propose a one-dimensional convolution neural network for feature extraction followed by a soft-max function for classification. The method successfully identifies one of five car types used in the controlled experiment with mean accuracy of 94%. The same method is applied to detect the size of the vehicles with binary classification of Small or Large. The results are encouraging with mean accuracy of 95%. The outcome of this pioneering work demonstrates the richness of DAS signals and their inherent role in identifying overground movement using a privacy-compliant approach.

ACKNOWLEDGMENT

The authors would like to thank the support of Fotech Solutions (www.fotech.com) and COMSA Corporación (www.comsa.com) for providing the dataset for this work.

REFERENCES

[1] Y. Jin, Z. Jia, P. Wang, Z. Sun, K. Wen, and J. Wang, "Quantitative assessment on truck-related road risk for the safety control via truck flow estimation of various types," *IEEE Access*, vol. 7, pp. 88799–88810, 2019.

[2] A. Fomin and S. Braeunig, "A wireless sensor system for traffic flow detection based on measurement of earth's magnetic field changes," in *Proc. 8th Medit. Conf. Embedded Comput. (MECO)*, Jun. 2019, pp. 1–4.

[3] B. Liu, Q. Li, D. Chen, and H. Sun, "Pattern recognition of vehicle types and reliability analysis of pneumatic tube test data under mixed traffic condition," in *Proc. 2nd Int. Asia Conf. Informat. Control, Autom. Robot. (CAR)*, Mar. 2010, pp. 44–47.

[4] C. Chen, B. Liu, S. Wan, P. Qiao, and Q. Pei, "An edge traffic flow detection scheme based on deep learning in an intelligent transportation system," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 3, pp. 1840–1852, Mar. 2020.

[5] S. Agarwal, S. Mustavee, J. Contreras-Castillo, and J. Guerrero-Ibañez, "Sensing and monitoring of smart transportation systems," in *The Rise of Smart Cities*, A. H. Alavi, M. Q. Feng, P. Jiao, and Z. Sharif-Khodaei, Eds. London, U.K.: Butterworth-Heinemann, 2022, ch. 20, pp. 495–522. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780128177846000102

[6] M. H. Sharif, "Laser-based algorithms meeting privacy in surveillance: A survey," *IEEE Access*, vol. 9, pp. 92394–92419, 2021.

[7] D. Vij and N. Aggarwal, "Transportation mode detection using cumulative acoustic sensing and analysis," *Frontiers Comput. Sci.*, vol. 15, no. 1, Feb. 2021, Art. no. 151311. [Online]. Available: https://journal.hep.com.cn/fcs/EN/abstract/article_25717.shtml

[8] S. Ntalampiras, "Moving vehicle classification using wireless acoustic sensor networks," *IEEE Trans. Emerg. Topics Comput. Intell.*, vol. 2, no. 2, pp. 129–138, Apr. 2018.

[9] J. Zuo, Y. Zhang, H. Xu, X. Zhu, Z. Zhao, X. Wei, and X. Wang, "Pipeline leak detection technology based on distributed optical fiber acoustic sensing system," *IEEE Access*, vol. 8, pp. 30789–30796, 2020.

[10] J. Tejedor, H. F. Martins, D. Piote, J. Macias-Guarasa, J. Pastor-Graells, S. Martin-Lopez, P. C. Guillén, F. De Smet, W. Postvoll, and M. González-Herráez, "Toward prevention of pipeline integrity threats using a smart fiber-optic surveillance system," *J. Lightw. Technol.*, vol. 34, no. 19, pp. 4445–4453, Oct. 1, 2016.

[11] Z. H. Warsi, S. M. Irshad, F. Khan, M. A. Shahbaz, M. Junaid, and S. U. Amin, "Sensors for structural health monitoring: A review," in *Proc. 2nd Int. Conf. Latest trends Electr. Eng. Comput. Technol. (INTELLECT)*, 2019, pp. 1–6.

[12] K. Hicke and K. Krebber, "Towards efficient real-time submarine power cable monitoring using distributed fibre optic acoustic sensors," in *Proc. 25th Optical Fiber Sensors Conf. (OFS)*, Apr. 2017, pp. 1–4.

[13] H. Liu, J. Ma, W. Yan, W. Liu, X. Zhang, and C. Li, "Traffic flow detection using distributed fiber optic acoustic sensing," *IEEE Access*, vol. 6, pp. 68968–68980, 2018.

[14] H. Liu, J. Ma, T. Xu, W. Yan, L. Ma, and X. Zhang, "Vehicle detection and classification using distributed fiber optic acoustic sensing," *IEEE Trans. Veh. Technol.*, vol. 69, no. 2, pp. 1363–1374, Feb. 2020.

[15] Y. Wang, P. Wang, K. Ding, H. Li, J. Zhang, X. Liu, Q. Bai, D. Wang, and B. Jin, "Pattern recognition using relevant vector machine in optical fiber vibration sensing system," *IEEE Access*, vol. 7, pp. 5886–5895, 2019.

[16] C. Xu, J. Guan, M. Bao, J. Lu, and W. Ye, "Pattern recognition based on time-frequency analysis and convolutional neural networks for vibrational events in ϕ-OTDR," *Opt. Eng.*, vol. 57, no. 1, pp. 1–7, 2018, doi: 10.1117/1.OE.57.1.016103.

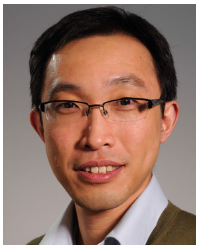
[17] F. Jiang, H. Li, Z. Zhang, and X. Zhang, "An event recognition method for fiber distributed acoustic sensing systems based on the combination of MFCC and CNN," in *Proc. SPIE*, 2018, pp. 15–21, doi: 10.1117/12.2286220.

