Lightweight Continuous Authentication via IMU Fingerprinting for V2X

Bei Gong, *Member, IEEE,* Zhe Li, Mowei Gong, Haotian Zhu, Weizhi Meng, *Senior, IEEE,* and Chong Guo

***Abstract*—Inertial measurement unit (IMU) fingerprinting is a promising physical authentication technique based on hardware imperfections produced during sensor manufacturing. This paper presents a two-stage feature extraction process that combines feature selection and mapping; the proposed approach is tailored for the lightweight vehicle-to-everything (V2X) application scenario. Specifically, the selected features are transformed into images via Gramian angular difference field (GADF), Gramian angular summation field (GASF), and Markov transition field (MTF) mappings, as well as feature extraction implemented via a convolutional neural network (CNN). Owing to the advances provided by the proposed scheme, a lightweight feature extraction system achieves satisfactory accuracy levels above 99.10% with fewer sample data and a short training time. The effectiveness and robustness of the developed approach were validated under various driving conditions via 20 IMU sensors, Arduino, and a Raspberry Pi across 20 vehicles. Additionally, tests conducted across different deep learning models demonstrated the generalizability of the proposed preprocessing and mapping methods.**

***Index Terms*—IMU fingerprint authentication, vehicle-to-everything (V2X), lightweight, feature mapping.**

I. Introduction

As Internet of Vehicles (IoV) technology has rapidly advanced, vehicles are increasingly becoming highly intelligent mobile information platforms, making vehicle-to-everything (V2X) communication crucial. These vehicles not only enable high-speed data exchanges and communications but also intelligently interact with urban infrastructure, other vehicles, and even pedestrians, thereby increasing road safety and traffic efficiency. However, the evolution of the IoV, along with the maturity of autonomous and driverless driving technologies, has also introduced several information security issues, including data breaches, illegal intrusions, and malicious attacks on driving systems, which pose serious threats to the safety of vehicles and passengers [1]. These challenges urgently demand the development of more advanced authentication mechanisms for ensuring the security and continuity of vehicle communications in the dynamic V2X environment while accommodating the technological progression from traditional to automated and assisted driving, thus safeguarding vehicles and passengers across all driving modes.

Currently, vehicle identity authentication typically employs several methods: cryptographic authentication, such as public key infrastructure (PKI) [2] and symmetric key technologies [3], relies on solving complex mathematical problems. However, these traditional methods are susceptible to quantum algorithms, which can quickly decrypt these problems, thereby overcoming existing encryption techniques. Additionally, trust-based authentication is conducted by evaluating the trust level of devices or users within the network. But trust relationships are often chained, and if any link in the chain is attacked or compromised, the entire trust chain may collapse [4]. There are also biometric or behavioral authentication methods, such as analyzing the driver's eye movement [5] or driving behaviors [6], like turning [7]. As assisted and autonomous driving technologies become more widespread, traditional verification methods that only target drivers are becoming increasingly insufficient for ensuring vehicle safety. Given the increasing complexity of attacks and the inherent vulnerabilities of traditional cryptographic and biometric methods in the evolving landscape of vehicular technology, there is a pressing need to explore authentication systems that are capable of adapting to advanced threats.

In response to this need, physical unclonable function (PUF) technology, using inertial measurement units (IMUs) as its basis, offers a potential solution. This approach leverages the inherent defects in the manufacturing of IMUs—sensors integral to modern vehicular systems. These IMU fingerprints characterized by high unpredictability and uniqueness, are suitable for generating encryption keys and authentication data that are difficult to replicate and predict. Additionally, non-cryptographic techniques based on a physical layer inherently have advantages in terms of handling delays, computational loads, and communication overhead. In [8] [9], microelectromechanical system (MEMS) devices were proven to be effective means of authentication. IMUs, which integrate MEMS accelerometers, gyroscopes, and magnetometers, are extensively employed in modern automotive technologies. IMUs are crucial for advanced driver assistance systems (ADASs), providing accurate positioning and navigation functions that are crucial for autonomous driving technologies.Notably, the

(Corresponding author: Weizhi Meng and Chong Guo.)

Bei Gong is with the Beijing Key Laboratory of Trusted Computing, College of Computer Science, Beijing University of Technology, Beijing 100124, China (e-mail: gongbei@bjut.edu.cn).

Zhe Li, Mowei Gong, Haotian Zhu and Chong Guo are with the College of Computer Science, Beijing University of Technology, Beijing 100124, China (e-mail: leezhe627@emails.bjut.edu.cn; gongmowei@emails.bjut.edu.cn; zhuhaotian@emails.bjut.edu.cn; chongguo@emails.bjut.edu.cn).

Weizhi Meng is with the School of Computing and Communications, Lancaster University and the Department of Applied Mathematics and Computer Science, Technical University of Denmark, Denmark. (emails: weizhi.meng@ieee.org)

integration of IMUs with the Global Navigation Satellite System (GNSS) is the key to achieving precise localization in environments with weak or unavailable signals [10] [11]. Our goal is to establish an IMU-based continuous vehicle authentication system that uses physical fingerprint technology to automatically and continuously monitor vehicle states without interfering with normal driving operations. This authentication method, which is based on physical properties, heralds new possibilities for vehicle safety, indicating the direction of future vehicle authentication technologies.

A. Related Work

1) MEMS hardware fingerprinting technology, especially methods based on PUFs, has been shown to have significant efficacy in security validation scenarios involving different devices, including accelerometers, gyroscopes, magnetometers, pressure sensors, and microphones. Dey et al. [12] first demonstrated a device authentication method using accelerometers in mobile devices, explicitly establishing the viability of MEMSs for identity verification applications. Lee et al. [13] introduced a novel method with a challenge-response structure for acoustic signal-based MEMS sensor fingerprint recognition. Shen et al. [14] developed a MEMS fingerprint technology called MotoPrint, which utilizes vibration motors. They also created a comprehensive device authentication system using MotoPrint, tailored for mobile payment systems. Abdolinezhad et al. [15] introduced a lightweight circuit based on MEMS with a piezoresistive bridge functioning as a weak PUF. The secret keys generated by this system successfully met the randomness standards set by the National Institute of Standards and Technology (NIST). Fereidooni et al. [16] presented AuthentiSense, a scalable and efficient behavioral biometrics authentication system that is user-agnostic. It leverages MEMS sensors to enable continuous authentication with a high accuracy of up to 97%. Additionally, Kussl et al. [17] engineered an innovative in-road sensor system using MEMS magnetic sensors. This system is capable of identifying and classifying vehicles by type and model with an overall classification accuracy exceeding 90%.

These studies not only validate the feasibility of identity verification but also underscore the significant advantages and practicality of MEMS hardware fingerprint technology in identity recognition scenarios. However, MEMS authentication technologies have seldom been applied in V2X contexts. Moreover, these systems often face limitations when processing large volumes of data and require extensive signal inputs and lengthy training periods for rapid response capabilities. This is particularly true in V2X environments, where such constraints may limit the practicality of these applications. Our approach builds on these findings and aims to overcome these limitations by offering a more lightweight and efficient solution for V2X scenarios.

2) Appropriately representing signals is crucial for enhancing the performance of IMU hardware-based fingerprint systems. Effectively representing time series signals enhances not only the discriminative capabilities of the utilized system but also its adaptability to complex environmental changes. Extensive research has explored various signal representation methods. One commonly used method is the recurrence plot (RP) technique, which represents time series data points in the phase space as a two-dimensional image. RPs can reveal the periodic and aperiodic structures within dynamic systems, making them particularly effective at recognizing complex dynamic behaviours [18]. Hatami et al. [19] transformed time series into 2D texture images via RPs and classified them with deep CNNs. Their results demonstrated that this method is competitively accurate. Baldini et al. [20] found that combining a deep CNN with RP images significantly outperforms traditional methods based on raw time series dissimilarity. A spectrogram is another powerful tool that analyses the spectral contents of signals as their frequencies change over time. Qi et al. [21] investigated the feasibility of using spectrograms as a signal representation method for LTE-V2X PLA, achieving up to 97.30% identification accuracy on thirteen identical and available RF devices within a short training period. Peralta-Braz et al. [22] proposed that CNN AlexNet be trained to identify traffic speed labels from voltage continuous wavelet transform (CWT) images, efficiently extracting essential features from the input image while reducing the utilization of the number of CWT images.

Despite the significant advancements provided by various signal representation methods, these techniques still face challenges, particularly in practical applications within V2X environments. First, their adaptability to environmental noise and dynamic changes remains limited. Second, most methods demand substantial computational resources, which can limit them when applied to devices with constrained computational capabilities. Our work proposes a hybrid signal representation method, improving accuracy while maintaining computational efficiency.

