Contents lists available at ScienceDirect

Pervasive and Mobile Computing

journal homepage: www.elsevier.com/locate/pmc

Passive Monitoring of Dangerous Driving Behaviors Using mmWave Radar

Argha Sen *, Avijit Mandal, Prasenjit Karmakar, Anirban Das, Sandip Chakraborty

Indian Institute of Technology, Kharagpur, India

ARTICLE INFO

ABSTRACT

Detecting risky driving has been a significant area of focus in recent years. Nonetheless, devising a practical, effective, and unobtrusive solution remains a complex challenge. Presently available technologies predominantly rely on visual cues or physical proximity, complicating the sensing. With this incentive, we explore the possibility of utilizing mmWave radars exclusively to identify dangerous driving behaviors. Initially, we scrutinize the attributes of unsafe driving and pinpoint distinct patterns in range-doppler readings brought about by nine common risky driving manoeuvres. Subsequently, we create an innovative Fused-CNN model that identifies instances of hazardous driving amidst regular driving and categorizes nine distinct types of dangerous driving actions. After conducting thorough experiments involving seven volunteers driving in real-world settings, we note that our system accurately distinguishes risky driving actions with an average precision of approximately 97% with a deviation of $\pm 2\%$. To underscore the significance of our approach, we also compare it against established state-of-the-art methods.

1. Introduction

The World Health Organization (WHO) has reported that a concerning 1.2 million individuals lose their lives annually on a global scale due to road accidents [1]; notably, approximately 45% of these accidents are attributed to dangerous driving. It is worth highlighting that substantial technological advancements have been made to monitor hazardous driving behaviors actively. Both commercial efforts like *Advanced Driver Assistance Systems (ADAS)* [2] and a plethora of research studies [3,4] contribute to a well-explored domain. For example, contemporary trends emphasize the significance of computer vision [5] and wearable devices mounted on the body [6] for monitoring driving conduct. However, privacy concerns, especially for public and shared vehicles, pose challenges for vision-based techniques. Additionally, the accuracy of detection is influenced by various factors such as lighting conditions and camera orientation [7]. Similarly, wearable-based strategies struggle with generalization as behavioral patterns from one age group might not seamlessly apply to another [8]. Furthermore, ensuring drivers consistently use necessary wearable devices is not straightforward. This leads us to a fundamental query: *How can we create a comprehensive and unobtrusive system to monitor dangerous driving, one that the community can widely embrace? Additionally, can an alternative approach effectively tackle this issue?* We embark on addressing these intriguing questions.

Recently, a notable shift has arisen towards adopting 5G technology, which is built upon the foundation of millimeter wave (mmWave) communication [9]. With an increasing influx of devices into this ecosystem, mmWave hardware is being integrated into a wide array of devices, establishing this technology as a widespread platform. This observation has sparked a crucial

---

* Corresponding author.

E-mail addresses: arghasen10@gmail.com (A. Sen), avijitmandal2001@gmail.com (A. Mandal), prasenjitkarmakar52282@gmail.com (P. Karmakar), anirbanfuture@gmail.com (A. Das), sandip.chakraborty@gmail.com (S. Chakraborty).

URL: http://www.arghasen10.github.io (A. Sen).

Fig. 1. A use-case scenario of *mmDrive* .

inquiry: *Is it feasible to utilize the properties of mmWave to capture a driver's dangerous driving behaviors?* A mmWave radar [10] can adeptly measure various parameters like the distance and velocity of objects. This concept has been applied to address a diverse range of challenges, including recognizing human activities [11], interpreting gestures [12], detecting vital signs [13], and even reconstructing voice [14], involving estimations of position and movement. Based on this premise, we introduce *mmDrive , a solution that employs a Frequency-modulated continuous-wave (FMCW) mmWave radar to monitor hazardous driving behaviors from the driver's perspective*. Upon detecting instances of distracted driving, *mmDrive* can promptly take actions such as alerting the driver or notifying nearby vehicles with relevant messages. As demonstrated by the use-case scenario in Fig. 1, *mmDrive* could play a crucial role in saving lives both in and outside a vehicle. As soon as a case of distracted driving is detected, *mmDrive* can be used to take immediate action, such as warning the driver or notifying nearby vehicles with appropriate messages.

Gains Compared to Current Approaches: We additionally note that employing mmWave sensing for the surveillance of hazardous driving holds several benefits when juxtaposed with prevailing cutting-edge methods. These advantages encompass the subsequent points. (i) Direct monitoring of a driver's actions is enabled by mmWave sensing, obviating the need for indirect inferences like vehicle conditions and kinematics. (ii) Unlike cameras, mmWave sensing maintains a commendable level of privacy as it abstains from capturing visual aspects of the surroundings [15]. (iii) Unlike wearables, mmWave sensing operates unobtrusively without necessitating any device to be worn or carried, permitting passive measurement without impeding the driver's natural movements and routines. (iv) Micro-movements such as *yawning,* crucial for discerning a drowsy state in a driver, can also be detected using mmWave sensing [16]. (v) The ability of mmWave to penetrate objects that could potentially obstruct light, like clothing [17], ensures its utility even when the driver is wearing a face mask. Nevertheless, alongside these merits, the practical implementation of *mmDrive* must surmount a set of challenges.

Challenges: When considering the utilization of a mmWave radar for overseeing activities within a moving vehicle, the road ahead is not devoid of complexities. Several challenges demand resolution to establish an effective system, as delineated below. *Initially*, the internal setting of a car is inherently noisy. Myriad movements stemming from objects within the vehicle, coupled with road and traffic conditions, can directly influence mmWave signals. *Secondarily*, a car might accommodate multiple occupants. However, exclusively, the driver's motions hold relevance for monitoring perilous driving circumstances, necessitating the separation of solely the driver's actions. *Tertiary*, the signatures captured by the FMCW radar arising from abrupt jolts, like encounters with road irregularities and potholes, exhibit substantial pattern fluctuations. As these patterns do not bear relevant information, their exclusion becomes imperative.

Our Contributions: In light of the aforementioned hurdles, this manuscript introduces an end-to-end system that employs mmWave sensing for unobtrusive, real-time monitoring of risky driving behaviors. In distinction from existing endeavors [18] that employ acoustic sonar to detect a limited range of basic driving behaviors exclusively, we encompass a spectrum of 9 intricate macro and micro-level actions that could potentially lead to fatal outcomes while driving. The contributions presented in this document unfold as follows.

(1) Definition of Dangerous Driving and Identification of Detectable Signatures. We establish definitions for *three* actions often linked to driver fatigue or drowsiness (specifically, *nodding, yawning, steering anomaly*), and *six* actions indicative of driver distraction (namely, *drinking/eating, looking back, picking up/dropping items, forward fetching from dashboard, talking on a mobile phone, engaging with passengers*). Collectively, these nine actions harbor potential danger and warrant avoidance while operating a vehicle. Utilizing a solitary Commercial Off-The-Shelf (COTS) FMCW mmWave radar, we assess an array of signal attributes encompassing *range-doppler*, *range profile*, and *noise profile* to distinguish these behaviors from typical driving actions. We underscore the necessity of amalgamating spatial and temporal fluctuations in these attributes to discern dangerous driving behaviors accurately.

(2) Creation of an end-to-end process for classifying driving behaviors in the presence of noise. Empirical tests in real-world scenarios under both rural and urban road conditions unveil that mmWave data contends with signal interference arising from adverse road conditions like bumps and potholes. To mitigate such disturbances, we exploit data from IMU sensors. Additionally, we introduce an innovative Fused-CNN classifier to ascertain hazardous driving versus regular driving. Notably, *mmDrive* not only discriminates between dangerous and normal driving but also distinguishes among the aforementioned nine distinct instances of dangerous driving conduct. To curtail computational demands and energy consumption, *mmDrive* exclusively classifies the potentially perilous driving actions upon detecting a case of hazardous driving.

(3) Implementation and Assessment with Real-world Driving Data. We put our proposed driver behavior model based on the Fused-CNN into action and compared it with a random forest and a VGG-16-based baseline. Furthermore, we compared with an acoustic-based method [18] to underscore the enhanced precision of mmWave sensing. To replicate our outcomes, we have made our implementation and the real-world driving dataset publicly accessible at: https://github.com/arghasen10/mmdrive.git. Employing a meticulous device deployment and field trials, we amassed 60 hours of driving data (comprising mmWave readings, IMU inputs, and dashcam footage for reference), approximately three times the previously collected version, from 7 individuals operating three distinct vehicles — three sedans, one SUV, and one Minivan, under both rural and urban road conditions.

In summary, we have significantly improved upon the previous publication by implementing key enhancements in this version [19]. We have expanded our driving dataset by approximately threefold, covering diverse real-world scenarios, including rural, suburban, and urban settings. This contrasts with the earlier dataset [19] that was limited to campus road dynamics, thereby enhancing the reliability of our findings. The modeling pipeline has also evolved due to the enlarged dataset and consideration of diversity in the driving signatures. We now divide the data based on timestamps, allocating 70% for training and the rest for testing, rectifying earlier data leakage issues and providing a more accurate representation of real-world driving. Our evaluation now includes comprehensive insights into performance on different road types and car models, the impact of IMU-based denoising, as well as real-time performance latency.

The rest of the paper is organized as follows. In Section 2, an examination of related works provides a context for the novelty of the proposed approach. Moving forward, Section 3 presents a thorough introduction to mmWave sensing and outlines the pilot experiments undertaken as part of this research. Drawing insights from these experiments, Section 4 elaborates on the chosen methodology and the underlying system architecture. The subsequent Section 5 delves into the implementation of *mmDrive* on COTS mmWave radar, followed by a detailed analysis of the module's performance compared to the state-of-the-art baselines. Ultimately, Section 6 encapsulates the paper with a conclusion summarizing the findings and hints at the potential future direction of work.

2. Related work

The existing literature concerning the surveillance of hazardous driving behaviors has expanded into various avenues. In this context, we provide an overview of the most notable works in each of these avenues.

Approaches Involving Orientation and Abnormality: These earlier methods [20,21], focus on evaluating the vehicle's orientation using a range of parameters, including position, velocity, acceleration, and more. Subsequently, these parameters are utilized to *indirectly* observe risky driving behaviors. The underlying premise is that deviations from normal driving patterns manifest as irregularities in these parameters, leading to the inference of hazardous driving actions. However, it is important to note that various factors outside of the driver's control, such as weather conditions and traffic congestion, can also result in irregular driving patterns. These patterns may not necessarily indicate dangerous driving behaviors.

Approaches Centered on Vision: Vision-centered methodologies [22,23] harness the capabilities of computer vision technology, utilizing various visual sources including RGB cameras [24], thermal imaging [25], and Infrared (IR) cameras [26]. These approaches are designed to monitor the driver's actions directly. Images captured within the surrounding environment undergo processing to extract motion details, encompassing facial cues like eye movements, talking, and yawning, as well as movements of other body components [27] like head and hand gestures. Such behaviors serve as indicators of inattentive or perilous driving conduct. However, it is worth noting that these vision-based methodologies compromise privacy [28], thereby limiting their viability, particularly for public vehicles, as a practical solution.

