Classification of Driver Head Motions Using a mm-Wave FMCW Radar and Deep Convolutional Neural Network

DREW G. BRESNAHAN®, (Member, IEEE), AND YANG LI®, (Senior Member, IEEE)
Department of Electrical and Computer Engineering, Baylor University, Waco, TX 76798, USA
Corresponding author: Drew G. Bresnahan (drew_bresnahan@baylor.edu)

This work was supported in part by the National Science Foundation under Grant 1609371.
This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Baylor University IRB Committee under IRB No. 463930-11.

ABSTRACT
Eight different driver head movements are measured using a millimeter-wave FMCW radar mounted in the dashboard of a car. The micro-Doppler signatures are converted into a spectrogram image format for analysis and classification purposes. The eight different head motions exhibit unique time-frequency profiles, which can be classified by deep learning algorithms. In this study, a convolutional neural network is used to classify the eight head motions with an optimized window size. Various dataset permutations are considered, such as the effect of window width on classification accuracy and the classification accuracy of head motions in a still car compared to a moving car.

INDEX TERMS Human motion, driver, head, micro-Doppler, spectrogram, millimeter wave, radar, convolutional neural network.

I. INTRODUCTION
Driver drowsiness and distraction is a significant cause of automobile traffic accidents. The United States National Highway Traffic Safety Administration (NHTSA) estimates that drowsiness accounted for approximately 91,000 crashes and at least 795 deaths from road accidents in 2017 [1]. Driver alertness safety features have begun to appear in several car models due to manufacturers realizing their market potential [2], but most of these safety features are driver control input-based and do not accurately monitor the driver’s true condition with high resolution. In contrast, analyzing the motion of the driver’s head allows one to more accurately infer the driver’s drowsiness condition [3], [4].

Previous works have investigated human head movements using a variety of sensing methods in different environments. Wearable sensors have proven viable to monitor head motion [4], [5], but this approach is too cumbersome for most natural driving applications. The two most popular noncontact approaches are camera-based video processing and radar sensing. Smith, et al. [6] laid foundational work in developing a video-fed algorithm to monitor driver head and eye motion. Tawari et al. [7] used a multiple camera vision algorithm to monitor and detect driver behavior in a variety of occlusion and lighting conditions. Additional studies have been conducted to develop models to analyze head motion using video capture methods, as seen in [8]–[10]. While video methods enable a high degree of visual clarity for processing driver gaze, the privacy concern of being recorded is undesirable. Poor lighting conditions and head-worn accessories like sunglasses also limit the usefulness of video processing in more general scenarios.

Radar has proven to be a reliable approach to sensing human head motion. Radar has the advantages of working well regardless of lighting condition or clothes covering the driver. Radar does not have the privacy concerns that a camera-based system does. Chae et al. [11] used a 5.8 GHz Doppler radar to measure six distinct head motions in a laboratory environment to analyze driver head motions. Cardillo et al. [12] used 120 GHz FMCW phase tracking to monitor head and eyelid motions for patients with neurodegenerative disorders. Jung et al. [13] collected radar data of four driver head motions in a stationary car environment and used a convolutional neural network to classify the resulting spectrogram images with over 80% accuracy. Despite these advancements, most previous studies have been...
The goal of this study is to measure and classify eight common head movement patterns exhibited by motor vehicle drivers using a millimeter-wave FMCW radar in a moving car interior environment with passengers sitting nearby. The radar measurement data are processed using a compressed two-dimensional fast-Fourier transform (FFT), creating joint time-frequency visualizations that reveal features of different head motion patterns. These time-frequency images are used to train a deep convolutional neural network (DCNN) to recognize and classify new data from the eight activity classes.

This paper is organized as follows: Section II will detail the radar measurement setup and parameters and illustrate the eight head motions to be considered. Section III will display the spectrogram images containing the human micro-Doppler information and will discuss relevant features for each image. Section IV will present the neural network modeling and classification performance. Section V will conclude the study and discuss future directions.