3) In recent years, hardware-based continuous authentication schemes have garnered increasing attention in the field of vehicular safety. Mekki et al. [23] developed a driver identification method that combines smartphone sensors with the OBD-II protocol, analysing the input data through a CNN and a recurrent neural network (RNN)/LSTM. This method was validated on real-world datasets via cross-validation techniques and was implemented within the Linux framework. Wang et al. [7] introduced a novel driver identification method that uses sensors in mobile devices to analyse the acceleration and angular velocity changes that occur during turns, employing a gradient boosting tree (GBT) classifier to achieve over 98% accuracy. Furthermore, Song et al. [24] proposed a multifactorial continuous authentication mechanism that uses facial recognition on smartphones coupled with a Bluetooth protocol to effectively perform real-time authentication; this approach offers protection against various attacks. Xun et al. [6] proposed a method that utilizes automotive diagnostic tools to collect data on vehicle operation. They then build a composite model using convolutional neural networks and support vector domain description, which effectively facilitates the fingerprint recognition of car drivers. The model has achieved an identification accuracy of 98.6% for 15 authorized drivers and has successfully detected unauthorized driving

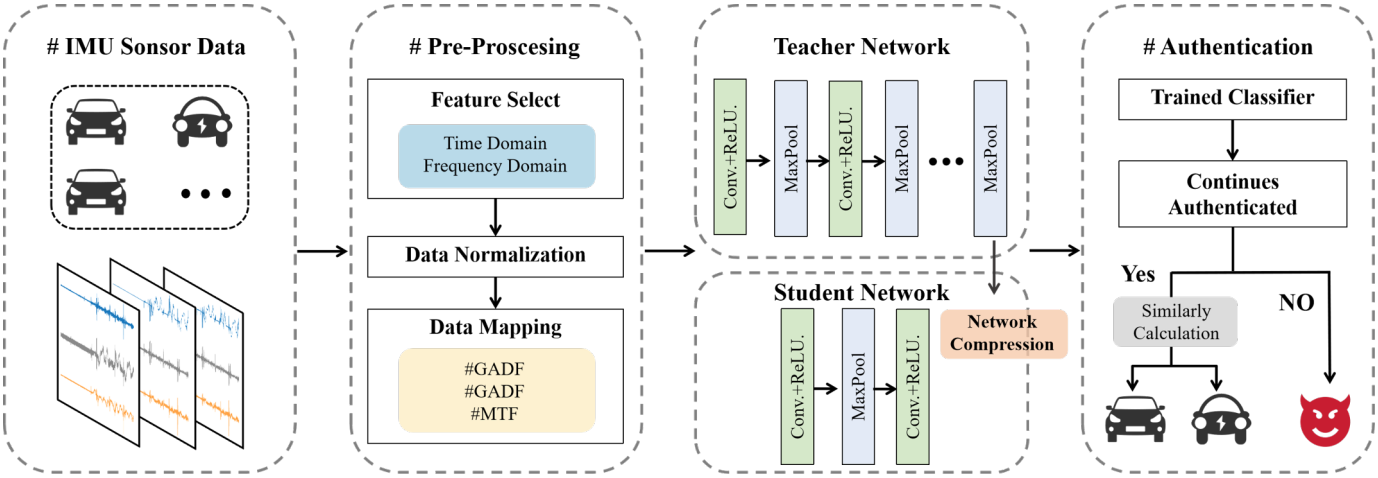


Fig. 1. Diagram of the proposed IMU-based system.

with an accuracy of 98.9% in five cases.

Although these methods have demonstrated high precision and robustness, they depend on additional hardware or require specific cooperation from drivers, which could interfere with driving or limit their applicability when such hardware resources are unavailable. Simultaneously, they still face limitations in adapting to future autonomous driving, where there is no driver. Our system, in contrast, does not rely on external hardware or explicit driver interactions, and it more adaptable to future V2X environments, including autonomous driving.

B. Main Contributions

Based on the above three issues, this paper introduces a novel continuous authentication system for V2X scenarios that uses IMU-based fingerprints. The main contributions of this work are fourfold.

1) We propose a two-stage feature extraction method that combines feature selection and deep learning to capture IMU hardware fingerprints. This approach reduces the number of required training samples and enhances the efficiency of data processing.

2) We propose a hybrid signal representation scheme that maps signal features onto a Gramian angular difference field (GADF), a Gramian angular summation field (GASF), and a Markov transition field (MTF). This method describes signal feature structures and variations from multiple perspectives. It achieves more than 99% authentication accuracy in complex driving environments.

3) The lightweight concept runs through our entire design architecture, and it is aimed at enhancing the suitability of the proposed approach for widespread deployment within real-world V2X environments. IMU sensors collect less data to ensure a lightweight input size for the model. A CNN is then applied to extract features, and through knowledge distillation, a lightweight student model achieves comparable performance while reducing complexity. The complexity and efficiency of our design are measured in terms of its time consumption, memory, number of floating-point operations (FLOPs), and number of multiply-add (MADD) operations.

4) The authentication system is comprehensively evaluated under various vehicle conditions and in multiple real-world environments. The experimental results demonstrate that this system is a highly promising solution for implicit user authentication, exhibiting robust performance across diverse environments without disrupting driving operations. To our knowledge, this study is the first to conduct such an extensive analysis in a range of practical situations.

The rest of the paper is organized as follows. Section II provides an overview of the system design and discusses the signal collection and preprocessing steps. Section III introduces the mapping processing of feature. Section IV details the employed deep learning strategies. Section V provides the experimental results. Section VI shows the generalizability of the proposed method to other DL models. Section VI discusses the limitations and future work. Finally, Section VIII concludes the paper.

II. System Overview and Data Preprocessing

In this section, we describe the System architecture, threat model and how to collect data for our system, preprocess the collected data, and then map the obtained feature data.

A. System Architecture

As mentioned before, this paper aims to build a lightweight IMU-based system for V2X environments by focusing on three issues of training data dependency, feature mapping, and model complexity. As shown in Fig. 1, the process includes data acquisition, preprocessing, network compression, and authentication stages. The data acquired from the IMU are processed through cubic spline interpolation, feature selection, and normalization before being mapped in image representations that help extract high-quality identity features. These representations are then processed through a CNN feature extraction module to authenticate and classify vehicles. To adapt to real-world applications, such as by reducing the

computational costs imposed on a roadside unit (RSU), the original CNN model is compressed via knowledge distillation, enabling a less computational student model to perform similarly to the teacher model. In the authentication phase, the RSU receives new data from the IMU sensors of the vehicle. The trained compressed classifier within the RSU then classifies these data, ensuring the reliability of the vehicle categorization process. Authentication tasks typically begin by checking the metric calculations concerning IMU fingerprints, followed by applying preset thresholds for admission judgement purposes. During classification, a classifier such as a softmax classifier is directly employed as the task head, selecting the candidate with the highest probability as the final prediction.

B. Threat Model

To explore potential security risks our system may encounter, focusing primarily on random, Spoofing and impersonation attacks. We assume that all vehicle IMUs are registered beforehand, which precludes attackers from compromising the registration phase. Consequently, attackers must rely on impersonation strategies to pass the authentication system.

Random Attacks: Lacking familiarity with legitimate driving patterns, attackers attempt to interact directly with RSUs or legitimate vehicles using IMUs of their own vehicl.

Spoofing Attacks: The attacker try to generate time-series signals resembling legitimate ones to perform the spoofing attack.

Impersonation Attacks: Before initiating assaults on RSUs or legitimate vehicles, attackers carefully observe and mimic the driving behaviors of legitimate vehicles. In some scenarios, they might install unauthorized IMUs to covertly collect dynamic vehicle data, which they then exploit to orchestrate attacks.

C. Data Acquisition

In the automotive industry, IMU sensors are typically placed near the centre of the coordinate system of a vehicle to better reflect the overall status of the vehicle. In our work, to more closely mirror real-world conditions, we positioned the IMU sensors near the centre of the vehicle as well. All the sensors were the same model from the same manufacturer (LSM9DS1, STMicroelectronics). They were installed in 20 different vehicles (15 gasoline vehicles and 5 electric vehicles). Data began to be collected from these sensors as soon as vehicle ignition occurred.

D. Data Preprocessing

Signal data preprocessing consists of two main steps: 1) feature selection and 2) a data normalization process.