Approaches Based on Wearables: Recognizing the wide-reaching presence and efficacy of wearable gadgets, researchers have harnessed these devices to detect body kinematics and other distinctive bodily markers. Variations within IMU (Inertial Measurement Unit) data, EEG (ElectroEncephaloGram) signals [29], heart rate, and more encapsulate diverse signatures specific to different activities. Consequently, these markers are employed to evaluate unusual and hazardous driving behaviors. Nonetheless, it is important to note that these signatures frequently diverge across diverse demographic groups (such as varying age brackets), introducing complexities in designing a universally applicable system. Moreover, these wearables necessitate active user engagement and may impose intrusive implications (such as EEG sensors attached to the body), potentially disrupting individuals' daily routines.

Approaches Involving Acoustics: Researchers have also delved into the realm of acoustics to explore distracted and hazardous driving situations. This involves the exploitation of phenomena like doppler shifts [30,31] and FMCW chirps [18] to gather insights. However, the presence of ambient noise significantly affects the accuracy of such acoustic-based sensing [32]. Even with widespread devices like smartphones being employed for acoustic sensing, it is crucial to recognize that distinct devices exhibit diverse sensitivity profiles [33] towards audio signals, which can impact the efficacy of these systems. Additional variables, including the placement and orientation of the sensor, and the intricate matter of privacy, also raise considerations with regard to acoustic sensing.
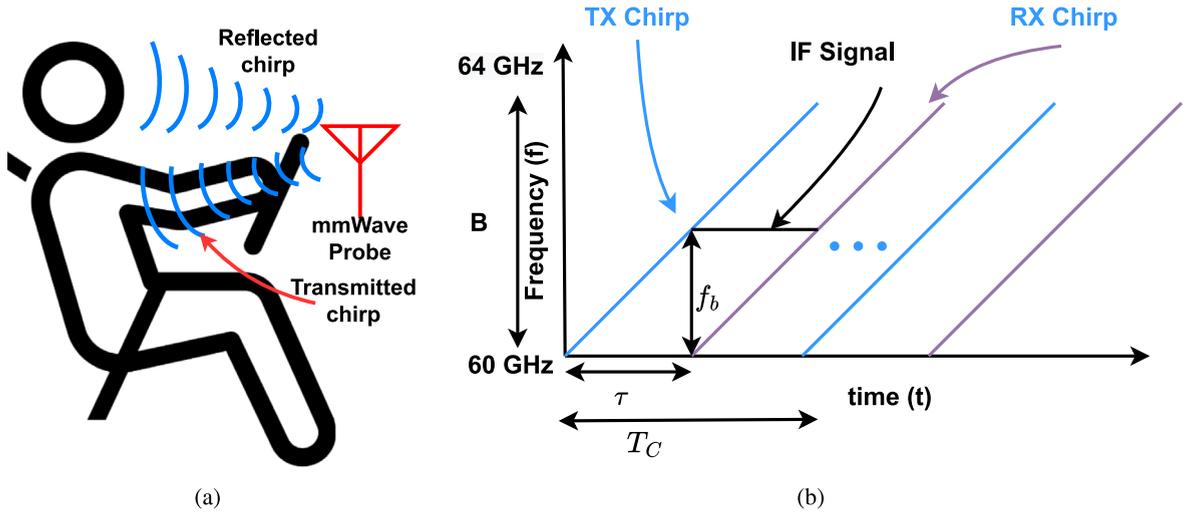
Fig. 2. Working principle of an FMCW radar — (a) Radar Placement (b) Frequency of multiple chirps vs. time.

3. Background and observations

This section introduces the technical elements of an FMCW mmWave radar that are pertinent to this study. Following that, we detail our findings regarding dangerous driving behaviors that can be associated with these factors.

3.1. Preliminaries

In a typical FMCW radar, a linear *chirp* is employed for transmitting continuous periodic signals. These signals interact with obstacles in the surroundings and get reflected. Subsequently, the radar engages in a process called *dechirping*, achieved by mixing the transmitted and reflected signals. The outcome of this process is an *Intermediate frequency* signal (IF), which then undergoes a series of procedures to extract necessary information.

Determining Subjects Position: When considering a transmitted *chirp* (T_X) and its corresponding received *chirp* (R_X), along with a transmission duration T_C, the distance d of the subject causing the reflection from the radar can be computed as follows:

$$d = \frac{c}{2} \cdot \frac{T_C}{B} \cdot f_b \tag{1}$$

In this equation, f_b denotes the beat frequency, which represents the difference in frequency between the T_X and R_X chirps. The value of c corresponds to the speed of light, and the fraction $\frac{T_C}{B}$ represents the slope S of the FMCW chirp.

To begin, a *Range Fast Fourier Transform* (Range-FFT) is initially applied to the IF signal to isolate all the reflectors, which are identified by their frequency peaks. Subsequently, Eq. (1) is utilized to calculate the distance of a reflector. Ultimately, the radar records the received power at various *range bins*, generating a one-dimensional array representing the range profile.

Estimating Subject Movements: In the context of an FMCW radar, a series of N chirps are transmitted in a continuous manner, each spaced by a transmission time of T_C. When the subject is in motion, the range-FFT associated with these chirps will exhibit peaks in their exact positions, albeit accompanied by a *phase change*. Assuming a velocity of v at which the subject is moving and denoting λ as the wavelength, the observed phase difference between two consecutive R_X chirps that corresponds to a motion of $v \times T_C$ can be expressed as follows:

$$\Delta\phi = \frac{4\pi(v \times T_C)}{\lambda} \tag{2}$$

Subsequently, a second FFT, referred to as the *Doppler-FFT*, can be applied to these N phases to analyze the subject's motion. The radar then captures information about both range and velocity within a 2D matrix termed a *range-doppler* heatmap.

Evaluating Noise Profile Characteristics: Doppler data across various range bins is utilized to compute the *noise profile*. This profile maintains the same structure as the *range profile*, but it is extracted from the maximum Doppler bin, where the most significant body movements occur.

Our methodology, *mmDrive* , makes use of the *range profile*, *range-Doppler* heatmap, and the *noise profile* to capture discernible patterns in *driver activities*. To achieve this, we conduct a preliminary investigation involving diverse driver behaviors, drawing insights from these key factors.

3.2. Pilot study

The initial challenge in devising *mmDrive* revolves around defining the scope of irregular and hazardous driving actions. Drawing on existing definitions of dangerous driving actions as outlined in literature [34], we expand these definitions to encompass a broader spectrum of intricate activities that serve as indicators of perilous driving. Our primary classification of these activities comprises two fundamental categories: (1) actions arising from fatigue or drowsiness and (2) actions arising from distraction.

Actions arising from Drowsiness: Within this category, actions typically emerge when a driver is experiencing fatigue or drowsiness, which could potentially lead to accidents. Some of these actions include:

- *Nodding*: When a driver feels drowsy or on the brink of falling asleep, they often exhibit rapid, cyclic bowing motions of the head in both the forward and backward directions (refer to Fig. 3(a)).
- *Yawning:* A common manifestation of driver fatigue is yawning. Yawning involves the sequential movements of widely opening the mouth during inhalation, lifting the head due to the reactive force during exhalation, and ultimately lowering the head to its original position (refer to Fig. 3(b)).
- *Steering Anomalies:* Typically, a vigilant driver navigates turns smoothly and consistently. However, a drowsy driver might refrain from adjusting the steering wheel for prolonged durations due to drowsiness. As the vehicle drifts laterally, the driver may abruptly turn the wheel to correct the course, resulting in distinctive movement patterns that are absent during regular driving (refer to Fig. 3(c)).

Actions Arising from Distractions: The category of distracted driving encompasses behaviors that indicate a lack of concentration on driving, potentially leading to hazardous situations. Below, we outline several such activities that *mmDrive* aims to identify:

- *Eating or Drinking:* Consuming food or beverages while driving often involves a minor forward inclination followed by retrieving items. In this scenario, one hand may briefly leave the steering wheel to pick up and consume the item (refer to Fig. 3(d)).
- *Looking Back:* A driver might turn their attention towards rear passengers, check items in the backseat, or attend to children seated behind them. This diverts focus from driving and usually entails head-turning and upper-body twisting (refer to Fig. 3(e)).
- *Retrieving Dropped Items:* A driver can become distracted when attempting to retrieve a fallen item from the car floor while driving. This action generally involves moving the hand and upper body downward (refer to Fig. 3(f)).
- *Fetching Objects Forward:* Drivers might reach out for something placed on the car dashboard while driving. This action usually requires leaning forward along with relevant hand movements (refer to Fig. 3(g)).
- *Using Mobile Phone:* The use of mobile phones is a well-known source of distraction. A driver might answer a call, send a text, or engage with an app by raising the mobile phone (refer to Fig. 3(h)).
- *Conversing with Front Passenger:* Engaging in conversation with the front-seat passenger can also divert the driver's attention (refer to Fig. 3(i)). This typically involves head-turning and mouth movements as the driver speaks.

Other Miscellaneous Movements: Apart from the previously discussed movements, there are additional factors in the vehicle environment that can introduce movements, such as *road bumps*. Although these movements do not carry pertinent information about dangerous driving, their patterns can be significant enough to confuse the classification model. Consequently, it is imperative to detect and filter out such movements to ensure accurate classification.

It is important to note that the highlighted movements associated with the aforementioned cases are not momentary but have a brief duration. Subsequently, we delve into the distinctive characteristics observed for the mentioned events. For each scenario, to capture these characteristics, we position an `AWR6843ISK`[1] mmWave FMCW radar on the car's dashboard in front of the driver, as depicted in Fig. 2(a).

3.2.1. Analysis of range-doppler heatmap

The *range-doppler* data is presented as a 2D image, where the *x-axis corresponds to the range (spanning* 32 *out of the* 64 *available indices, following the* `AWR6843ISK` *specifications)*, the *y-axis represents the Doppler speed,* and the contained value indicates the power magnitude for the velocity across various range and Doppler bins. In our setup, the typical distance between the driver and the mmWave sensor is approximately 0.6 to 0.8 m, translating to range bins between 6 and 16. As a result, higher power values are observed within this range, as illustrated in Fig. 4.

For bodily movements like reaching forward, picking up items, nodding, looking back, etc., the driver engages in specific movements, maintains that posture for several seconds, and then returns to a normal position. During such activities, we notice that as the driver initiates the action, the Doppler frequency shifts upwards (or downwards, contingent upon the movement direction). Upon reverting to the standard position, the Doppler frequency shifts downward (or upward, depending on the movement sequence) before eventually returning to zero as the activity concludes. This pattern results in an evident circular configuration within the range-doppler heatmap for these actions, as depicted in Fig. 4.