II. EXPERIMENT SETUP AND DATA COLLECTION

A. RADAR PARAMETERS

In this study, the Texas Instruments AWR1642BOOST [14] and DCA1000EVM [15] evaluation boards are used together to collect raw radar measurements inside a car environment. The FMCW radar can capture both range and velocity information simultaneously on a frame-by-frame basis. For each frame, the radar emits several linear chirps with specific slope and bandwidth, as shown in Table 1. The received chirps are mixed with transmitted chirps to compute the intermediate frequency signal, which is then processed using the FFT to generate the radar range profile. This mm-radar has four receiver antenna elements and can resolve targets in the cross-range dimension as well. A two-dimensional FFT across the multiple chirps allows a range-Doppler map to be created from each frame.

B. EXPERIMENT PROCEDURE

Fig. 1 shows the experiment setup inside a full-sized sedan. The radar assembly is positioned on the dashboard behind the steering wheel facing the driver. This placement allows the radar to capture a wide view of the driver’s range of motion. In the first set of experiments, the car is not moving and there are no passengers present in the car. Fig. 2 shows the range-azimuth profile when the driver is sitting still inside the car. The strong red return at 0.25 m is the steering wheel, the yellow return at 0.8 m is the driver, and the soft blue return at 1.2 m is the roof of the car.

Five driver test subjects perform eight distinct head movements in a repeating manner for 60 seconds each while the radar collects data. Testing takes place in three different vehicle types: a sedan, a small SUV, and a pickup truck. Three subjects test in the sedan, one subject tests in the SUV, and one subject tests in the pickup truck. The period of each repeated activity is controlled and listed below. We do not control the precise movement pattern of each test subject, allowing the subject to act out the motions in a natural way. The slight differences of motion across the subjects enriches our dataset diversity and enhances the classification capability for a general population. These eight head movements are chosen to represent a wide set of common human actions while seated in a car. The head movements are described as follows and are illustrated in Fig. 3.
FIGURE 3. Illustration of eight driver head motions: (a) still, (b) head droop forward, (c) head turn left-right, (d) body turn left-right, (e) chewing, (f) lean back-forth, (g) emergency jolt, and (h) sudden head jerk.

1) Still—the driver head remains upright and motionless.
2) Head Droop Forward—the head falls forward at a moderate speed and then is pulled back upright, simulating drowsiness or falling asleep at the wheel. Period controlled to be about 3 seconds.
3) Head Turn Left-Right—the head turns to look left, then right, continuing to alternate between the positions. The torso remains still. Period controlled to be about 1.5 seconds.
4) Body Turn Left-Right—the head alternates turning to look left and right with the shoulders moving along with the head. Period controlled to be about 2 seconds.
5) Chewing—the jaw moves up and down repeatedly. Period controlled to be about 0.5 seconds.
6) Lean Back-Forth—the driver slowly leans forward with the entire torso and then slowly leans back to the original position. Period controlled to be about 5 seconds.
7) Emergency Jolt—the entire body jumps/convulses with an impulse as if the car has collided with a solid surface. Period controlled to be about 5 seconds.
8) Sudden Head Jerk—a sharp motion throws the head forward and backward, as if a car collision has occurred or the driver is startled while drifting to sleep. Period controlled to be about 4 seconds.

This experiment procedure is repeated twice for two different scenarios: a “still car” case and a “moving car” case. In the “still car” case, the test subject sits alone in the parked car while the engine is off, and he performs the head activities in front of the radar. In the “moving car” case, the test subject is driving the car under 25 miles per hour in an empty parking lot while turning randomly and breaking periodically. Two additional passengers are also present in the passenger seat and the right-back seat. The passengers are instructed not to make large motions but are allowed to move their heads and arms naturally. The “moving car” case simulates a more realistic driving environment with interference and extra vibration, the effects of which are explored in section III.