Feature Selection: IMU sensors provide a 3×n matrix representing the data collected over a period, where x, y, and z correspond to readings derived from the three axes of each sensor. We segment these raw data into various sliding windows, computing the root sum of squares (RSS) for the accelerometer, gyroscope, and magnetometer readings within each window as follows:

$$I(k) = \sqrt{I_x^2(k) + I_y^2(k) + I_z^2(k)}. \tag{1}$$

Given that the sampling intervals are not uniform, the derived values are not equally distant, complicating the frequency-domain feature extraction step. To address this issue, we employ cubic spline interpolation to create equidistant data points for I(k).

The features we construct include both time-domain and frequency-domain features. Time-domain features are used to analyse signal characteristics over time, detecting time-related properties that might include critical unique identifiers provided by the sensors. Compared with time-domain features, frequency-domain features are better suited for extracting information about the frequency components and their behaviours within the sensor signals, dynamically revealing the contributions and distributions of these components. For each time window, we construct 21 statistical features, including 10 time-domain and 11 frequency-domain features in Table I, yielding a total of 63 features per window across all three sensors.

Data Normalization: To address the scale disparities among features and enhance the accuracy and learning efficiency of the model, we normalize the extracted features. Utilizing the min-max normalization method, each feature is scaled to a range of [0, 1]. The transformation formula is as follows:

$$y = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{2}$$

where x represents the original feature value and y is the normalized value. All normalized sensor data are then represented as $I_{norm} = [I_{a_{norm}}, I_{g_{norm}}, I_{m_{norm}}]$.

III. Feature Mapping

The GADF and GASF schemes are techniques for transforming time series data into 2D images [25]. A GADF constructs an image by calculating the cosine of the angular differences between two normalized time series points, whereas a GASF does so by calculating the cosine of their angular sums. These angles, θ_i and θ_j, represent the angles of any two points in the input time series (encoded in polar coordinates), and each sequence value is mapped to an angle to ultimately form a matrix. The mathematical expressions for the GADF and GASF techniques are as follows:

$$GADF(i,j) = \cos\Big(\arccos\left(\frac{x_i'}{\sqrt{\sum_{k=1}^{N} {x_k'}^2}}\right) - \arccos\left(\frac{x_j'}{\sqrt{\sum_{k=1}^{N} {x_k'}^2}}\right)\Big) \tag{3}$$

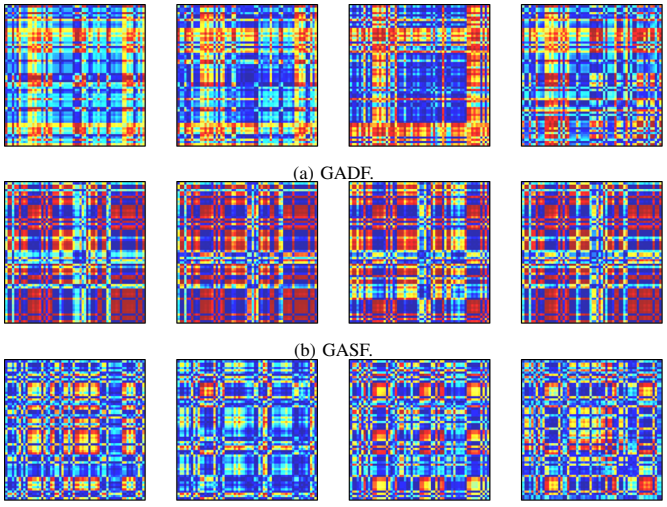
TABLE I
CNN STRUCTURE OF THE STUDENT MODEL

Domain category	Feature Name	Description
Time-domain feature	Mean	$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$
	Standard deviation	$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i-\mu)^2}$
	Average deviation	$\text{AD} = \frac{1}{N}\sum_{i=1}^{N}\lvert x_i-\mu\rvert$
	Skewness	$\text{Skew} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i-\mu}{\sigma}\right)^3$
	Kurtosis	$\text{Kurt} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i-\mu}{\sigma}\right)^4 - 3$
	RMS amplitude	$\text{RMS} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2}$
	Minimum	$\text{Min} = \min(x_1, x_2, \ldots, x_N)$
	Maximum	$\text{Max} = \max(x_1, x_2, \ldots, x_N)$
	Energy	$E = \sum_{i=1}^{N} x_i^2$
	Energy entropy	$\text{Entropy} = -\sum_{i=1}^{N} p_i \log(p_i) \quad \text{where } p_i = \frac{x_i^2}{E}$
Frequency-domain feature	Spec. standard deviation	$\sigma_{\text{spec}} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(X_i-\mu_{\text{spec}})^2}$
	Spec. skewness	$\text{Skew}_{\text{spec}} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{X_i-\mu_{\text{spec}}}{\sigma_{\text{spec}}}\right)^3$
	Spec. kurtosis	$\text{Kurt}_{\text{spec}} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{X_i-\mu_{\text{spec}}}{\sigma_{\text{spec}}}\right)^4 - 3$
	Spectral crest	$\text{Crest} = \frac{\max(X_1, X_2, \ldots, X_N)}{\mu_{\text{spec}}}$
	Spec. centroid	$C = \frac{\sum_{i=1}^{N} f_i X_i}{\sum_{i=1}^{N} X_i}$
	Rolloff	$\text{Rolloff} = f_k, \quad \text{where} \quad \sum_{i=1}^{k} X_i \geq 0.85 \sum_{i=1}^{N} X_i$
	Flatness	$\text{Flatness} = \frac{\exp(\frac{1}{N}\sum_{i=1}^{N}\log(X_i))}{\frac{1}{N}\sum_{i=1}^{N} X_i}$
	Smoothness	$\text{Smooth} = \sum_{i=2}^{N-1}\lvert X_i - 2X_{i+1} + X_{i+2}\rvert$
	Irregularity-K	$\text{Irreg-K} = \sum_{i=1}^{N-1}\frac{\lvert X_i - X_{i+1}\rvert}{X_i}$
	Irregularity-J	$\text{Irreg-J} = \sum_{i=1}^{N-1}\lvert X_i - X_{i+1}\rvert$
	Peak valley ratio	$\text{PVR} = \frac{\max(X_1, X_2, \ldots, X_N)}{\min(X_1, X_2, \ldots, X_N)}$

$$GASF(i,j) = \cos\left(\arccos\left(\frac{x_i'}{\sqrt{\sum_{k=1}^{N} {x_k'}^2}}\right) + \arccos\left(\frac{x_j'}{\sqrt{\sum_{k=1}^{N} {x_k'}^2}}\right)\right) \tag{4}$$

Where x_i' and x_j' are the normalized data points of the time series, N is The total number of data points in the time series, and $\sum_{k=1}^{N} {x_k'}^2$ is the sum of the squares of the normalized data points, the square root of which is used for further polar coordinate transformation. GADF and GASF mappings present distinguishable textural and shape features in GAF images derived from different IMU devices, as shown in Figs. 2(a) and (b).

An MTF maps images by considering the Markov transition properties of time series, specifically by quantifying time series data into a certain number of discrete states and then calculating the probability matrix of state transitions [26]. This method effectively encodes the dynamic transition characteristics of time series into images. A visualization of an MTF is shown in Fig. 2(c). To construct an MTF from a one-dimensional signal, we begin by discretizing the processed sequence data into quantized state spaces. This involves dividing the value range of the time series into Q equal-frequency bins, with each data point x_t being assigned to a specific bin q_i on the basis of its value. Each bin q_i represents a discrete state within the Markov model. Then, a Q x Q transition matrix W as:



(a) GADF.

(b) GASF.

(c) MTF.

Fig. 2. Comparison among the GADF/GASF/MTF plots derived from four different devices.

$$W = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1Q} \\ w_{21} & w_{22} & \cdots & w_{2Q} \\ \vdots & \vdots & \ddots & \vdots \\ w_{Q1} & w_{Q2} & \cdots & w_{QQ} \end{bmatrix} \tag{5}$$

where each element w_ij denotes the relative frequency of

Fig. 3. Experimental environments and devices within vehicles.

transitioning from state q_i to state q_j.

The MTF M is subsequently defined as a T x T matrix, with each element m_st derived from the transition frequencies w_ij based on the bins q_i and q_j to which the time points x_s and x_t, respectively, belong.

$$M = \begin{bmatrix} w_{ij} \mid x_1 \in q_i, x_1 \in q_j & \cdots & w_{ij} \mid x_1 \in q_i, x_T \in q_j \\ w_{ij} \mid x_2 \in q_i, x_1 \in q_j & \cdots & w_{ij} \mid x_2 \in q_i, x_T \in q_j \\ \vdots & \ddots & \vdots \\ w_{ij} \mid x_T \in q_i, x_1 \in q_j & \cdots & w_{ij} \mid x_T \in q_i, x_T \in q_j \end{bmatrix} \tag{6}$$

IV. CNN-Based IMU Fingerprinting

Deep learning models that provide authentication via IMUs have shown excellent performance but often involve slow processing steps and high resource usage rates, limiting their practical V2X applications. In the previous section, we discussed the feature extraction capabilities of MTF, GADF, and GASF, which form the basis of our approach. To further address the challenges of computational overhead, we apply model compression techniques in this chapter to enhance the efficiency of our CNN architecture. Specifically, we apply model compression techniques to overcome these challenges and reduce the imposed computational overhead via a teacher-student network setup.