In instances involving sudden steering changes, such as a steering anomaly, the driver's abrupt manoeuvring of the steering wheel leads to a rapid alteration in the range-doppler heatmap across various range bins. In contrast, activities like yawning, drinking, or using a mobile phone generally entail fewer bodily movements while driving. Consequently, the variations in the range-doppler exhibit a lower magnitude in power values.

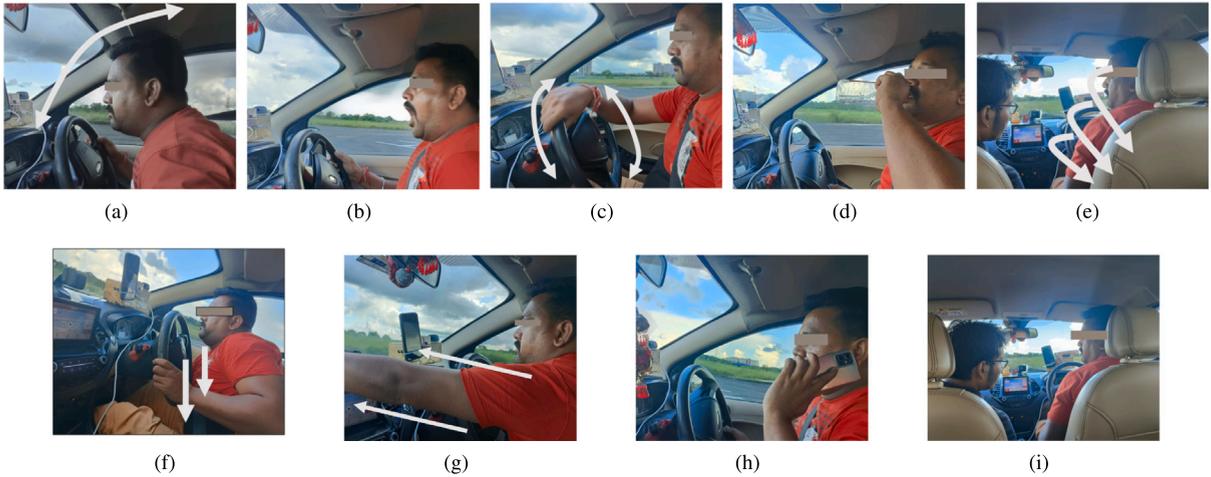[1] https://www.ti.com/tool/AWR6843ISK (Accessed: June 7, 2024).

Fig. 3. Dangerous driving activities — (a) Nodding, (b) Yawning, (c) Steering anomaly, (d) Drinking, (e) Talking to the rear passenger, (f) Picking a drop, (g) Fetching from the dash, (h) Using mobile, (i) Talking sideways.
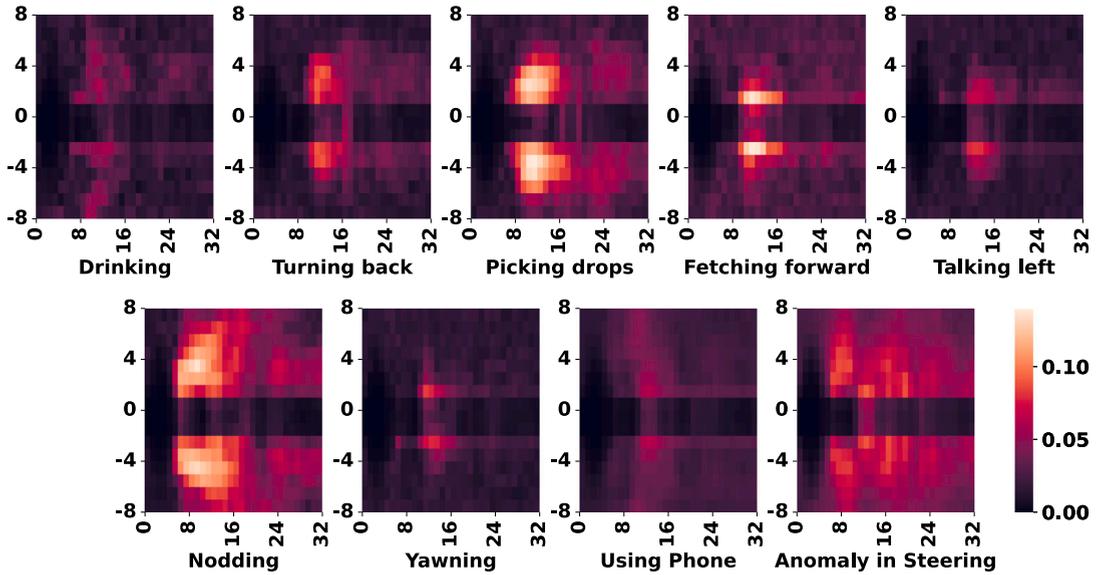

Fig. 4. Range-doppler heatmaps for driving actions.

3.2.2. Analysis of range-doppler, range profile, and noise profile for different drivers

To delve deeper into the observed variations, we present the mean distribution of the range-doppler, range profile, and noise profile for distinct driving actions performed by two different drivers in Fig. 5. Analyzing Figs. 5(a) and 5(b), it becomes evident that actions such as yawning, using a phone, conversing with the person on the left (adjacent seat), or drinking exhibit lower variation, as indicated by the denser distribution within the interquartile range. In contrast, for other driving actions, the distribution tends to be less dense around the median, signifying notable *macro body movements* made by the driver.

Furthermore, Fig. 5(c) showcases the distribution of the range profile across diverse driving actions. The variation across activities is relatively less pronounced. However, the vertical separation between the medians of the range profiles for the two drivers is considerably noticeable, attributable to disparities in driver height, seating posture, and position.

3.2.3. Analysis of different movements over time

Continuing our investigation, we proceed to analyze the influence of distinct movements and their corresponding signatures over time. This assessment encompasses even the effects of abrupt motions due to road bumps, which we monitor through an IMU sensor positioned on the car dashboard. The *z*-axis of the IMU's accelerometer exhibits a peak during road bumps, indicating the extent of the induced movement. Fig. 6 displays the mean values of range-doppler, noise profile, and IMU data over each second across a three-minute span.
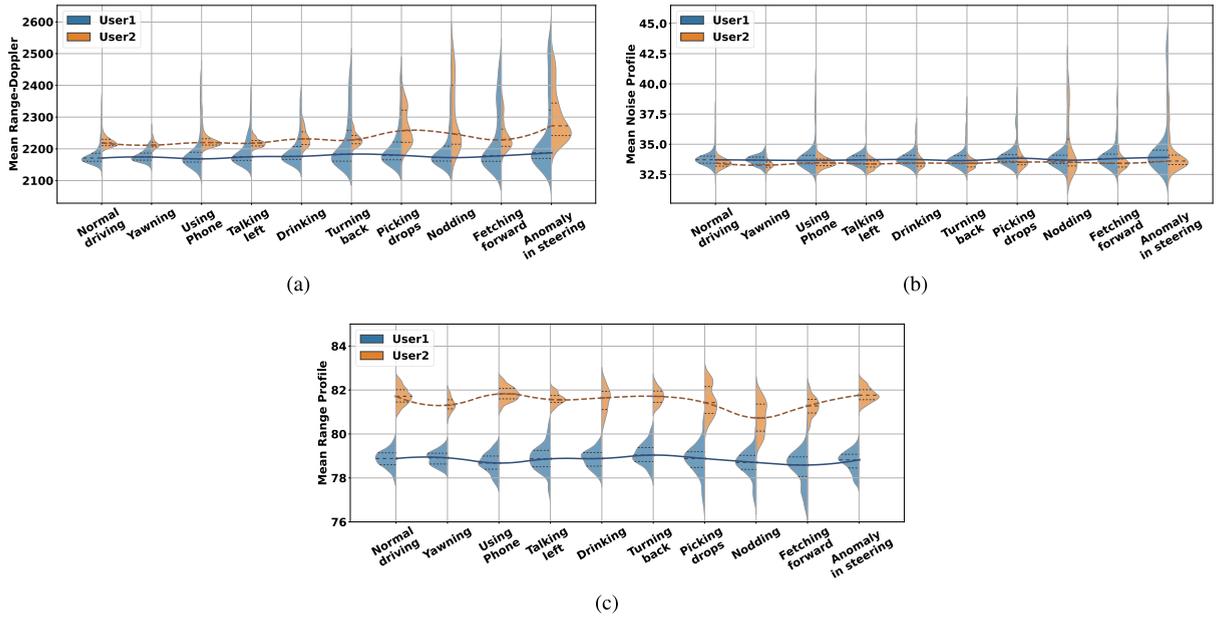
Fig. 5. Variation in the mean of (a) range-doppler, (b) noise profiles, and (c) range profiles across different driving actions for two users.
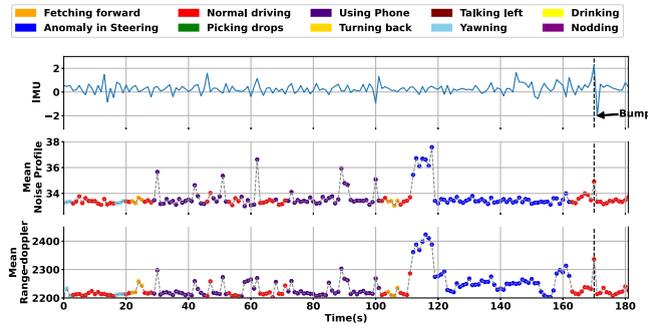

Fig. 6. Variation in the IMU data and radar measurements.

The depicted figure unveils that every activity exhibits its own temporal duration and mean value. For instance, yawning occurs briefly and results in a low mean value (close to the mean during standard driving) within the range-doppler. When a driver utilizes a phone, the behavior spans a duration similar to normal driving, barring instances where substantial body movements transpire, like picking up the phone or adjusting the hand from the steering. Conversely, a steering anomaly endures for an extended period, featuring significant mean variations compared to normal driving. This insight underscores that each driving behavior etches its distinct signature within the range-doppler or noise profile, accompanied by a unique time frame.

Transitioning to the IMU data, it becomes apparent that while the driver enacts these activities, the fluctuations are not particularly significant. Nonetheless, the IMU data registers a peak when encountering a road bump (around the 170th second in Fig. 6). Interestingly, this peak coincides with peaks in the mean values of both the range-doppler and noise profile data. This observation underscores the direct impact of road bumps on mmWave measurements.

These findings underscore the individuality of each driving behavior in terms of both spatial and temporal variations. Consequently, an effective classification of driving behaviors necessitates the consideration of both *spatial and temporal variations* within the feature space. Simultaneously, instances involving road bumps need to be effectively addressed.

3.2.4. Analysis of driving behavior distribution

By evaluating the mean values of range-doppler, range profile, and noise profile, the viability of classifying driving behaviors can be assessed. The range-doppler heatmap is structured as a 16×64 2D matrix, representing 64 range bins and 16 doppler bins.