III. MEASUREMENT RESULTS AND ANALYSIS

A. DATA PROCESSING

The radar data is post-processed using a customized MATLAB script: a two-dimensional FFT is taken along each frame, resolving the range-Doppler profile of the car cabin at each time sample. The range-Doppler frames are then compressed into one unified range bin using a sum operation along the range dimension. Each compressed frame becomes a column of the spectrogram, ordered from left to right to create the slow time axis. A difference filter is applied across the frames to remove the strong spectral component at zero Doppler frequency. The colormap is plotted on a dB scale for visual clarity.

B. SPECTROGRAMS FOR STILL CAR WITH NO PASSENGERS

Fig. 4 displays the spectrograms corresponding to the eight driver head motion activities for the “still car” case. The horizontal axis is time in seconds, and the vertical axis is radial velocity in meters per second. A positive velocity implies forward motion toward the radar, and a negative velocity implies backward motion away from the radar.

Fig. 4(a) shows the still results. This is the baseline for no head motion. A weak return around 0 m/s is visible.

Fig. 4(b) shows the head droop forward results. This motion creates a spectral profile with distinct upward-curve shapes in the positive domain (head moving forward) followed by shorter triangular shapes in the negative domain (head moving backward).

Fig. 4(c) shows the head turn left-right motion, which is characterized by small symmetrical lobes appearing along the center of the plot.

Fig. 4(d) shows body turn left-right, which looks similar to the Head Turn Left-Right results but with larger velocity values and longer duration. The lobes alternate between having a positive and negative velocity skew—this is due to the driver moving his shoulder toward the radar as he looks to the left and then returning to the neutral position before repeating the motion while looking to the right.
Fig. 4(e) shows the *chewing* motion. Chewing is characterized by small, short protrusions from the 0 m/s line that repeat quite rapidly.

Fig. 4(f) shows the *lean back-forth* motion. This motion exhibits slower, sinusoidal shapes in the spectrogram. The velocity does not exceed ±0.5 m/s.

Fig. 4(g) shows the *emergency jolt* motion. This motion generates a tall, sharp, and quick Doppler response in both the positive and negative directions. This is due to the sudden convulsion of the body moving its parts in more than one direction at once. The head moves away from the radar while the torso moves slightly toward the radar. These spikes are followed by a small lobe where the body relaxes and returns to its original position.

Fig. 4(h) shows the *sudden head jerk* motion. This motion’s spectrogram exhibits a sharp positive Doppler response as the head is sharply pulled back up from a slouching position. This motion is similar to *head droop forward* but is more focused on the aggressive snapping-back motion of the head and less focused on the slow, drowsy, forward motion of the head droop.

**C. SPECTROGRAMS FOR MOVING CAR WITH PASSENGERS**

Fig. 5 displays the spectrograms corresponding to the “moving car case.” The test subject performs the eight driver head activities while driving the car approximately 25 miles per hour in an empty parking lot. These spectrograms share many similarities to their “still car” counterparts from Fig. 4. One difference is the extra energy centered about the 0 m/s line caused by the vibration of the moving car. For example, compare Fig. 5(a) to Fig. 4(a). This phenomenon is present on all the Fig. 5 spectrograms. Another difference between these sets of data is the presence of low-velocity energy randomly appearing as time progresses, i.e., Fig. 5(a) at 10 and 15 seconds. This is due to the driver’s hands turning the
steering wheel 7 inches from the radar. However, the action of turning the steering wheel is primarily tangential to the radar, which causes negligible Doppler shift.

Two docile passengers were sitting in the moving vehicle during the measurement trials. The presence of the passengers seems to have little visible effect on the spectrograms. Based on the distinctive traits of the data in Fig. 4 and Fig. 5, we next employ a deep learning algorithm, the deep convolutional neural network, for classification of image data.