First, we use a complex CNN-based feature extractor as our teacher network for extracting features from IMU sensor data. The teacher model is crafted as a deep network with an input image size of 49x49x3, initiating with convolutional layers that have 8 channels and progressively increasing to 256 channels. Each convolutional layer includes batch normalization and a rectified linear unit (ReLU) activation layer and selectively utilizes max pooling to reduce the dimensions of features. In contrast, the student model listed in Table II starts with convolutional layers possessing 8 channels, eventually reaching 32 channels. Other models include a fully connected layer that outputs dimensions for predetermined classifications. To integrate the information derived from MTF, GADF, and GASF, each mapping is processed through separate branches in our CNN architecture. These branches handle the unique properties of each data representation, ensuring that the network can extract and learn from each's nuances. Post-feature extraction, the one-dimensional feature vectors from each branch are amalgamated. This is done through an addition layer that combines the vectors $\vec{v}_{MTF}, \vec{v}_{GADF}, \vec{v}_{GASF}$ into a single comprehensive feature vector: $\vec{v}_{(n)} = \vec{v}_{GADF} + \vec{v}_{GASF} + \vec{v}_{MTF}$. The combined feature vector then feeds into a subsequent layer which further processes the information for the final classification task, leveraging the strength of the compressed, yet richly informative feature set provided by the teacher-student architecture.

TABLE II
CNN STRUCTURE OF THE STUDENT MODEL

INPUT: 49×49×3			
Layer	**Activation**	**Filter/Stride/Padding**	**Dimension**
Conv_2D	BN+ReLu	3x3/1/1	49x49x8
Pool	-	2x2/2/-	24x24x8
Conv_2D	BN+ReLu	3x3/1/1	24x24x16
Pool	-	2x2/2/-	12x12x16
Conv_2D	ReLu	3x3/1/1	12x12x32
FC	-	-	20

The student network is guided through a loss function to extract task-relevant information from the input features as follows:

$$DL = \sum_{i=1}^{M} \alpha \frac{\left\| \vec{T}_i - \vec{S}_i \right\|^2}{\left\| \vec{T}_i \right\|^2} + (1-\alpha) \frac{\vec{T}_o \cdot \vec{S}_o}{\left\| \vec{T}_o \right\| \left\| \vec{S}_o \right\|} \tag{7}$$

The loss function for the neural network model is structured in two distinct parts. The first component calculates the squared Euclidean distance between each layer of the teacher and student networks, providing a quantitative measure of their differences. The second component evaluates the cosine similarity at the final layer, measuring the alignment between the directional outputs of the teacher and student models. The overall formula integrates these two divergent loss calculation methods, employing a parameter α to balance their contributions to the total loss. Given a teacher network T and a student network S, where M is the number of intermediate KD layers in the teacher network, the parameter α is used to balance the contributions of these networks to the total loss, and $\vec{T}_o$ and $\vec{S}_o$ are the output layers of the model. In this work, we set M to 0.3 and α to 0.7.

The complexity of a model is a critical metric for assessing its feasibility in practical applications. This complexity is measured through several key indicators: the number of FLOPs, the number of MADDs, and the overall memory

requirements. The number of FLOPs represents the number of floating-point operations required to execute a single inference process, serving as a reliable gauge of the inference efficiency of a model. Models with fewer FLOPs demand less computational resources and consequently take less time to process data. MADDs often assess the computational complexity of networks. Models with a high MADD counts typically incur greater computational expenses and may require more robust hardware to function efficiently. Memory requirements indicate the total amount of storage needed for a model to operate, including its parameters, activation values, and other transient data used during computations. Models with high memory demands may face challenges when operating on devices with limited resources. The performance achieved in terms of system overhead is shown in Sections V-H.

V. Experimental Evaluation

A. Experimental Setup

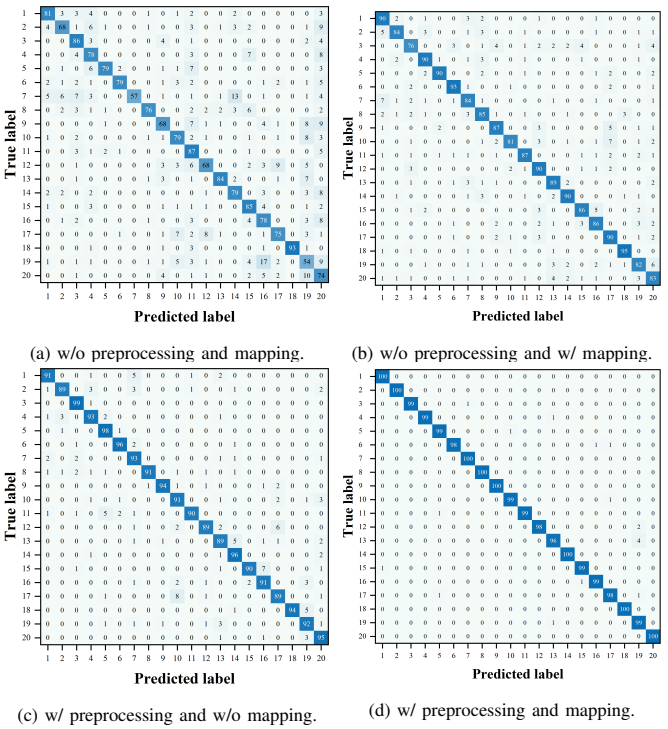
The scenes of vehicle and RSU authentication and the device for an on-board unit (OBU) to collect and send data from the IMU sensors are shown in Fig. 3. This device included three main components: 1) IMU sensors. A total of 20 IMU sensors were installed and distributed across 15 gasoline vehicles and 5 electric vehicles, and they were positioned near the centre of each vehicle. 2) An Arduino device connected to the IMU sensor to capture real-time data. 3) A Raspberry Pi 4b connected via USB to the Arduino to process the data, convert them into GASF/GADF/MTF images, and transmit them via WiFi to our assumed RSU for vehicle authentication.

The features five driving scenarios: starting the ignition (stationary), driving straight and turning at speeds of 0-30 km/h, driving straight at 30-60 km/h, and driving straight at 60-90 km/h. The experimental period spanned more than nine months (from September 2023 to May 2024) for exploring the effects of IMU sensor ageing and varying temperature conditions. To further test the robustness of the proposed approach under various driving conditions, such as uphill, downhill, miry, and bumpy roads, we collected sensor data in this external scene and conducted vehicle authentication directly on a workstation because our current experimental scene could not satisfy all these conditions. In each scene, approximately 200 data packets were collected per device.

The assumed RSU was a laboratory workstation with an Intel 14th Core i7KF CPU and an NVIDIA GeForce RTX 4070TiS GPU. All models were implemented using Python 3.7 and utilized the Adaptive Moment Estimation (Adam) optimizer with a learning rate of 10^{-4}. The mini-batch size was set to 32. For the dataset division, data was randomly allocated into a 70% training set and a 30% test set.

B. Performance Metrics

With respect to classification, we measured the performance of the proposed method with a confusion matrix, accuracy, and the F1 score. By definition, accuracy is calculated as $\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$, which represents the proportion of correctly predicted vehicle samples out of the total number of vehicle samples. The F1 score, defined as



(a) w/o preprocessing and mapping. (b) w/o preprocessing and w/ mapping.

(c) w/ preprocessing and w/o mapping. (d) w/ preprocessing and mapping.

Fig. 4. Confusion matrices produced under different experimental steps; overall accuracies: (a) 76.4%, (b) 86.9%, (c) 92.5%, and (d) 99.1%.

$\text{F1} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$, represents the harmonic mean of precision and recall.

For authentication purposes, the receiver operating characteristic (ROC) curve assesses the ability of a classifier system by plotting the true-positive rate (TPR) against the false-positive rate (FPR) under threshold settings. The area under the ROC curve (AUC) measures the area covered by the ROC curve, with values ranging from 0.5 to 1. The equal error rate (EER) is a common judgement metric in which the false-positive rate equals the missing rate (1-TPR). A higher AUC value and a lower EER indicate a model with stronger

C. Ablation Experimental Design

To evaluate the impact of each experimental step on the results, we designed corresponding ablation studies. 1) We bypassed the preprocessing and mapping stages for signals collected from the IMU sensors, adapting the network architecture to fit one-dimensional signals. 2) We omitted either the preprocessing step or the mapping step independently. 3) Last, we retained the preprocessing step and utilized a GADF for signal mapping.