In contrast, both the range profile and noise profile take the form of 1D arrays, each comprising 64 elements to represent 64 range bins. Consequently, a total of $(16 \times 64) + (64 \times 2) = 1152$ dimensions are implicated. We apply a T-distributed neighbor embedding (t-SNE) based dimensionality reduction technique to condense this higher-dimensional (1152) space into a 2D space, allowing us
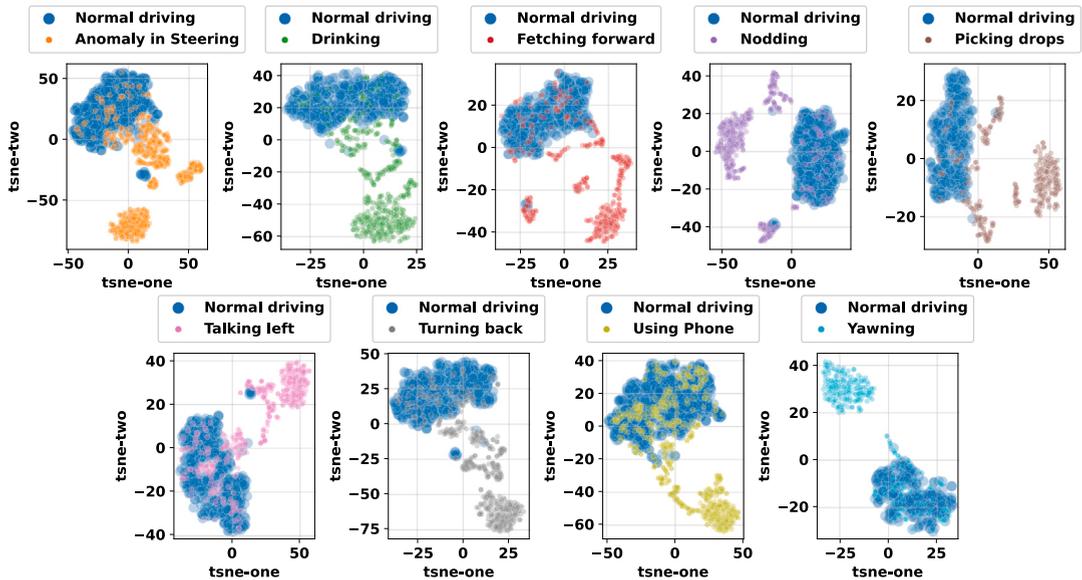
Fig. 7. Distributions of normal driving with other dangerous driving actions in two-dimensional feature space.

to visualize the separation between each hazardous driving action and normal driving behavior. The outcome, as demonstrated in Fig. 7, reveals that hazardous driving actions involving pronounced bodily movements are distinctively discernible from regular driving behavior. Nonetheless, some overlap can be observed in scenarios such as normal driving and less substantial activities like using a phone or conversing with a passenger.

The principal takeaways from this analysis can be summarized as follows:

- Range-doppler, noise profile, and range profile exhibit distinct patterns across diverse driving actions and drivers.
- Each driving action leaves behind a unique signature upon completion.
- Features characterizing hazardous driving actions with micro-body movements tend to share similarities with normal driving behavior.
- Real-world driving conditions, including road bumps, can introduce variations within the feature space, potentially complicating model predictions.

Building on these insights, our subsequent step involves the development of *mmDrive* for effectively classifying distinct driving behaviors.

4. Methodology

A broad overview of the procedural stages comprising the development of *mmDrive* is presented in Fig. 8. In this methodology, an FMCW radar captures measurements tied to the driver's bodily movements, while an IMU sensor records mobility patterns specific to the vehicle. These data sources collectively serve as input for subsequent processing within *mmDrive* . The ensuing steps are elucidated as follows.

4.1. Pre-processing

Driving actions are not isolated instantaneous events; rather, they unfold over a period characterized by distinct temporal patterns in the relevant features. To effectively capture this temporal variation, we aggregate 10 consecutive feature frames together. This quantity is derived through empirical analysis, as elaborated upon in Section 5.4. Note to mirror real-world driving patterns more closely; we have now revised our approach. Our dataset is divided based on timestamps, with the initial 80% allocated for training and the remainder for testing. In our previous approach, we trained the model using a 70%–30% train–test split with random sampling. However, this could lead to data leakage, as the testing set might contain samples with instances of mmWave data from activities used for model training. By switching to a timestamp-based train–test split, we rectify this potential data leakage issue, resulting in a more accurate representation of real-world driving scenarios. Subsequently, we subject the concatenated features to normalization using a *min–max scaler,* which scales the features to fall from 0 to 1.
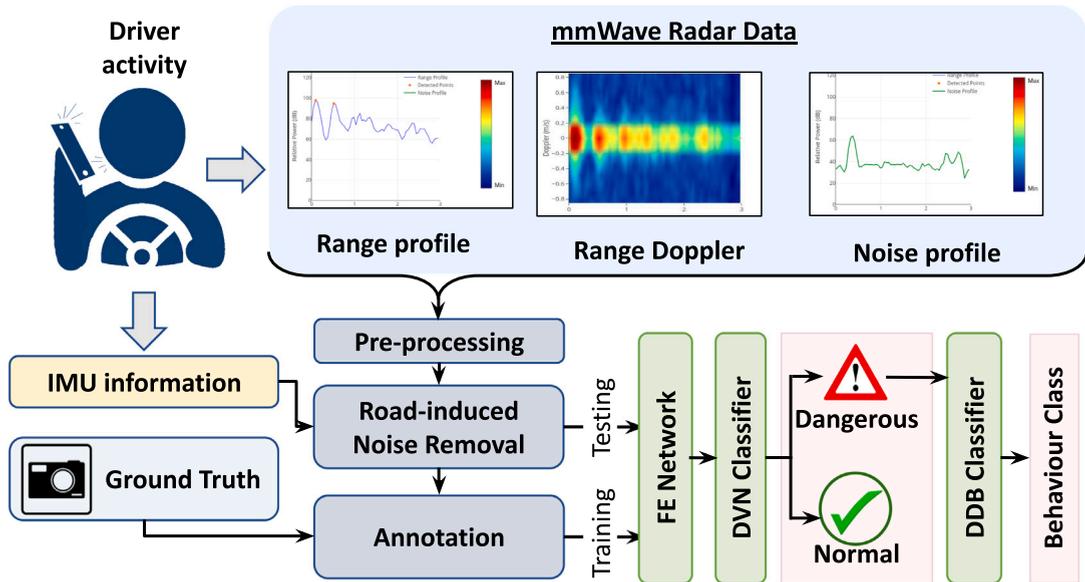
Fig. 8. Overview of *mmDrive* .

4.2. Removal of road-induced noise

As discussed in Section 3, unfavorable road conditions like road fractures, speed bumps, etc., can directly influence the noise profile and range-doppler heatmap. This variability in the feature space mirrors the impact stemming from the driver's driving actions, thereby introducing ambiguity if not addressed adequately. To identify and address such undesired variations, we employ an alternative modality through IMU sensors located on the car dashboard. Following the approaches outlined in previous studies [30], we utilize the z-axis acceleration of the vehicle to detect instances of jerkiness attributed to suboptimal road conditions or bumps. This is achieved by establishing thresholds for the acceleration data. Subsequently, within the framework of *mmDrive* , the corresponding data derived from the mmWave sensor is filtered to mitigate the presence of undesirable noise in the feature space.

4.3. Classification pipeline

Distinguishing itself from conventional approaches that rely solely on techniques like Convolutional Neural Networks (CNNs) and range-doppler spectrograms [35–37], our approach leverages the combined potential of range-doppler, range profile, and noise profile features to classify both normal and hazardous driving behaviors. In this context, we introduce an innovative Fused-CNN architecture, illustrated in Fig. 9. The specifics of this model are elucidated in the subsequent sections.

4.3.1. Feature Extraction (FE) network

The range profile contributes to the maximum power value at zero-doppler regions, whereas the noise profile supplies noise floor power at non-zero Doppler regions. This concatenation aids in comprehending the spatial dynamics of the driver across the entire range bins. We concatenate 10 feature frames, as previously mentioned, to encompass temporal variations. Consequently, the range and noise profiles manifest as vectors sized $64 \times 1 \times 10$, and their concatenation yields an array of dimensions $64 \times 2 \times 10$. In contrast, the range-doppler heatmap constitutes a stacked 2D feature resembling an image with dimensions $16 \times 64 \times 10$.

Our approach entails three Convolutional 2D layers employing valid padding and ReLU activation. This is followed by global average pooling, yielding 96-dimensional range-noise-based embeddings. Similarly, the range-noise feature extraction employs 2D convolutional operations to capture dependencies among neighboring values within all possible $k \times k$ regions within each range-doppler frame. These operations also consider the temporal interplay among the preceding 10 frames. This procedure generates multiple cross-channel feature maps beneficial for subsequent model layers. Subsequently, four 2D Convolutional layers are employed, maintaining the same padding and ReLU activation. These are succeeded by global averaging, resulting in 128-dimensional range-doppler feature embeddings. Moving forward, the concatenated range-noise and range-doppler feature embeddings are processed through two consecutive modules within the proposed architecture.
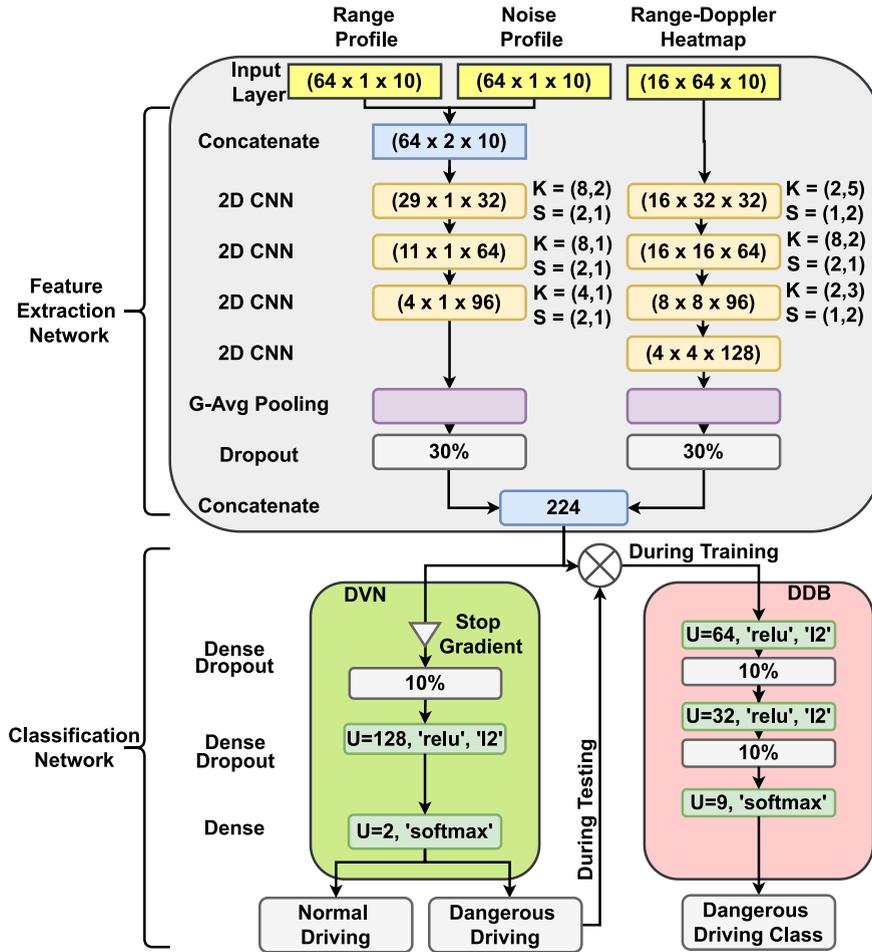
Fig. 9. Fused-CNN model architecture.