IV. NEURAL NETWORK CLASSIFICATION TECHNIQUE

A. NETWORK STRUCTURE

For real-world applications, a vehicle safety system must be able to recognize and classify driver head motions in an accurate and efficient way. Deep learning can accomplish most types of narrow, well-defined classification challenges given a large, diverse set of training data [16]. In this work, a deep convolutional neural network is trained to classify the eight driver head activities given the spectrogram images from section III as training data. Previous works have employed DCNN models for human body motion classification and have shown that well-trained models produce highly accurate and computationally efficient results for spectrogram-based image classification [5], [13], [17]. Fig. 6 illustrates the architecture of the DCNN. This network is implemented using MATLAB® and the Deep Learning toolbox available from Mathworks® [18]. This network is composed of an input layer for raw pixel information, three convolutional layers, three maximum pooling layers to reduce model dimensionality, and a fully connected layer that predicts which of the eight head motion options the input image belongs to.

B. DATASET

The experiment described in section II involving the still car is iterated a total of three times for three different test subjects, collecting three complete data sets for all eight driver head motions where each motion is recorded for 60 seconds. A fourth dataset is collected from one test subject driving a moving car while acting out the eight head motions. In total, there is a combined dataset of four trials with three unique test subjects where each trial contains eight head activity recordings that are each 60 seconds in duration.

To generate a properly large training data pool for the DCNN, each 60-second recording is split into overlapping windows of various widths. The resulting number of unique images per recording using this technique is given as

\[ N = \frac{\text{Time measured} - \text{Window width}}{\text{Step size}} + 1. \]  

With a 4-second window and a step size of 0.5 seconds, each 60-second recording generates 113 spectrogram windows. These values are selected based on heuristic search, and the effect of varying the window size is quantified in section IV-C. The total combined dataset including eight activity recordings for three subjects sums to 4,520 unique images. These data are labeled according to their activity and randomly partitioned into a training set and validation set using a 70-30 split: 70% of the data are used to train the model to recognize the unique features of each activity, and 30% of the data are withheld from the network until the validation step to help prevent overfitting to the training data. The training and validation sets are shuffled every epoch. The network predicts the activity label on each of the validation images to compute the classification accuracy. The model is trained for 30 epochs, with 30 iterations per epoch. Table 2 lists the key parameters for the DCNN model training procedure.

C. DCNN CLASSIFICATION RESULTS

Fig. 7 shows the confusion matrix for the DCNN model for the “still car” case. The average classification accuracy, referring to the mean of the diagonal elements, is 99.36%. The model had the most difficulty recognizing the chewing action, misclassifying it as still 3.8% of the time. This network displays excellent classification performance given the distinct spectrogram patterns for each of the head motions.

| Model Parameter            | Value Used |
|----------------------------|------------|
| Number of Epochs           | 20         |
| Iterations per Epoch       | 30         |
| Learning Rate              | 0.01       |
| Optimizer                  | sgd        |
| Training-Validation Split  | 70%-30%    |

TABLE 2. List of DCNN model parameters.
We repeat the sliding-window dataset generation procedure for the “moving car” data shown in Fig. 5. Fig. 8 shows the confusion matrix for the “moving car” case. The average classification accuracy is slightly lower than the “still car” case, at 94.64%, which can be attributed to the unpredictable hand movements near the steering wheel confusing the DCNN. Nevertheless, this network model still performs well and is expected to improve with additional training data from a multitude of unique test subjects.

Next, we combine the datasets for the “still car” case and the “moving car” case into one dataset to be used to quantify the effect of window size on classification accuracy. We consider window widths of 1, 2, 3, 4, and 5 seconds. A new DCNN model is trained for each window width, and the DCNN parameters are the same as those shown in Table 2. The accuracy vs. window size results are shown in Fig. 9. Evidently, the classification accuracy improves as the window size increases. At smaller windows, such as 1 or 2 seconds, the regions of no motion between the spikes of emergency jolt (Fig. 4(g)) may look like still (Fig. 4(a)) to the network, leading to poor classification accuracy. At larger window sizes of 4 seconds or more, the window captures at least one full period of each repeated motion, resulting in high classification accuracy due to the distinctive features of the full motions. Based on Fig. 9, one can select the appropriate window width depending on the classification accuracy requirement.

The training process using MATLAB takes approximately 1 hour to complete. Once finished, the model can classify new images almost instantly.