The confusion matrix derived from the results shown in Fig. 4 shows that compared with bypass preprocessing and mapping or the lack of preprocessing, the accuracies were 76.4%, 86.9%, and 92.5%, respectively. The experiment that included preprocessing and mapping achieved a significantly higher accuracy of 99.1%. These findings demonstrate that our design approach can effectively attain enhanced vehicle classification performance. Furthermore, achieving higher classification accuracy not only advances our capabilities in vehicle identification but also mitigates potential threats from spoofed

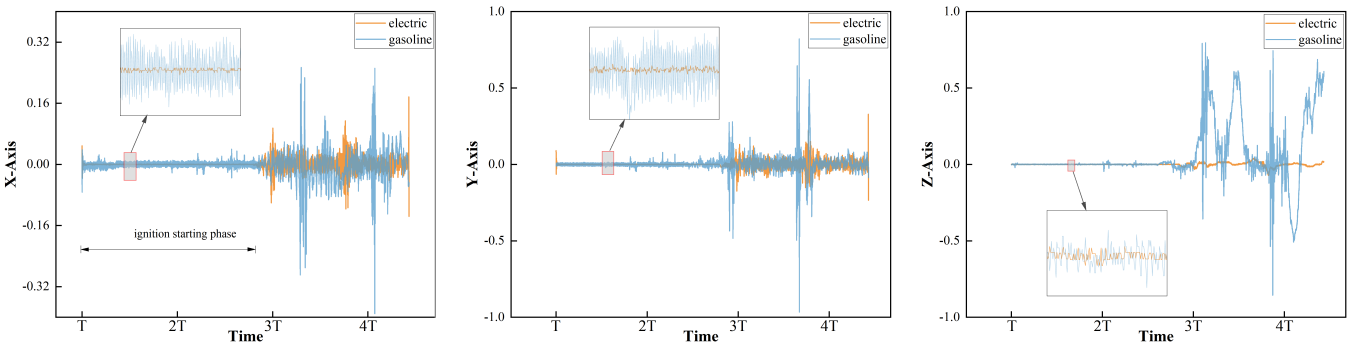
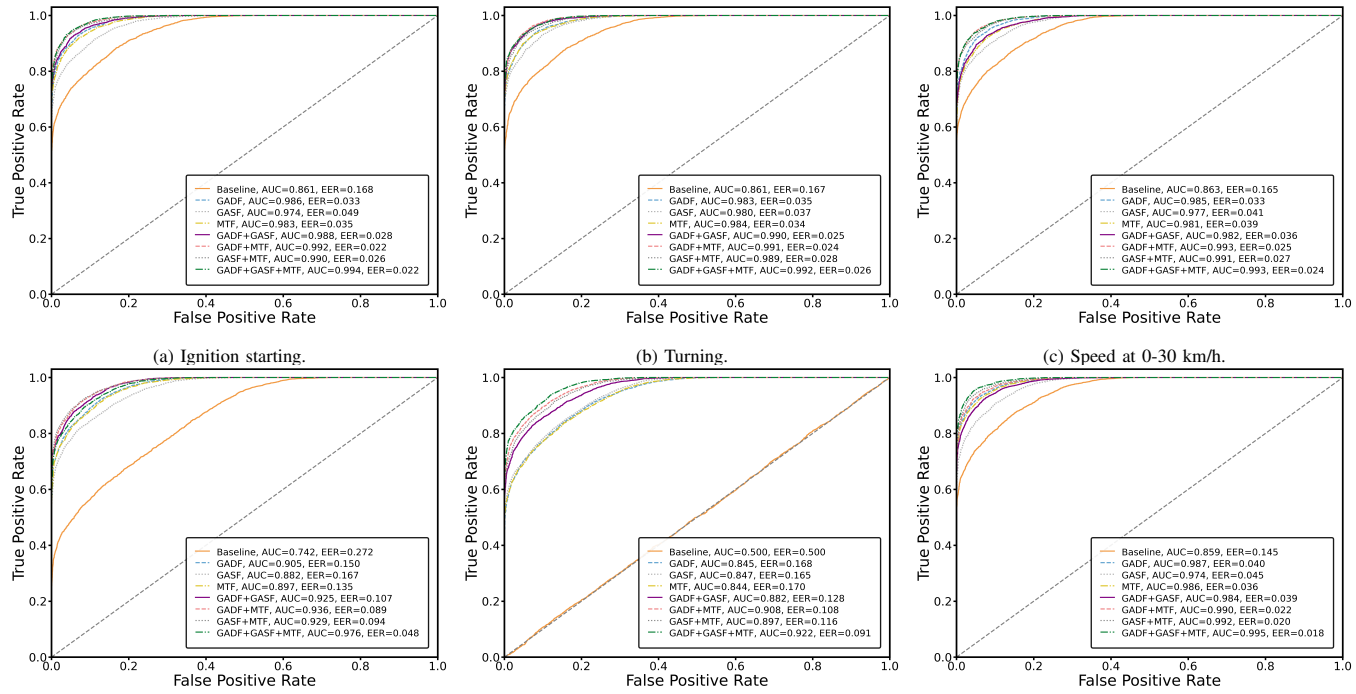


Fig. 5. Comparison between the vibrations of gasoline vehicles and electric vehicles.



(a) Ignition starting. (b) Turning. (c) Speed at 0-30 km/h.

(d) Speed at 30-60 km/h. (e) Speed at 60-90 km/h. (f) Retraining speed at 60-90 km/h.

Fig. 6. Comparison among the actual performances of three different schemes.

or unauthorized access, thereby strengthening the security of V2X communications through more precise authentication of vehicle identities

D. Performance Achieved Under Different Vehicle Conditions

This subsection evaluates the performance of our authentication scheme under various driving scenarios, including vehicle ignition, driving at different speeds, and turning.

In a real V2X environment, authentication and connection with the RSU should begin once the vehicle ignition occurs. Therefore, verifying whether successful authentication can occur before the vehicle is mobile is crucial. Notably, the ignition processes of gasoline vehicles cause significant vibrations due to their engines, whereas electric vehicles, powered by electric motors, do not induce such noticeable vibrations. We collected IMU data from 20 vehicles, including 15 gasoline and 5 electric vehicles. We closely monitored the three axes of the IMU sensor: X-axis (longitudinal vibration), Y-axis (lateral vibration), and Z-axis (vertical vibration). Figure 5, shows the variations in the X, Y and Z axis values of the accelerometer during the startup of different types of vehicles. Since electric vehicles produce significantly lower vibrations during startup, especially along the X and Y axes, this could lead to weakened signals in these two directions, further affecting the accuracy of the authentication process. Therefore, it is important to investigate whether reliable authentication can still be achieved for electric vehicles, despite these lower vibration levels.

As shown in Fig. 6(a), gasoline and electric vehicles yielded high accuracy, indicating that even the fewer vibrations induced by the startup processes of electric vehicles were sufficient for the system to capture IMU hardware fingerprints. This ensured secure authentication with the RSU as soon as the vehicles started up. Fig. 6(b) shows the authentication process during vehicle turning, which achieved average AUC rates above ninety-eight percent. Figs. 6(c)-(e) show that the average AUC of the system classification results decreased

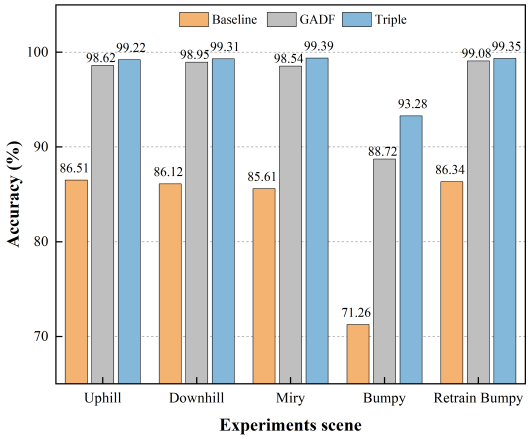


Fig. 7. Comparing among the performances achieved in different scenes.