4.3.2. Dangerous Driving Behaviors (DDB) classifier

The DDB (Dangerous Driving Behavior) classifier receives the cross-channel feature embeddings from the FE Network and guides them through a sequence of two consecutive layers: Dropout and Dense. A dropout rate of 10% is implemented in the Dropout layers to counteract overfitting. The final layer, featuring softmax activation, consists of 9 neurons, culminating in an output representing a comprehensive probability distribution encompassing the nine distinct dangerous driving actions.

It is noteworthy that the FE Network undergoes training exclusively via backward gradients originating from the DDB Classifier. This approach facilitates the acquisition of efficient feature extraction capabilities optimized for classifying the 9 dangerous driving behaviors.

4.3.3. Dangerous vs. Normal driving (DVN) classifier

The DVN Classifier capitalizes on the acquired feature embeddings from the pre-trained FE Network to effectively categorize all variants of hazardous driving behaviors when contrasted with normal driving. Importantly, the weights that the FE Network learns during the DDB Classifier training remain unaltered, given the halt in backward gradient propagation beyond the DVN Classifier. This strategic decision optimizes overall training duration and maintains the model's compactness, rendering it suitable for real-time implementation.

For further insights into the intricate architecture of the Fused-CNN network, you can refer to Fig. 9.

4.3.4. Lazy inferencing

In the context of real-world application, the DVN Classifier plays a pivotal role in discerning whether the ongoing driving behavior falls within the category of normal or hazardous. When the DVN Classifier detects normal driving behavior, *mmDrive* optimizes its execution by deferring the involvement of the DDB Classifier. This strategic decision serves to curtail computational load and conserve energy resources. The DDB Classifier is activated solely in scenarios where the DVN Classifier identifies indications of potentially dangerous driving behavior. Subsequently, the DDB Classifier refines its assessment by determining the specific class of dangerous driving behavior.

Table 1
Radar configuration.

Parameters	Value
Start frequency (GHz)	60
Range resolution (m)	0.0375
Maximum unambiguous range (m)	2.41
Maximum radial velocity (m/s)	1
Radial velocity resolution (m/s)	0.13
Frames per second	5
Number of chirps per frame	64
Baud-rate	921 600 bps

5. Implementation and evaluation

In order to assess the real-world implications of dangerous driving behaviors, we put into operation the experimental setup within actual driving scenarios. This endeavor encompasses 5 distinct cars and involves 7, different drivers. To facilitate data collection, the radar is strategically positioned on the car's dashboard. Following a comprehensive data-gathering process and meticulous annotation of ground truth labels, we subject our novel Fused-CNN model to evaluation, contrasting its performance against two contemporary state-of-the-art baselines. Further elaboration on the implementation and evaluation specifics is provided in the subsequent sections.

5.1. Hardware setup

Our *mmDrive* approach has been implemented utilizing the commercially available off-the-shelf (COTS) mmWave radar system named `AWR6843ISK`, manufactured by Texas Instruments. This FMCW-based mmWave radar operates within the frequency range of 60 to 64 GHz and boasts a range resolution of approximately 4 cm, which is suitably adept at capturing the activities of interest outlined in Section 3.2. The radar's maximum detectable range spans up to 10 meters, encompassing a Field-of-View spanning $-70°$ to $+70°$ both azimuthally and in elevation. This Field-of-View configuration is conducive to detecting hazardous driving behaviors taking place within the confines of the vehicle.

mmDrive effectively measures variations across 64 range bins, each representing a distance of 2.4 meters from the dashboard. This provides a refined range resolution of 3.75 cm. Furthermore, the system captures doppler information spanning 16 doppler bins via the 2D Fast Fourier Transform (FFT), affording a velocity resolution of 0.13 m/s. This resolution proves sufficient for accurately capturing both micro and macro body movements of the driver within the realm of real-world driving scenarios. The configuration parameters of the mmWave radar are detailed in Table 1.

For real-time driving behavior detection, a Raspberry Pi 4 Model-B with 8 GB RAM is employed. The mmWave radar and IMU sensor are connected to the Raspberry Pi 4 via USB and the I2C bus, respectively. The IMU sensor serves to filter out road-induced noises, as expounded upon in Section 4.2. Parsing of the mmWave radar's serial data is executed, and the resultant data is forwarded to the classification pipeline for the inference of driving behavior. Fig. 10(c) shows the latest hardware setup with the 3D-printed stand for incorporating mmWave radar with the Raspberry Pi-4.

To establish the ground truth for driving actions, the Nexar Pro Smart Dash Camera[2] is utilized. Videos collected from the Nexar camera serve as a reference for annotating the features generated by the mmWave Radar. The hardware setup employed for field experimentation is visually depicted in Fig. 10(a). Further elucidation on the data collection strategy is presented in the subsequent subsection.

5.2. Data collection

To procure raw mmWave features employing the aforementioned hardware components, we harnessed the mmWave-Demo-Visualizer[3] and customized its source code to facilitate the collection of raw data based on user input. The data collection initiative encompassed 7 distinct users, spanning a cumulative time span of 60 h. Each user contributed ≈8.57 h to the endeavor. This data accumulation endeavor was conducted across three diverse vehicles, including three sedans, one SUV and one minivan. Among the driving cohort, one participant was female, while the remaining were male, with ages ranging from 24 to 55.

The amassed dataset manifests a distribution across various dangerous driving behaviors, as depicted in Fig. 10(b). This comprehensive dataset encompasses mmWave radar data, IMU readings, GPS data, video footage, and audio recordings captured from the roadside, as well as the car cabin's rearview perspective. These records were obtained through the utilization of the Nexar system. The process of annotating hazardous driving behaviors from the interior car camera footage was facilitated by the involvement of three volunteers.

---

[2] https://www.getnexar.com/global/the-dash-cams.
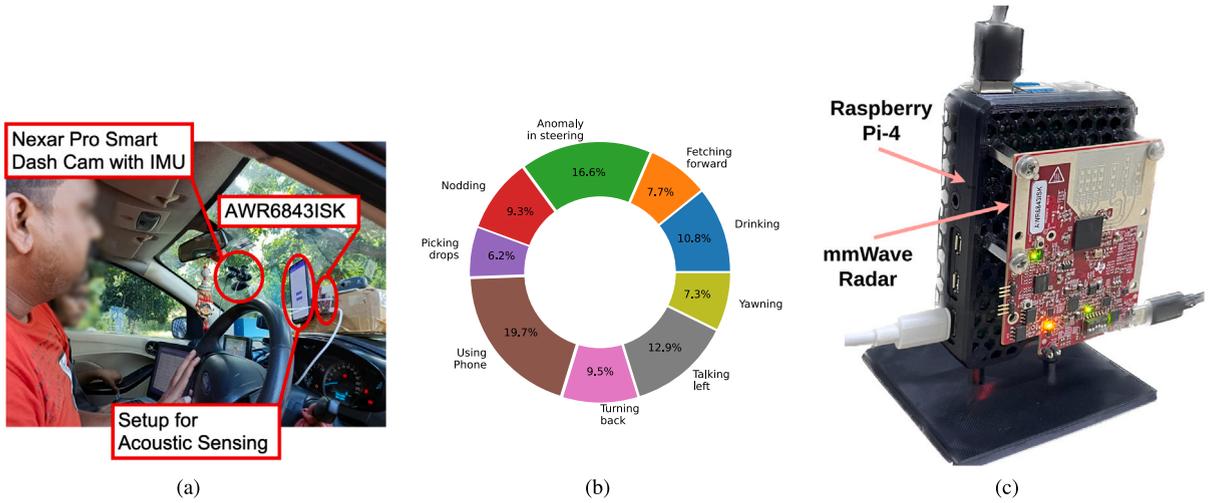[3] https://dev.ti.com/gallery/view/mmwave/mmWave_Demo_Visualizer/.

Fig. 10. (a) Hardware setup (b) Collected data distribution, (c) 3D Printed *mmDrive* probe.

5.3. Software setup and baselines

In the implementation of our driver behavior classifier model, we utilized Python 3.9, Tensorflow v2.10, and Scikit-learn v1.1. Alongside our novel Fused-CNN-based approach, we also employed two established state-of-the-art baselines: Random Forest (RF) and the VGG-16 network [38]. The model training procedures were executed on an iMac boasting 16 GB of primary memory, running MacOS v12.6 with a base-kernel version of 21.6.0.

For the Random Forest model, features were engineered by computing statistics such as minimum, maximum, mean, and standard deviation for the range profile, noise profile, and range-doppler data, within a kernel size of 16×1, 16×1 and 16×16 for the total 10 frames with an array size of 64×1, 64×1 and 16×64 respectively, resulting $(\frac{(64 \times 1)}{(16 \times 1)} + \frac{(64 \times 1)}{(16 \times 1)} + \frac{(16 \times 64)}{(16 \times 16)}) \times 4 \times 10 = 480$ features. The Random Search Cross Validation technique [39] was employed to explore the optimal hyperparameters for the RF classifier. The model was trained with 400 estimators based on the outcome of this hyperparameter search.

Additionally, the VGG-16 [38] network was employed as another baseline. This network was initialized with pre-trained weights from the ImageNet dataset [40], leveraging transfer learning to facilitate feature extraction from high-dimensional image-like input data. In the Fused-CNN architecture, we augmented the base VGG-16 model with global average pooling, followed by successive Dropout and Dense layers for the purpose of classifying hazardous driving behaviors. Model training was conducted utilizing a train–test split of 70%–30%, with a validation split of 20% derived from the training set.

An alternative baseline was established employing an acoustic modality based on FMCW radar. Previous works such as [18,31,41] use an acoustic sensing approach to detect dangerous driving behaviors with fewer behavior classes. This approach utilized the smartphone's embedded microphone and speaker to transmit and receive raw chirps in the near-ultrasound range of 16–19 kHz. Subsequent post-processing encompassed applying Range-FFT to derive the amplitude and phase of the IF (Intermediate Frequency) signal across distinct range bins. A 2D-doppler FFT was subsequently executed to generate the range-doppler heatmap. Lastly, a Random Forest-based classifier was employed to evaluate and compare its performance against our *mmDrive* approach.