D. COMBINED MOTIONS

Finally, real-life driver head monitoring would involve different head motions displayed in an unpredictable order, and it is necessary to continually monitor the driver’s head motions in real-time during the trip. Fig. 10(a) shows a test case where the driver subject chooses to act out several of the eight actions in random order. The actions are performed for a naturally short duration. The first 7 seconds are still, and then we see the small, rapid perturbations of chewing. Next, at 11 seconds, we see the symmetrical lobes characteristic of head turn left-right. At 20 seconds, we find the sinusoidal curve of lean back-forth. Finally, at 26 seconds, we see the impulse-like signature of sudden head jerk. A sliding window is used to segment this 30-second spectrogram into overlapping windows of 4-second width and 0.5-second step size to be compatible with the DCNN model we already trained.
Each of the overlapping spectrogram windows are fed into the pre-trained DCNN model, and the prediction results are shown in Fig. 10(b). We can see that the model correctly predicts most of the activity regions of Fig. 10, starting with still, then transitioning to chewing, followed by head turn left-right, then lean back-forth, and finally sudden head jerk.

The transitions between each main activity region may cause some transient confusion for the network, as seen by sudden deviations to other class predictions such as sudden head jerk at 16 seconds or emergency jolt at 23 seconds. Fig. 11(a) and (b) show another case of random combined head motions and their classification. The classification results in Fig. 11(b) closely match the respective spectrogram patterns but become temporarily confused with spurious predictions around 15 and 20 seconds due to similar head motions so close together.

V. CONCLUSION

This study investigated eight common head motions of a human driver in a realistic motor vehicle environment and computed visualizations of the time-frequency signatures generated from the body interacting with the millimeter-wave radar. We trained a deep convolutional neural network to recognize and classify the activities based on the features exhibited in their respective spectrograms. This work has shown that the high velocity resolution offered by a mm-wave FMCW radar is capable of differentiating many human body motions for practical applications in real-world environments.

Future directions for this research include measuring and classifying driver motions under various real driving scenarios like parking lot, highway cruising, bumpy roads, bad weather, etc. Furthermore, the effects of vigorous passenger motion inside the car while driving need to be studied. The effects of different driver body shapes and sizes may also be explored. Finally, conducting a large-scale measurement campaign with many driver test subjects will enable the subsequent DCNN training to produce robust, high-quality models capable of detecting driver behavior in commercial safety applications. The end goal is to develop practical sensors and edge computing devices to detect head motions and alert the driver in real time.