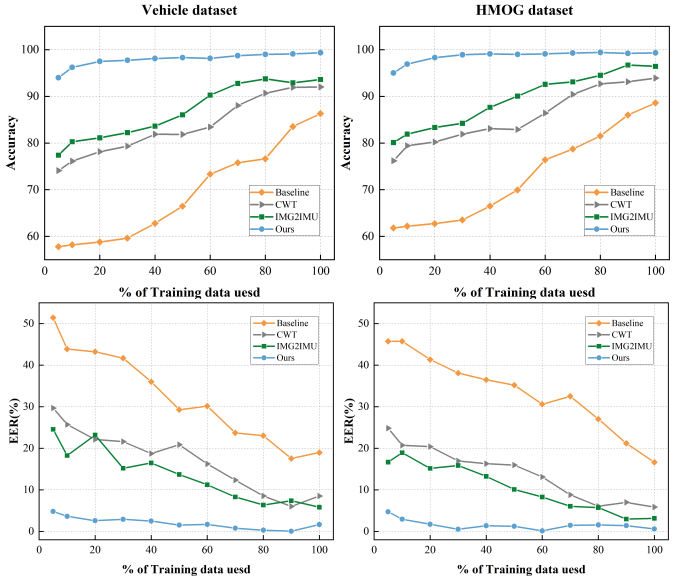


Fig. 8. Performance comparison between our approach and other schemes with different training data.

with increasing speed, with a notable reduction at approximately 90 km/h. The severe conditions show that the performance hierarchy is a triple>dual>single hierarchy. Moreover, for dual-channel setups, the MTF method outperformed the combination of the GADF and GASF, likely owing to the conceptual similarity between the GADF and GASF.

Although the best performance was achieved with the triple representation at approximately 90 km/h, the results were still unsatisfactory under real-world conditions. This suggests that high-speed driving is a major obstacle to perturbation performance. To address this issue, we utilized IMU sensor traces collected at high speeds of 90 km/h to retrain the model, as shown in Fig. 6(f), where the average AUC of the system returned to above ninety-eight percent. This change even impacted the vehicle classification accuracy achieved in the 30-60 km/h range, with significant improvements, Specifically, the AUC with the triple representation improved from 0.976 to 0.995, and the EER dropped to below 0.03. This demonstrates the robustness of our system across multiple driving scenarios with minimal additions to the training set.

E. Performance Achieved in an External Environment

This subsection examines how different external environments influence the IMU authentication performance achieved in vehicles. Experiments were conducted at a speed of approximately 30 km/h under four types of road conditions, namely, uphill, downhill, miry, and bumpy, and the results are displayed in Fig. 7.

At a constant speed, uphill, downhill, and miry roads did not significantly affect the vehicle classification results. However, when running on unusually bumpy roads, the model trained on smooth roads exhibited a slight decrease in recognition accuracy (approximately 88%). Accordingly, if traces from bumpy roads were combined with traces from flat sections for hybrid training, the system could still classify all vehicles with no less than 98% accuracy. This indicates that the road conditions themselves do not affect vehicle classification. Additionally, before the vehicles left the factory, their manufacturers performed whole-vehicle stress relief procedures on bumpy roads, and we could obtain the related IMU data in this phase before conducting training.

F. Performance Achieved with Various Training Samples Sizes

In the context of classification problems, deep learning approaches have demonstrated effectiveness with fewer training samples than those of traditional machine learning methods [27] [28]. To prove the reduced dependency of our approach on labelled data, we used both our dataset and a publicly available dataset, the HMOG dataset [29]. This dataset included data collected from 100 participants via 10 Samsung Galaxy S4 smartphones equipped with accelerometers, gyroscopes, and magnetometers. In our experiments, which focused on device classification verification, we amalgamated various poses for the individual devices and randomly extracted samples across 11 different proportions from 5% to 100% of the total dataset. We evaluated the performance of the proposed method against that of the existing technologies, including the following approaches.

- Baseline: Inputting raw signals into the GADF mapping process.
- Utilizing the continuous wavelet transform (CWT) [30] coupled with a CNN for feature extraction.
- IMG2IMU [31], a spectrogram-based feature extraction method for enhancing IMU sensor data, which achieved good results.

Comparisons between our approach and advanced technologies were performed based on accuracy and the ERR, as shown in Fig. 8. Even with 5% of the training data (i.e., 100 samples), our method yielded significantly higher accuracy and lower EER values than spectrogram and CWT. This clear advantage highlights the effectiveness of our preprocessing technique, particularly when within equivalent training volumes. Similar results were observed on the HMOG public dataset, which includes rich data categories and only ten device types;

TABLE III
PERFORMANCE ACHIEVED IN AGEING AND THERMAL TESTS CONDUCTED OVER 9 MONTHS

Condition		Accuracy	F1	EER
Date	Temperature			
2023/9/20	27°C	99.12%	99.13%	2.23%
2023/10/20	18°C	99.17%	99.17%	1.96%
2023/11/20	15°C	98.93%	98.95%	3.47%
2023/12/20	-7°C	99.01%	99.01%	2.84%
2024/01/26	5°C	99.25%	99.23%	2.33%
2024/02/20	-1°C	99.19%	99.20%	1.84%
2024/03/20	15°C	99.17%	99.08%	2.05%
2024/04/20	21°C	98.89%	98.99%	3.61%
2024/05/22	30°C	99.15%	99.16%	1.96%

however, our method achieved an accuracy of 99.34%. These experiments indicate the robust performance of our approach with limited training samples.

G. Performance Achieved with Ageing and Temperature

Semiconductor manufacturers perform ageing and thermal testing on chips during production, which includes Monte Carlo and process corner simulations, as well as testing in thermal chambers, confirming the impacts of time and temperature factors on chip performance. Similarly, we tested various timings and temperature ranges to mimic real-world conditions. For the ageing tests, the sensor chips were powered for approximately 10 hours daily. The data in Table III show how the accuracy shifted with time and temperature variations. The performance of our system remained stable over the nine months despite temperature and duration changes. These results are promising, showing that during prolonged operational periods, the IMU fingerprint of a vehicle remains stable, yielding consistently high accuracy.

H. System Overhead

To evaluate the deployment overhead of the proposed network compression-based continuous authentication system, we analyzed it from six perspectives: accuracy, EER, time cost, memory, FLOPs, and MADDs. As shown in Fig. 9, we improved the accuracy and EER of the student model, closely matching the teacher model by distilling knowledge from the teacher model. However, compared with the memory requirements, time cost, FLOPs, and MADDs of the teacher network, those of the student network were 83%, 51%, 83.5% and 82.7% lower, which are 1.769M, 6.91 seconds, 3.98M and 2.338M, respectively. These results validate the effectiveness of our network compression method. Student networks can replace teacher networks when deployed on terminals with limited computing resources. Additionally, As in Table V, we compared our approach with other schemes, and our resource consumption is one of the lowest.

I. Security Analysis

Based on the threat model in Section 2, we designed real scenario to assess the resistance of our proposed continuous

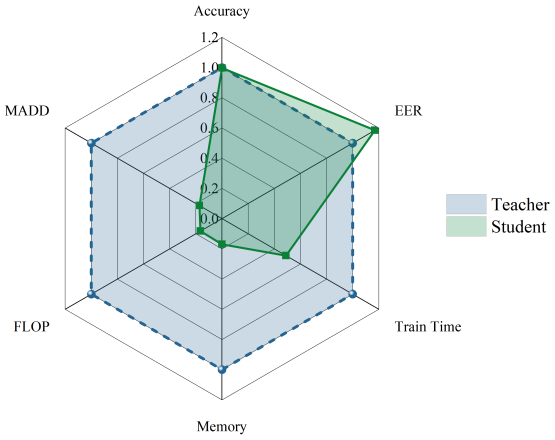


Fig. 9. Comparison between the overheads of the teacher and student models.

TABLE IV
SEVEN TEST DEVICES CORRECTLY MATCHED THE RESULTS OF THEIR RESPECTIVE CATEGORIES

Vehicles	A	B	C	D	E	F	G
A	99.99	0.49	0.00	0.28	0.00	0.02	1.36
C	0.00	0.00	99.96	0.02	0.00	0.79	0.04
E	0.00	0.03	0.04	0.00	99.23	0.00	0.00
Eves	0.01	99.21	0.00	99.70	0.77	99.19	98.60

authentication system to attacks, this scenario involving six cars, three of which (A, C, E) were legitimate vehicles (Alices) and three of which (B, D, F) were illegal vehicles (Eves). These vehicles requested communication validation from the RSU, which played the role of Bob. Vehicles A and B, and C and D, were grouped, with the illegal vehicles following the legitimate vehicles to closely observe their driving habits. For a more extreme test, we placed an unauthorized IMU (G) inside a legitimate vehicle E to capture its driving environment and initiate impersonation attacks. Illegal vehicle F sends its own IMU data to the Roadside Unit (RSU) without knowledge of the legitimate vehicle's driving patterns.