5.4. Results

Initially, we explore the influence of combining (stacking) sequential temporal feature frames acquired from the mmWave radar on the precision of detection. Subsequently, we delve into an elaborate examination of outcomes, focusing on the outcomes obtained through the most effective frame-stacking strategy.

5.4.1. Impact of frame stacking on F1-score

As depicted in Fig. 11(a), the duration of each driving action exhibits a notable skew, with a median value of approximately 16 feature frames derived from the mmWave radar data. This underscores the temporal influence of driving actions on the radar measurements. To effectively capture this temporal impact, multiple frames are aggregated through stacking.

In order to determine the optimal frame stacking approach for detecting dangerous driving behaviors, an exploration of the number of frames stacked is conducted, ranging from 1 to 16, as illustrated in Fig. 11(b). The figure notably illustrates the direct correlation between frame stacking and the overall performance of the classification pipeline. This trend reveals that as the number of frames increases, activities characterized by longer durations exhibit improved F1-Scores. For instance, activities such as using a phone or experiencing an anomaly in steering, which entail lengthier time intervals, demonstrate enhanced performance with
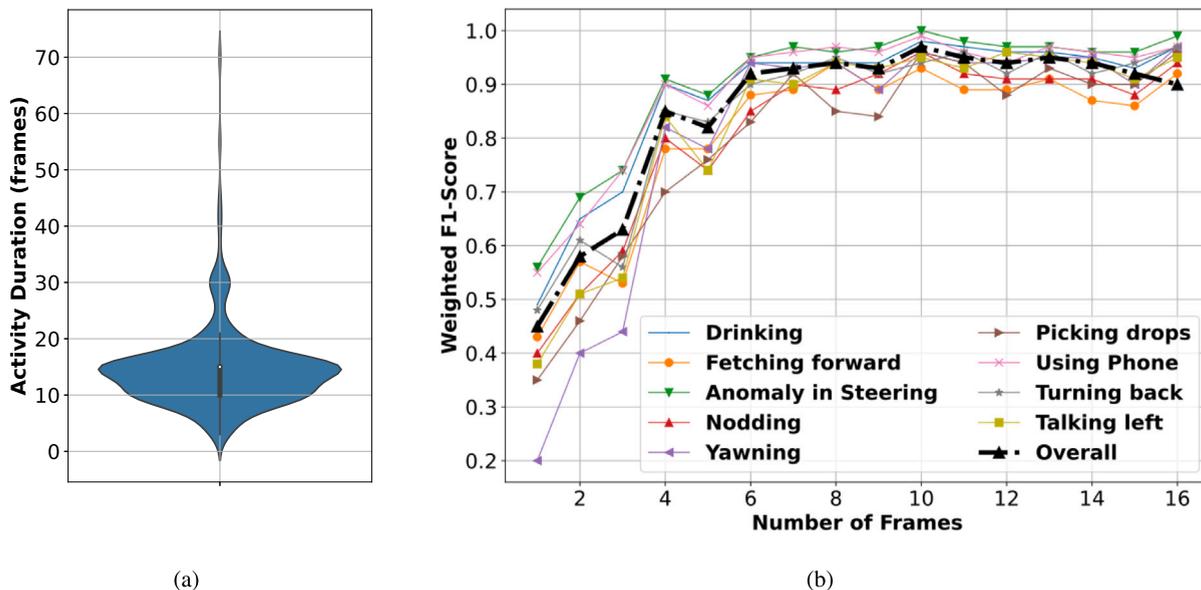
(a)

(b)

Fig. 11. (a) Distribution of Frames required to complete individual driving actions, (b) Variation in the weighted F1-Score with number of stacked frames.
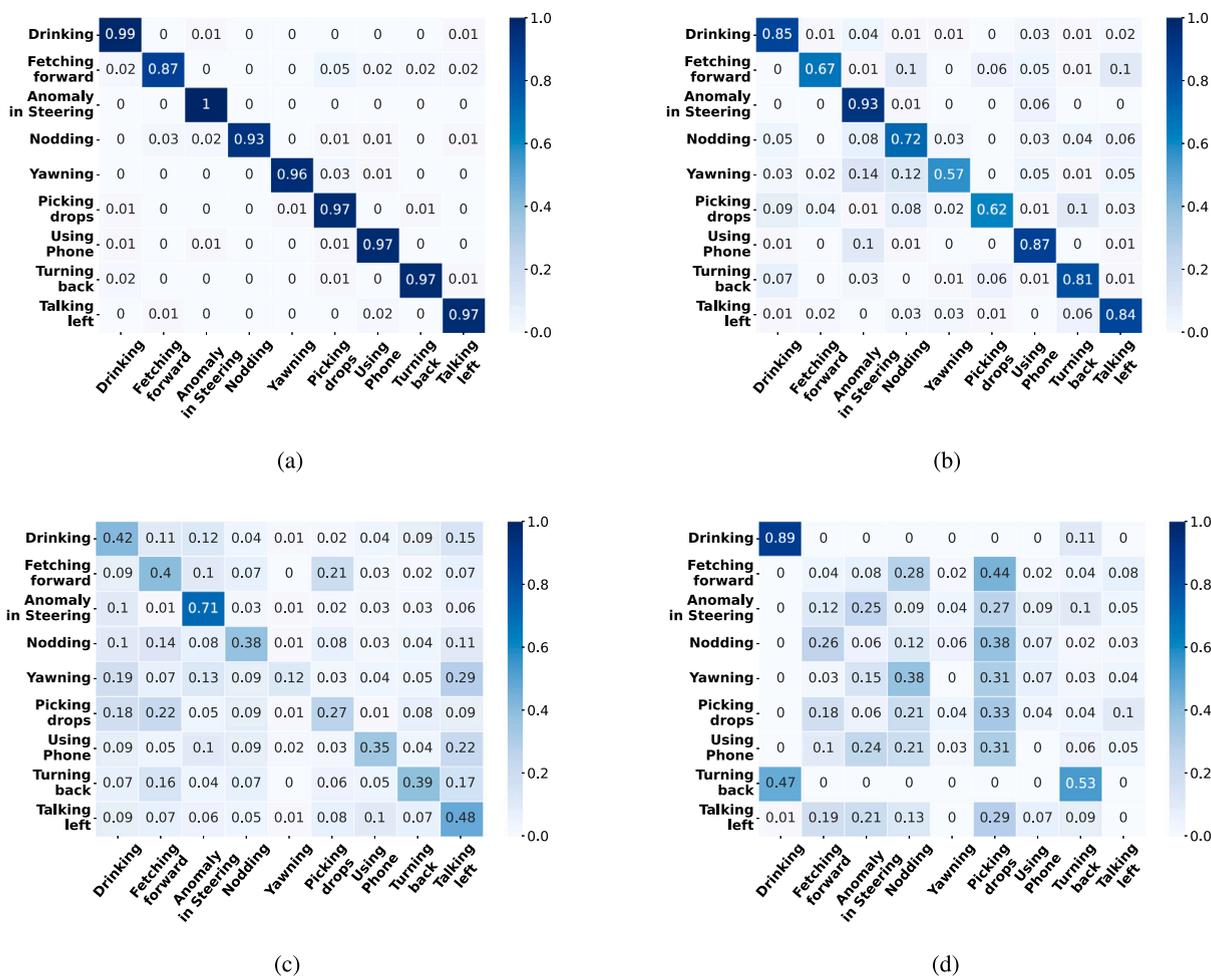

(a)

(b)

(c)

(d)

Fig. 12. Confusion matrix for all the dangerous driving behaviors — (a) Fused-CNN, (b) RF, (c) VGG-16, (d) Acoustic-FMCW.
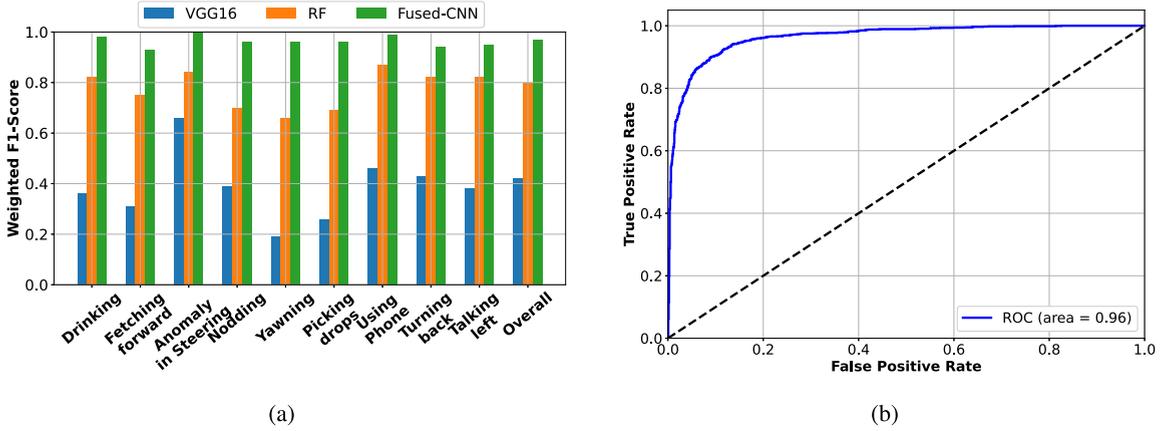
Fig. 13. (a) Weighted F1-Score across driving behaviors, (b) AUC-ROC for classifying dangerous vs. normal driving.

increased frame stacking. However, behaviors with shorter durations, like fetching forward or nodding, experience a decline in F1-Score.

The data presented in Fig. 11(b) highlights that our proposed method (*mmDrive*) attains the highest overall F1-Score when employing a frame stacking of 10 frames, equating to a 2-second activity time window. Consequently, for the remainder of this paper, we employ a frame-stacking strategy of 10 frames.

5.4.2. Performance of DDB classifier

We conducted a performance comparison between our proposed Fused-CNN model and two baseline approaches: (i) Random Forest (RF) and (ii) VGG-16. The confusion matrix for all three classifiers is presented in Fig. 12. The outcomes from this matrix unmistakably demonstrate that our proposed Fused-CNN model outperforms the baseline methods.

Fig. 13(a) showcases the average weighted F1-Score for each individual dangerous driving action. Among the two baselines, RF exhibits a relatively better performance in contrast to VGG-16. The main reason for VGG-16's subpar performance lies in its expectation of a 2D input feature, such as a 2D range-doppler heatmap image in this scenario. As a result, it cannot fully leverage the features derived from range or noise profiles. Additionally, VGG-16's pre-training on the *imagenet* dataset does not align well with engineering features from a 2D range-doppler heatmap, which tends to be predominantly sparse across range bins, except where the driver is situated.