[18] Z. Wang, S. Lou, S. Liang, and X. Sheng, "Multi-class disturbance events recognition based on EMD and XGBoost in ϕ-OTDR," *IEEE Access*, vol. 8, pp. 63551–63558, 2020.

[19] H. Wu, J. Chen, X. Liu, Y. Xiao, M. Wang, Y. Zheng, and Y. Rao, "One-dimensional CNN-based intelligent recognition of vibrations in pipeline monitoring with DAS," *J. Lightw. Technol.*, vol. 37, no. 17, pp. 4359–4366, Sep. 1, 2019.
[20] E. Catalano, A. Coscetta, E. Cerri, N. Cennamo, L. Zeni, and A. Minardo, "Automatic traffic monitoring by φ-OTDR data and Hough transform in a real-field environment," *Appl. Opt.*, vol. 60, no. 13, pp. 3579–3584, May 2021. [Online]. Available: https://opg.optica.org/ao/abstract.cfm?URI=ao-60-13-3579
[21] N. Zeghidour, O. Teboul, F. De Chaumont Quitry, and M. Tagliasacchi, "LEAF: A learnable frontend for audio classification," 2021, *arXiv:2101.08596*.
[22] C. E. Kayan, K. Y. Aldogan, and A. Gumus, "An intensity and phase stacked analysis of Φ-OTDR system using deep transfer learning and recurrent neural networks," 2022, *arXiv:2206.12484*.
[23] A. Hartog, *An Introduction to Distributed Optical Fibre Sensors*. Boca Raton, FL, USA: CRC Press, 2017.
[24] V. Handerek, "Distributed optical fibre sensor," patent US 9 304 017 B2, Feb. 16, 2016.
[25] N. Zeghidour, O. Teboul, F. De Chaumont Quitry, and M. Tagliasacchi, "LEAF: A learnable frontend for audio classification," in *Proc. Int. Conf. Learn. Represent.*, 2021, pp. 1–16. [Online]. Available: https://openreview.net/forum?id=jM76BCb6F9m
[26] C.-Y. Chiang, M. Jaber, and P. Hayward, "A distributed acoustic sensor system for intelligent transportation using deep learning," 2022, *arXiv:2209.05978*.
[27] C. Sánchez-Giraldo, C. L. Bedoya, R. A. Morán-Vásquez, C. V. Isaza, and J. M. Daza, "Ecoacoustics in the rain: Understanding acoustic indices under the most common geophonic source in tropical rainforests," *Remote Sens. Ecology Conservation*, vol. 6, no. 3, pp. 248–261, Sep. 2020, doi: 10.1002/rse2.162.
[28] K. E. Winters, M. C. Quinn, and J. R. Piccuci, *Diurnal Changes Signal-to-Noise Ratio in a Distributed Acoustic Sensing System*. Los Angeles, CA, USA: Geo-Congress, 2022, pp. 74–81, doi: 10.1061/9780784484067.008.

CHIA-YEN CHIANG is currently pursuing the Ph.D. degree with the Queen Mary University of London, with a focus on AI applications for smart cities. She was an Intern Researcher with 2Excel Geo for her master's thesis project. She has taken part in a number of projects, including forest analysis in high-resolution imagery, people re-identification in live CCTV images, and deep learning on cloud computing. Her current research interests include deep learning and machine learning modeling in multi-media resources, such as video and audio, including distributed acoustic sensing (DAS) time series, in the context of smart cities. Her research interest includes computer vision.

MONA JABER (Senior Member, IEEE) received the B.E. degree in computer and communications engineering and the M.E. degree in electrical and computer engineering from the American University of Beirut, Lebanon, in 1996 and 2014, respectively, and the Ph.D. degree from the 5G Innovation Centre, University of Surrey, in 2017. Her Ph.D. research was on 5G backhaul innovations. She was a telecommunication consultant in various international firms with a focus on the radio design of cellular networks, including GSM, GPRS, 3G, and 4G. She led the IoT Research Group, Fujitsu Laboratories of Europe, from 2017 to 2019, where she researched IoT-driven solutions for the automotive industry. She is currently a Lecturer in IoT with the School of Electronic Engineering and Computer Science, Queen Mary University of London. Her research interests include zero-touch networks, the intersection of ML and IoT in the context of sustainable development goals, and the IoT-driven digital twins. She received the title of N2Women Rising Star in Computer Networking and Communications, in 2022.



KOK KEONG CHAI received the B.Eng. (Hons.), M.Sc., and Ph.D. degrees, in 1998, 1999, and 2007, respectively. He joined the School of Electronic Engineering and Computer Science (EECS), Queen Mary University of London (QMUL), in August 2008. He is currently a Professor in the Internet of Things, the Queen Mary Director of Joint Programme with the Beijing University of Posts and Telecommunications (BUPT), and a Communication Systems Research Group Member of QMUL. He has authored more than 65 technical journals and conference papers in his research areas. His current research interests include sensing and prediction in distributed smart grid networks, smart energy charging schemes, applied blockchain technologies, dynamic resource management, wireless communications, and medium access control (MAC) for M2M communications and networks.

JONATHAN LOO (Member, IEEE) received the M.Sc. degree (Hons.) in electronics and the Ph.D. degree in electronics and communications from the University of Hertfordshire, Hertfordshire, U.K., in 1998 and 2003, respectively. He is currently a Chair Professor in computing and communication engineering with the School of Computing and Engineering, University of West London, U.K. He has successfully graduated 18 Ph.D. students and has coauthored over 350 journals and conference papers in his research areas. His research interests include machine learning and AI, natural language and image processing, the IoT/cyber-physical systems, cyber-security, information-centric networking, wireless/mobile systems, and communication networks.