.

The results in Table IV show that all seven test devices were correctly associated with their respective categories, achieving satisfactory outcomes; in particular, the IMU of the attacker (G) placed inside a legitimate vehicle was also accurately identified. This finding demonstrates that even if an attacker mimics the driving style of the victim vehicle under similar external conditions, the unique hardware fingerprint of its IMU sensor remains distinct from those of legitimate vehicles. Notably, hardware fingerprints are determined by uncontrollable random factors during the manufacturing process, making it practically impossible for attackers to replicate these characteristics accurately, even if they manage to simulate most of the signal. Thus, the Spoofing attack is also deemed unsuccessful. They confirm that our designed system can effectively defend against random, Spoofing and impersonation attacks.

TABLE V
COMPLEXITY ANALYSIS WITH DIFFERENT APPROACH

Baseline	Memory (MB)	FLOP (M)	MADD (M)
Mekki [23]	8.63	153.15	126.20
Wang [7]	0.658	0.796	0.637
Xun [6]	4.661	13.27	10.25
Fereidooni [16]	17.33	264.72	201.89
Jiang [32]	11.56	193.34	178.64
Weshabi [33]	2.417	58.291	48.77
Arshad [30]	6.422	90.67	72.19
Yoon [31]	8.26	130.39	110.42
Reddy [34]	19.72	342.89	326.45
Ours	1.769	3.98	2.338

TABLE VI
COMPARISON AMONG DIFFERENT DL MODELS

Deep learning models		**Accuracy(%)**	**Training time(s)**
Categories	Baseline		
LSTM	LSTM [35]	90.91	20.34
MLP	CycleMLP [36]	89.66	144.07
Transformer	Swin-Trans [37]	**99.31**	155.43
	Crossvit [38]	**99.15**	178.26
BLS	ConvBLS [39]	99.01	78.51
	ABL [40]	92.55	**4.65**
CNN	ResNet-18 [32]	**99.19**	40.33
	VGG16 [34]	96.41	56.29
	MobileNet-V2 [33]	98.67	13.62
	Ours	99.10	**6.19**

VI. GENERALIZABILITY TO OTHER DL MODELS

To assess the generalizability of the proposed preprocessing and representation methods across other deep learning models, we considered nine baseline methods belonging to five categories. The models included LSTM networks, multilayer perceptrons (MLPs), transformers, broad learning systems (BLSs) and CNNs. We adjusted the input to suit each baseline, conducted ten training sessions for each method and calculated the average performance.

The classification results obtained on a dataset consisting of 20 vehicles are shown in Table VI; the models using the CNN and transformer methods performed almost perfectly, whereas the models using the MLP and LSTM methods achieved mediocre performance. The adaptive breadth learning (ABL) model completed the training process in less time but with intermediate accuracy, possibly due to parameterization challenges. Notably, after the proposed signal processing preprocessing and representation methods were applied, most of the models required only short training periods to achieve satisfactory performance. Our model achieved encouraging results in only 5.46 seconds and ranked second. The results indicate that the proposed lightweight approach achieved a better balance between training time and accuracy. This makes it highly suitable for real-world applications where rapid training and model adjustment processes are crucial. In the V2X communication environment, the ability of models to adapt in real time is crucial. Most models rely on offline training; however, offline data often fails to cover all variables encountered on-site, such as extreme road and equipment ageing, leading to decreased performance in practical applications. In contrast, the lightweight design proposed in this study makes it possible to adapt instantly to environmental changes on-site.

VII. LIMITATIONS AND FUTURE WORK

In this section, we discuss some of the limitations and corresponding future work emerging from the above experiments, as follows:

Extreme Scenarios: In Sections V-D and F, while the system performs well under typical driving conditions, challenges persist in more extreme scenarios, such as high-speed driving and driving on bumpy roads. Under these conditions, the accuracy remains insufficient, which necessitates the collection of training data that are specific to these scenarios. To address the challenge, future research may focus on developing more robust feature extraction methods that are less sensitive to changes in extreme scenarios, using adaptive filtering techniques could help reduce the impact of harsh driving conditions on system performance.

Extensive Testing: Our paper demonstrates the potential of IMU-based device fingerprinting and introduces a two-stage feature mapping method. However, due to current experimental limitations, we are unable to validate the system across thousands of vehicles. Nevertheless, with the continued development of device fingerprinting technology, IMU-based continuous authentication for vehicles shows considerable promise. In future work, we aim to test the system on a broader range of vehicles to further validate its scalability and effectiveness.

Resource Overhead: Although we apply a network compression model to address resource-limited terminals, this approach still introduces some overhead due to the compression process itself. Future work should focus on developing models that eliminate the need for compression, aiming for even lower overhead levels.

VIII. CONCLUSION

In this paper, we proposed a lightweight continuous authentication system based on IMU fingerprints, which significantly enhanced the safety of V2X communication. By combining the fingerprints of IMU devices with our proposed two-stage feature mapping method, the system achieved an average accuracy and F1 score of over 99.10% in nine months of real-world driving tests. Additionally, our method was validated using the HMOG dataset, and it similarly achieved an impressive average accuracy of 99.34%. In the resource-constrained environments encountered in V2X environments, our model compression approach dramatically reduces the demand for computing resources, and the system to complete model training in 6.91s while maintaining robustness and high accuracy. This rapid training capability is crucial for onsite training and updates, enhancing the system's practical deployment. A series

of extensive experiments demonstrated the effectiveness and adaptability of our approach across various driving phases and road conditions. Furthermore, through evaluations conducted across various deep learning models, this study shows the generalizability of the proposed preprocessing and data representation methods.

ACKNOWLEDGMENTS

This work was supported in part by the Major Science and Technology Projects in Yunnan Province (202202AD080013).

REFERENCES

[1] B. Padmaja, C. V. K. N. S. N. Moorthy, N. Venkateswarulu, and M. M. Bala, "Exploration of issues, challenges and latest developments in autonomous cars," *journal of big data*, vol. 10, no. 1, 2023.

[2] M. Khodaei, H. Jin, and P. Papadimitratos, "Secmace: Scalable and robust identity and credential management infrastructure in vehicular communication systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 5, pp. 1430–1444, 2018.

[3] S. Lv and Y. Liu, "Plva: Privacy-preserving and lightweight v2i authentication protocol," *IEEE Transactions on Intelligent Transportation Systems*, vol. PP, no. 99, pp. 1–7, 2021.

[4] A. Alnasser, H. Sun, and J. Jiang, "Cyber security challenges and solutions for v2x communications: A survey," *Computer Networks*, vol. 151, pp. 52–67, 2019.

[5] B. Taha, S. N. A. Seha, D. Y. Hwang, and D. Hatzinakos, "Eyedrive: A deep learning model for continuous driver authentication," *IEEE Journal of Selected Topics in Signal Processing*, vol. 17, no. 3, pp. 637–647, 2023.

[6] Y. Xun, J. Liu, N. Kato, Y. Fang, and Y. Zhang, "Automobile driver fingerprinting: A new machine learning based authentication scheme," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 2, pp. 1417–1426, 2020.

[7] Y. Wang, T. Zhao, F. Tahmasbi, J. Cheng, Y. Chen, and J. Yu, "Driver identification leveraging single-turn behaviors via mobile devices," in *2020 29th International Conference on Computer Communications and Networks (ICCCN)*, pp. 1–9, IEEE, 2020.

[8] T. Zhu, Z. Qu, H. Xu, J. Zhang, Z. Shao, Y. Chen, S. Prabhakar, and J. Yang, "Riskcog: Unobtrusive real-time user authentication on mobile devices in the wild," *IEEE Transactions on Mobile Computing*, pp. 1–1, 2020.

[9] S. H. Sánchez, R. F. Pozo, and L. A. H. Gómez, "Driver identification and verification from smartphone accelerometers using deep neural networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 1, pp. 97–109, 2020.

[10] P. Yan, J. Jiang, H. Tan, Q. Zheng, and J. Liu, "High precision time synchronization strategy for low-cost embedded gnss/mems-imu integrated navigation module," *IEEE Transactions on Intelligent Transportation Systems*, 2024.

[11] Z. Wu, D. Yuan, F. Zhang, and M. Yao, "Low-cost attitude estimation using gps/imu fusion aided by land vehicle model constraints and gravity-based angles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 8, pp. 13386–13402, 2021.

[12] S. Dey, N. Roy, W. Xu, R. R. Choudhury, and S. Nelakuditi, "Accelprint: Imperfections of accelerometers make smartphones trackable," in *Network and Distributed System Security Symposium*, 2014.