On the contrary, the RF model benefits from features passed through various kernel sizes, effectively capturing the spatial variations in range, noise profiles, and range-doppler heatmaps. This attribute contributes to its relatively improved performance compared to VGG-16. In contrast, our proposed Fused-CNN model possesses the advantage of autonomously learning spatiotemporal cross-features, as its features are not manually engineered like RF's (minimum, maximum, mean, standard deviation). Moreover, the Fused-CNN model demonstrates a lower inference time, courtesy of its fewer convolutional layers compared to VGG-16.

5.4.3. Performance of DVN classifier

Fig. 13(b) illustrates the ROC curve for the DVN classifier. As depicted, the area under the curve (AUC) stands at 0.96, signifying a high level of accuracy in distinguishing dangerous driving from normal driving behavior. Additionally, the observed weighted F1-Score amounts to $90\% \pm 0.5\%$.

5.4.4. Driver demographics

Moving forward, we proceed to evaluate the Fused-CNN model on an individualized level, considering all five drivers. Table 2 presents the action-specific F1-Scores of the model. We observe consistent and commendable performance across all five drivers, resulting in an overall weighted F1-Score of $91\% \pm 1\%$. It is noteworthy that certain activities pose more challenging detection tasks for specific drivers due to factors such as their height, sitting posture, and seating arrangement. Furthermore, driving behaviors naturally differ among individual drivers. Despite these driver-specific nuances, *mmDrive* exhibits robust generalization capabilities, delivering noteworthy performance across all drivers.

5.4.5. Evaluation on different cars

Assessing the proposed FUsed-CNN model's ability to identify hazardous actions exhibited by various car types accurately is crucial for enhancing driving behavior analysis. As shown in Fig. 14(a), the Sedan1 and Sedan2 categories exhibit the highest weighted F1-scores, both achieving an impressive score of 0.98. This outcome underscores the model's robust ability to detect hazardous actions within these sedan subtypes accurately. The Sedan3 category follows closely with a weighted F1-score of 0.96, further establishing the model's efficacy in classifying dangerous driving behaviors in this car type. The SUV category maintains a
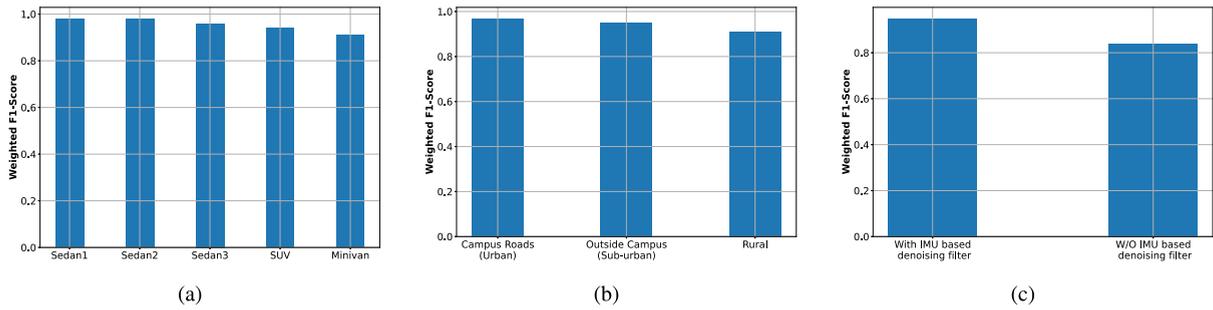
Fig. 14. Weighted F1-Score under (a) different cars, (b) different roads, (c) impact of IMU-based denoising filter.

Table 2
Driver-wise weighted F1-score.

Driving behavior	Driver-1	Driver-2	Driver-3	Driver-4	Driver-5	Driver-6	Driver-7
Drinking	0.90	0.96	0.80	0.88	0.92	0.82	0.88
Fetching forward	0.94	0.86	0.91	0.94	0.87	0.97	0.95
Anomaly in steering	0.96	0.91	0.98	0.91	0.93	0.98	0.96
Nodding	0.95	0.89	0.97	0.95	0.95	0.98	0.94
Yawning	0.96	0.86	0.81	0.87	0.88	0.85	0.82
Picking drops	0.96	0.85	0.87	0.82	0.84	0.98	0.96
Using phone	0.91	0.95	0.93	0.92	0.93	0.90	0.89
Turning back	0.91	0.95	0.92	0.94	0.91	0.93	0.94
Talking left	0.90	0.86	0.89	0.92	0.93	0.83	0.88
Overall	0.93	0.90	0.90	0.91	0.91	0.92	0.91

weighted F1-score of 0.94, indicating reliable performance in recognizing hazardous actions despite the variation in car design and features. The Minivan category possesses the lowest weighted F1-score among the groups at 0.91, as the Minivan is used only in rural road dynamics with more road-bumps or potholes.

5.4.6. Performance across rural and urban roads

Fig. 14(b) shows the weighted F1-scores insights into the Fused-CNN model's performance in detecting dangerous driving behaviors under varying road conditions. The highest weighted F1-score is achieved on Campus Roads (Urban), standing at an impressive 0.97. This result suggests that the model excels in identifying hazardous actions on well-maintained urban roads commonly found within a campus environment. Outside Campus (Suburban) roads exhibit a slightly lower but still commendable weighted F1-score of 0.95. This outcome underscores the model's capability to accurately classify dangerous driving behaviors even in moderately varying road conditions typically encountered in suburban areas. However, the most noticeable contrast arises in the Rural road category, where the model achieves a weighted F1-score of 0.91. This lower accuracy can be attributed to the increased presence of potholes and road bumps in rural areas. These road conditions introduce complexities that can lead to challenges in identifying hazardous behaviors accurately.

5.4.7. Impact of IMU in denoising road-induced noise

Fig. 14(c) shows the difference in weighted F1-scores between the two scenarios, highlighting the significant impact of the IMU-based denoising filter on classification accuracy. With the filter in place, the model achieves a weighted F1-score of 0.95, indicating a substantial improvement in identifying dangerous driving behaviors. This enhancement can be attributed to the filter's ability to detect and mitigate the noise caused by road irregularities, allowing the model to focus more effectively on actual driving behavior patterns. In contrast, without the IMU-based denoising filter, the model's performance exhibits a lower weighted F1-score of 0.84. This reduction in accuracy underscores the challenges posed by road-induced noise, especially from road bumps and potholes, which can lead to misinterpretation of driving behavior and decreased model effectiveness. The key advantage of the IMU-based denoising filter is its capacity to distinguish between true dangerous driving behaviors and noise originating from road conditions. By utilizing IMU data to identify and remove the effects of road irregularities, the filter enables the model to make more precise classifications, resulting in improved accuracy.

5.4.8. Performance of acoustic-FMCW

In Fig. 12(d), we present the confusion matrix depicting the classification results of the nine dangerous driving behaviors using the acoustic-FMCW approach [18]. The outcome of this classification is an accuracy of only 36%, indicating a significant struggle in distinguishing between the different dangerous driving behaviors. For instance, *fetching forward* is incorrectly identified as *picking/drops* and *nodding*, likely due to the similarity in the driver's body movements. Moreover, regarding the classification between dangerous and normal driving, the acoustic-FMCW approach yields F1-Scores of 60% and 73%, respectively. These outcomes emphasize the limitations of existing acoustic-based methodologies, which might perform reasonably well in scenarios involving fewer behavior classes, but prove impractical for the differentiation of nine distinct driving behaviors.
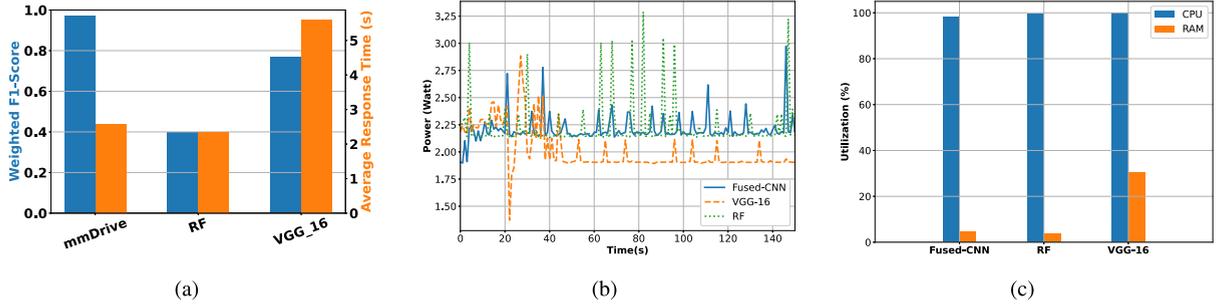
Fig. 15. (a) Weighted F1-Score and latency, (b) Power Consumption, (c) Resource Utilization.

5.4.9. Real-time system evaluation

We deployed *mmDrive* in a real driving environment to assess its efficiency in real-time inference. As illustrated in Fig. 15(a), *mmDrive* demonstrates superior accuracy compared to the baselines. Regarding latency, the DDB classifier of *mmDrive* takes approximately 2.5 s to infer the dangerous driving activity class. Among the baselines, RF exhibits the shortest inference time, approximately 2.33 s, albeit with a lower accuracy (weighted F1-Score of 78%). In contrast, VGG-16 performs the worst with an average latency of 5 s. It is worth mentioning that the average latency of the DVN classifier is even less, ($\approx$1.2) s, and only in case it detects a dangerous driving behavior, it will initiate the DDB classifier. Having a latency of less than 3 s is sufficient to classify dangerous driving activities, as per our study in Section 5.4.1 the activities usually take 3 (median at 16 frames) to 6 s.

5.4.10. Resource and power consumption

Lastly, the power consumption profiles, as well as the CPU and memory utilization of *mmDrive* 's Fused-CNN in comparison with the baselines, are presented in Fig. 15. As depicted in Fig. 15(b), it is noticeable that both Fused-CNN and RF exhibit higher power consumption when compared to VGG-16. Nonetheless, due to their lower computational demands, RF and Fused-CNN manage to utilize CPU and memory resources more effectively, resulting in lower latency for real-time inference. The peaks in Fig. 15(b) correspond to the initiation of an inference process. Furthermore, the longer latency of VGG-16 compared to Fused-CNN and RF-based classifiers underscores its unsuitability for real-time deployment. As shown in Fig. 15(c) on analyzing the memory utilization of all classifiers, it becomes evident that Fused-CNN employs CPU and memory resources in a similar manner to RF while achieving significantly higher accuracy.

6. Conclusion

Given the growing concern for road safety and the persistent issue of dangerous driving behaviors and driver inattention, it is imperative to address this problem effectively. Over the years, researchers have dedicated substantial effort to finding solutions that are not only efficient and widely applicable but also timely. In this study, we present a novel approach that leverages carefully selected features extracted from the measurements of a commercially available mmWave Frequency-Modulated Continuous-Wave (FMCW) radar. Our method, named *mmDrive* , stands out as a comprehensive solution to tackle this issue. Not only is it compact, adaptable, and entirely on-device, preserving user privacy, but it also boasts an impressive accuracy exceeding > 95% in detecting a critical range of driver actions indicative of hazardous driving situations.