REFERENCES

[1] Drowsy Driving. Accessed: Feb. 24, 2021. [Online]. Available: https://www.nhtsa.gov/risky-driving/drowsy-driving
[2] Drowsy Driver Detection Systems Sense When You Need a Break. Accessed: Feb. 25, 2020. [Online]. Available: https://www.cars.com/articles/drowsy-driver-detection-systems-sense-when-you-need-a-break-1420684409199/
[3] A. Sathyarayana, S. Nagesswaren, H. Ghasemzadeh, R. Jafari, and J. H. L. Hansen, “Body sensor networks for driver distraction identification,” Proc. IEEE Int. Conf. Veh. Electron. Saf., Columbus, OH, USA, Sep. 2008, pp. 120–125.
[4] E. Murphy-Chutorian and M. M. Trivedi, “Head pose estimation and augmented reality tracking: An integrated system and evaluation for monitoring driver awareness,” IEEE Trans. Intell. Transp. Syst., vol. 11, no. 2, pp. 300–311, Jun. 2010.
[5] D. Bresnahan, Y. Li, S. Koziol, and Y. Kim, “Monitoring human head and neck-based motions from around-neck creeping wave propagations,” IEEE Trans. Intell. Transp. Syst., vol. 11, no. 7, pp. 1199–1203, Jul. 2010.
[6] P. Smith, M. Shah, and N. da Vitoria Lobo, “Monitoring head/eye motion for driver alertness with one camera,” in Proc. 15th Int. Conf. Pattern Recognit., Barcelona, Spain, vol. 4, 2000, pp. 635–642.
[7] A. Tawari, S. Martin, and M. M. Trivedi, “Continuous head movement estimator for driver assistance: Issues, algorithms, and on-road evaluations,” IEEE Trans. Intell. Transp. Syst., vol. 15, no. 2, pp. 818–830, Apr. 2014.
[8] Y. Zhu and K. Fujimura, “Head pose estimation for driver monitoring,” in Proc. IEEE Intell. Veh. Symp., Parma, Italy, Jun. 2004, pp. 501–506.
[9] A. Benoit and A. Caplier, “Biological approach for head motion detection and analysis,” in Proc. 13th Eus. Signal Process. Conf., Antalya, Turkey, 2005, pp. 1–4.
[10] Y. Zhu and K. Fujimura, “Head pose estimation for driver monitoring,” in Proc. IEEE Intell. Veh. Symp., Parma, Italy, Jun. 2004, pp. 501–506.
[11] R. Chae, A. Wang, and C. Li, “FMCW radar driver head motion monitoring based on Doppler spectrogram and range-Doppler evolution,” in Proc. IEEE Topical Conf. Wireless Sensors Sensor Netw. (WiSNet), Orlando, FL, USA, Jan. 2019, pp. 1–4.
[12] E. Cardillo, C. Sapienza, C. Li, and A. Caddemi, “Head motion and eyes blinking detection: A mm-wave radar for assisting people with neurodegenerative disorders,” in Proc. 50th Eur. Microw. Conf. (EuMC), Utrecht, The Netherlands, Jan. 2021, pp. 925–928.
[13] J. Jung, S. Lim, B.-K. Kim, and S. Lee, “CNN-based driver monitoring using millimeter-wave radar sensor,” IEEE Sensors Lett., vol. 5, no. 3, pp. 1–4, Mar. 2021.
[14] AWR1642BOOST. Accessed: Jan. 22, 2022. [Online]. Available: https://www.ti.com/tool/AWR1642BOOST
[15] DCA1000EVM. Accessed: Jan. 22, 2022. [Online]. Available: https://www.ti.com/tool/DCA1000EVM
[16] C. C. Aggarwal, “Teaching deep learners to generalize,” in Neural Networking Deep Learning: A Textbook, 1st ed. Yorktown Heights, NY, USA: Springer, 2018, pp. 162–175.

[17] R. P. Trommel, R. I. A. Harmanny, L. Cifola, and J. N. Driessen, “Multi-target human gait classification using deep convolutional neural networks on micro-Doppler spectrograms,” in Proc. Eur. Radar Conf., London, U.K., 2016, pp. 81–84.

[18] MATLAB for Deep Learning. Accessed: Apr. 15, 2021. [Online]. Available: https://www.mathworks.com/solutions/deep-learning.html?s_tid=srchtitle

DREW G. BRESNAHAN (Member, IEEE) was born in Plano, TX, USA, in 1994. He received the B.S. and M.S. degrees in electrical and computer engineering from Baylor University, in 2016 and 2018, respectively, where he is currently pursuing the Ph.D. degree in electrical and computer engineering.

Since 2014, he has been working with Raytheon Technologies Corporation, Dallas, TX, as a Hardware Test Engineer and an MMIC Design Engineer. His research interests include electromagnetic propagation, machine learning, and analysis of human body interactions with radar.

YANG LI (Senior Member, IEEE) received the B.S. degree in electrical engineering from the University of Science and Technology of China, in 2005, and the M.S. and Ph.D. degrees in electrical and computer engineering from The University of Texas at Austin, in 2007 and 2011, respectively.

In 2011, he was a Faculty Member with Baylor University, where he is currently an Associate Professor of electrical and computer engineering. He has authored 34 journal articles and 57 conference papers and abstracts. His research interests include body area electromagnetic wave propagation and antenna radiation, wearable antenna analysis and design, effective medium theory and homogenization technique, and micro-Doppler radar. He serves as the Technical Program Committee Chair for the 2021 IEEE Texas Symposium on Wireless and Microwave Circuits and Systems.

* * *

VOLUME 9, 2021 100479