[13] S. Lee and D. H. Lee, "From attack to identification: Mems sensor fingerprinting using acoustic signals," *IEEE Internet of Things Journal*, vol. 10, no. 6, pp. 5447–5460, 2022.

[14] Y. Shen, F. Lin, C. Wang, T. Liu, Z. Ba, L. Lu, W. Xu, and K. Ren, "Motoprint: Reconfigurable vibration motor fingerprint via homologous signals learning," *IEEE Transactions on Dependable and Secure Computing*, vol. 21, no. 1, pp. 372–387, 2023.

[15] S. Abdolinezhad, V. Stavrov, L. Zimmermann, and A. Sikora, "Piezoresistive-based physical unclonable function," *IEEE Sensors Journal*, 2024.

[16] H. Fereidooni, J. König, P. Rieger, M. Chilese, B. Gökbakan, M. Finke, A. Dmitrienko, and A.-R. Sadeghi, "Authentisense: A scalable behavioral biometrics authentication scheme using few-shot learning for mobile platforms," *arXiv preprint arXiv:2302.02740*, 2023.

[17] S. Kussl, K. S. Omberg, and O.-I. Lekang, "Advancing vehicle classification: A novel framework for type, model, and fuel identification using nonvisual sensor systems for seamless data sharing," *IEEE Sensors Journal*, vol. 23, no. 17, pp. 19390–19397, 2023.

[18] N. Marwan, M. Carmen Romano, M. Thiel, and J. Kurths, "Recurrence plots for the analysis of complex systems," *Physics Reports*, vol. 438, 2007.

[19] N. Hatami, Y. Gavet, and J. Debayle, "Classification of time-series images using deep convolutional neural networks," 2017.

[20] Baldini, G., Gentile, C., Giuliani, R., and Steri, "Comparison of techniques for radiometric identification based on deep convolutional neural networks," *Electronics Letters*, 2019.

[21] X. Qi, A. Hu, and Z. Zhang, "Data-and-channel-independent radio frequency fingerprint extraction for lte-v2x," *IEEE Transactions on Cognitive Communications and Networking*, 2024.

[22] P. Peralta-Braz, M. M. Alamdari, E. Atroshchenko, and M. Hassan, "On the joint optimization of energy harvesting and sensing of piezoelectric energy harvesters: case study of a cable-stayed bridge," *IEEE Transactions on Intelligent Transportation Systems*, 2023.

[23] A. E. Mekki, A. Bouhoute, and I. Berrada, "Improving driver identification for the next-generation of in-vehicle software systems," *IEEE Transactions on Vehicular Technology*, no. 68-8, 2019.

[24] Y. Song, F. Jiang, S. W. A. Shah, and R. Doss, "Multi-factor continuous authentication of drivers leveraging smartphone," in *2023 IEEE Smart World Congress (SWC)*.

[25] L. Ye, S. Hu, T. Yan, and Y. Xie, "Gaf representation of millimeter wave drone rcs and drone classification method based on deep fusion network using resnet," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 59, no. 1, pp. 336–346, 2022.

[26] Z. Wang and T. Oates, "Encoding time series as images for visual inspection and classification using tiled convolutional neural networks," in *Workshops at the twenty-ninth AAAI conference on artificial intelligence*, 2015.

[27] M. Hu, K. Zhang, R. You, and B. Tu, "Multisensor-based continuous authentication of smartphone users with two-stage feature extraction," *IEEE Internet of Things Journal*, vol. 10, no. 6, pp. 4708–4724, 2022.

[28] A. Berdich, P. Iosif, C. Burlacu, A. Anistoroaei, and B. Groza, "Fingerprinting smartphone accelerometers with traditional classifiers and deep learning networks," in *2023 IEEE 17th International Symposium on Applied Computational Intelligence and Informatics (SACI)*, pp. 000039–000044, IEEE, 2023.

[29] Z. Sitová, J. Šeděnka, Q. Yang, G. Peng, G. Zhou, P. Gasti, and K. S. Balagani, "Hmog: New behavioral biometric features for continuous authentication of smartphone users," *IEEE Transactions on Information Forensics Security*, vol. 11, no. 5, pp. 877–892, 2016.

[30] M. Z. Arshad, D. Jung, M. Park, H. Shin, J. Kim, and K. R. Mun, "Gait-based frailty assessment using image representation of imu signals and deep cnn," 2021.

[31] H. Yoon, H. Cha, C. H. Nguyen, T. Gong, and S.-J. Lee, "Img2imu: Applying knowledge from large-scale images to imu applications via contrastive learning," *arXiv preprint arXiv:2209.00945*, 2022.

[32] A. Jiang and J. Ye, "Selfvis: Self-supervised learning for human activity recognition based on area charts," *IEEE Transactions on Emerging Topics in Computing*, vol. PP.

[33] F. N. Al-Wesabi, A. A. Albraikan, A. M. Hilal, A. A. Al-Shargabi, S. Alhazbi, M. Al Duhayyim, M. Rizwanullah, and M. A. Hamza, "Design of optimal deep learning based human activity recognition on sensor enabled internet of things environment," *IEEE Access*, vol. 9, pp. 143988–143996, 2021.

[34] B. Reddy, "Multi-feature embedding and deep classification for elderly activity recognition," in *2023 International Conference on Data Science and Network Security (ICDSNS)*, pp. 1–7, IEEE, 2023.

[35] H. Liu, X. Y. Li, L. Zhang, Y. Xie, Z. Wu, Q. Dai, G. Chen, and C. Wan, "Finding the stars in the fireworks: Deep understanding of motion sensor fingerprint," pp. 126–134, 2018.

[36] S. Chen, E. Xie, C. Ge, R. Chen, D. Liang, and P. Luo, "Cyclemlp: A mlp-like architecture for dense visual predictions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023.

[37] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, and B. Guo, "Swin transformer: Hierarchical vision transformer using shifted windows," in *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 10012–10022, 2021.

[38] C.-F. R. Chen, Q. Fan, and R. Panda, "Crossvit: Cross-attention multi-scale vision transformer for image classification," in *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 357–366, 2021.

[39] C. Lei, C. Chen, J. Guo, and T. Zhang, "Convbls: An effective and efficient incremental convolutional broad learning system for image classification," *arXiv preprint arXiv:2304.00219*, 2023.
[40] Z. Xu, G. Han, L. Liu, H. Zhu, and J. Peng, "A lightweight specific emitter identification model for iiot devices based on adaptive broad learning," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 5, pp. 7066–7075, 2022.

Bei Gong (Member, IEEE) received a B.S. degree from Shandong University in 2005 and a Ph.D. degree from the Beijing University of Technology in 2012. In the past five years, he has published more than 30 papers in first-class SCI/EI and other international famous journals and top international conferences in relevant research fields. His research interests include trusted computing, Internet of Things security, Industrial Internet of Things, and privacy data protection.

Zhe Li pursuing his Ph.D degree in Beijing University of Technology, Beijing China. He received his M.S. degree from the Chemnitz University of Technology, Chemnitz, Germany. His research interests include Internet of Vehicles security, identity authentication, and cryptography.

Mowei Gong pursuing her Ph.D degree in Beijing University of Technology, Beijing China. She received her M.S. degree from the Beijing University of Technology, Beijing, China. Her research interests include Industrial Internet of Things security, privacy data protection, and cryptography.

Haotian Zhu pursuing his Ph.D degree in Beijing University of Technology, Beijing China. He received his M.S. degree from the Beijing University of Technology, Beijing, China. His research interests include trust management, Internet of Things security, and cryptography.

Weizhi Meng is a Full Professor in the School of Computing and Communications, Lancaster University, United Kingdom, and an adjunct faculty in the Department of Applied Mathematics and Computer Science, Technical University of Denmark, Denmark. He received the IEEE ComSoc Best Young Researcher Award for Europe, Middle East, Africa Region (EMEA) in 2020. His primary research interests are intersections among cyber security, artificial intelligence and blockchain technology, such as intrusion detection, IoT security, biometric authentication, and blockchain. He serves as associate editors / editorial board members for many reputed journals such as IEEE TDSC and IEEE TIFS, as well as chair for various international conferences such as ACM CCS'23 and ESORICS'22. He is a senior member of IEEE.

Chong Guo received his M.S. degree from Beijing University of Technology, Beijing, China, in 2021. He is working toward the Ph.D degree in computer science and technology with Beijing University of Technology. His research interests including network and information security, Internet of things security and cryptography. (Email: chongguo@emails.bjut.edu.cn)