We introduce the concept of *mmDrive* and illustrate its potential benefits, highlighting its real-world applicability. By utilizing data from a single mmWave FMCW radar, we demonstrate how our solution can effectively detect a variety of dangerous driving behaviors. Through extensive evaluations conducted in diverse real-world scenarios, we compare *mmDrive* 's performance against established baselines, showcasing its superiority. Given these promising results, we are confident that *mmDrive* holds the potential to make a significant contribution to enhancing road safety by accurately identifying hazardous driving behaviors and thereby playing a pivotal role in preventing potential accidents and safeguarding lives across a wide range of driving conditions. In our future work, we would like to extend *mmDrive* by incorporating an effective notification mechanism to alert the users. We also want to capture driver vital signs using the mmWave reflections and want to study how these physiological parameters correlate with dangerous driving actions. We can further introduce vital signs-related features for classifying the severity of dangerous driving actions.

CRediT authorship contribution statement

Argha Sen: Data curation, Formal analysis, Methodology, Writing – original draft, Writing – review & editing. **Avijit Mandal:** Data curation, Formal analysis, Software. **Prasenjit Karmakar:** Methodology. **Anirban Das:** Conceptualization, Software, Validation. **Sandip Chakraborty:** Conceptualization, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

I have shared the data in the github repository shared in the manuscript.

References

[1] I.T. Forum, Road Safety Annual Report 2020, OECD, 2020, p. 64, URL: https://www.oecd-ilibrary.org/content/publication/f3e48023-en. (Accessed 10 October 2022).

[2] A. Shaout, D. Colella, S. Awad, Advanced driver assistance systems-past, present and future, in: 2011 Seventh International Computer Engineering Conference, IEEE, 2011, pp. 72–82.

[3] J.-L. Yin, B.-H. Chen, K.-H.R. Lai, Y. Li, Automatic dangerous driving intensity analysis for advanced driver assistance systems from multimodal driving signals, IEEE Sens. J. 18 (12) (2017) 4785–4794.

[4] L. Jiang, W. Xie, D. Zhang, T. Gu, Smart diagnosis: Deep learning boosted driver inattention detection and abnormal driving prediction, IEEE Internet Things J. 9 (6) (2021) 4076–4089.

[5] H. Yang, L. Liu, W. Min, X. Yang, X. Xiong, Driver yawning detection based on subtle facial action recognition, IEEE Trans. Multimed. 23 (2020) 572–583.

[6] Y. Cao, F. Li, X. Liu, S. Yang, Y. Wang, Towards reliable driver drowsiness detection leveraging wearables, ACM Trans. Sensor Netw. (2022).

[7] J. Wang, W. Chai, A. Venkatachalapathy, K.L. Tan, A. Haghighat, S. Velipasalar, Y. Adu-Gyamfi, A. Sharma, A survey on driver behavior analysis from in-vehicle cameras, IEEE Trans. Intell. Transp. Syst. (2021).

[8] T. Kundinger, P.K. Yalavarthi, A. Riener, P. Wintersberger, C. Schartmüller, Feasibility of smart wearables for driver drowsiness detection and its potential among different age groups, Int. J. Pervasive Comput. Commun. (2020).

[9] S. Hur, S. Baek, B. Kim, Y. Chang, A.F. Molisch, T.S. Rappaport, K. Haneda, J. Park, Proposal on millimeter-wave channel modeling for 5G cellular system, IEEE J. Sel. Top. Signal Process. 10 (3) (2016) 454–469.

[10] G. Li, Z. Zhang, H. Yang, J. Pan, D. Chen, J. Zhang, Capturing human pose using mmwave radar, in: 2020 IEEE International Conference on Pervasive Computing and Communications Workshops, PerCom Workshops, IEEE, 2020, pp. 1–6.

[11] Y. Wang, H. Liu, K. Cui, A. Zhou, W. Li, H. Ma, M-activity: Accurate and real-time human activity recognition via millimeter wave radar, in: ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing, IEEE, 2021, pp. 8298–8302.

[12] S. Palipana, D. Salami, L.A. Leiva, S. Sigg, Pantomime: Mid-air gesture recognition with sparse millimeter-wave radar point clouds, Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 5 (1) (2021) 1–27.

[13] Z. Yang, P.H. Pathak, Y. Zeng, X. Liran, P. Mohapatra, Monitoring vital signs using millimeter wave, in: Proceedings of the 17th ACM International Symposium on Mobile Ad Hoc Networking and Computing, 2016, pp. 211–220.

[14] T. Liu, M. Gao, F. Lin, C. Wang, Z. Ba, J. Han, W. Xu, K. Ren, Wavoice: A noise-resistant multi-modal speech recognition system fusing mmwave and audio signals, in: Proceedings of the 19th ACM Conference on Embedded Networked Sensor Systems, 2021, pp. 97–110.

[15] I. Dua, A.U. Nambi, C. Jawahar, V. Padmanabhan, AutoRate: How attentive is the driver? in: 2019 14th IEEE International Conference on Automatic Face & Gesture Recognition, FG 2019, IEEE, 2019, pp. 1–8.

[16] S. Liu, L. Zhao, X. Yang, Y. Du, M. Li, X. Zhu, Z. Dai, Remote drowsiness detection based on the mmwave FMCW radar, IEEE Sens. J. 22 (15) (2022) 15222–15234.

[17] P. Zhao, C.X. Lu, B. Wang, C. Chen, L. Xie, M. Wang, N. Trigoni, A. Markham, Heart rate sensing with a robot mounted mmwave radar, in: 2020 IEEE International Conference on Robotics and Automation, ICRA, IEEE, 2020, pp. 2812–2818.

[18] H. Jiang, J. Hu, D. Liu, J. Xiong, M. Cai, DriverSonar: Fine-grained dangerous driving detection using active sonar, Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 5 (3) (2021) 1–22.

[19] A. Sen, A. Mandal, P. Karmakar, A. Das, S. Chakraborty, mmDrive: mmWave sensing for live monitoring and on-device inference of dangerous driving, in: 2023 IEEE International Conference on Pervasive Computing and Communications, PerCom, IEEE, 2023, pp. 2–11.

[20] T. Ersal, H.J. Fuller, O. Tsimhoni, J.L. Stein, H.K. Fathy, Model-based analysis and classification of driver distraction under secondary tasks, IEEE Trans. Intell. Transp. Syst. 11 (3) (2010) 692–701.

[21] J. Dai, J. Teng, X. Bai, Z. Shen, D. Xuan, Mobile phone based drunk driving detection, in: 2010 4th International Conference on Pervasive Computing Technologies for Healthcare, IEEE, 2010, pp. 1–8.

[22] B. Cyganek, S. Gruszczyński, Hybrid computer vision system for drivers' eye recognition and fatigue monitoring, Neurocomputing 126 (2014) 78–94.

[23] G. Borghi, M. Venturelli, R. Vezzani, R. Cucchiara, Poseidon: Face-from-depth for driver pose estimation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 4661–4670.

[24] I. Dua, T.A. John, R. Gupta, C. Jawahar, Dgaze: Driver gaze mapping on road, in: 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems, IEEE, 2020, pp. 5946–5953.

[25] S. Kajiwara, Driver-condition detection using a thermal imaging camera and neural networks, Int. J. Automot. Technol. 22 (6) (2021) 1505–1515.

[26] C. Yan, Y. Wang, Z. Zhang, Robust real-time multi-user pupil detection and tracking under various illumination and large-scale head motion, Comput. Vis. Image Underst. 115 (8) (2011) 1223–1238.

[27] T. Hoang Ngan Le, Y. Zheng, C. Zhu, K. Luu, M. Savvides, Multiple scale faster-rcnn approach to driver's cell-phone usage and hands on steering wheel detection, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2016, pp. 46–53.

[28] L. Zhang, K. Liu, X.-Y. Li, C. Liu, X. Ding, Y. Liu, Privacy-friendly photo capturing and sharing system, in: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, 2016, pp. 524–534.

[29] G. Cisotto, A.V. Guglielmi, L. Badia, A. Zanella, Joint compression of EEG and EMG signals for wireless biometrics, in: 2018 IEEE Global Communications Conference, IEEE, 2018, pp. 1–6.

[30] X. Xu, H. Gao, J. Yu, Y. Chen, Y. Zhu, G. Xue, M. Li, ER: Early recognition of inattentive driving leveraging audio devices on smartphones, in: IEEE INFOCOM 2017-IEEE Conference on Computer Communications, IEEE, 2017, pp. 1–9.

[31] Y. Xie, F. Li, Y. Wu, S. Yang, Y. Wang, D 3-guard: Acoustic-based drowsy driving detection using smartphones, in: IEEE INFOCOM 2019-IEEE Conference on Computer Communications, IEEE, 2019, pp. 1225–1233.

[32] Y. Bai, L. Lu, J. Cheng, J. Liu, Y. Chen, J. Yu, Acoustic-based sensing and applications: A survey, Comput. Netw. 181 (2020) 107447.

[33] V. Hromadová, P. Kasák, R. Jarina, P. Brída, Frequency response of smartphones at the upper limit of the audible range, in: 2022 ELEKTRO, IEEE, 2022, pp. 1–5.

[34] H. Jiang, J. Hu, D. Liu, J. Xiong, M. Cai, DriverSonar: Fine-grained dangerous driving detection using active sonar, Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 5 (3) (2021) http://dx.doi.org/10.1145/3478084, URL: https://doi.org/10.1145/3478084.

[35] S. Wang, J. Song, J. Lien, I. Poupyrev, O. Hilliges, Interacting with soli: Exploring fine-grained dynamic gesture recognition in the radio-frequency spectrum, in: Proceedings of the 29th Annual Symposium on User Interface Software and Technology, 2016, pp. 851–860.

[36] H. Liu, A. Zhou, Z. Dong, Y. Sun, J. Zhang, L. Liu, H. Ma, J. Liu, N. Yang, M-gesture: Person-independent real-time in-air gesture recognition using commodity millimeter wave radar, IEEE Internet Things J. 9 (5) (2021) 3397–3415.

[37] G. Tiwari, S. Gupta, An mmwave radar based real-time contactless fitness tracker using deep CNNs, IEEE Sens. J. 21 (15) (2021) 17262–17270.

[38] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, 2014, arXiv preprint arXiv:1409.1556.

[39] W. Koehrsen, Hyperparameter tuning the random forest in python, Towards Data Sci. (2018).

[40] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, L. Fei-Fei, Imagenet: A large-scale hierarchical image database, in: 2009 IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2009, pp. 248–255.

[41] Y. Xie, F. Li, Y. Wu, S. Yang, Y. Wang, Real-time detection for drowsy driving via acoustic sensing on smartphones, IEEE Trans. Mob. Comput. 20 (8) (2020) 2671–2685.