Develop algorithms to determine the status of car drivers using built-in accelerometer and GBDT

Article *in* International Journal of Electrical and Computer Engineering · February 2022

DOI: 10.11591/ijece.v12i1.pp785-792

International Journal of Electrical and Computer Engineering (IJECE)
Vol. 12, No. 1, February 2022, pp. 785~792
ISSN: 2088-8708, DOI: 10.11591/ijece.v12i1.pp785-792

Develop algorithms to determine the status of car drivers using built-in accelerometer and GBDT

Thi Thu Nguyen[1], Phuc Thinh Doan[2,3], Anh-Ngoc Le[4], Kolla Bhanu Prakash[5], Subrata Chowdhury[6], Duc-Nghia Tran[7], Duc-Tan Tran[8]
[1]Department Faculty of Electronic Engineering, Hanoi University of Industry, Hanoi, Vietnam
[2]NTT Hi-Tech Institute, Nguyen Tat Thanh University, Ho Chi Minh City, Vietnam
[3]Faculty of Mechanical, Electrical, Electronic and Automotive Engineering, Nguyen Tat Thanh University, Ho Chi Minh, Vietnam
[4]Faculty of Electronics-Telecommunications, Electric Power University, Hanoi, Vietnam
[5]Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation (KLEF), Vaddeswaram, India
[6]Computer Science Department, Sri Venkateswara College of Engineering and Technology (SVCET), Chittoor, India
[7]Institute of Information Technology, Vietnam Academy of Science and Technology, Hanoi, Vietnam
[8]Faculty of Electrical and Electronic Engineering, Phenikaa University, Hanoi, Vietnam

Article Info

Article history:

Received Jan 27, 2021
Revised Jul 15, 2021
Accepted Jul 29, 2021

Keywords:

Acceleration
CarSafe
Gradient booting decision tree
android introduction

ABSTRACT

In this paper, we introduce a mobile application called CarSafe, in which data from the acceleration sensor integrated on smartphones is exploited to come up with an efficient classification algorithm. Two statuses, "Driving" or "Not driving," are monitored in the real-time manner. It enables automatic actions to help the driver safer. Also, from these data, our software can detect the crash situation. The software will then automatically send messages with the user's location to their emergency departments for timely assistance. The application will also issue the same alert if it detects a driver of a vehicle driving too long. The algorithm's quality is assessed through an average accuracy of 96.5%, which is better than the previous work (i.e., 93%).

Corresponding Author:

Duc-Tan Tran
Faculty of Electrical and Electronic Engineering, Phenikaa University
Hanoi, Vietnam
Email: tan.tranduc@phenikaa-uni.edu.vn

1. INTRODUCTION

According to statistics, on an average year, Vietnam has about 8,000 deaths, 15,000 injured due to traffic crashes. Thus, economic losses are estimated at 5 to 12 billion US $. There are many reasons for crashes, such as using the phone to watch movies, listen to music or text, call while driving [1], [2]. Those behaviors are very potential to cause crashes.

CarPlay and Google Android Auto are two intelligent systems of Apple and Android installed on automobiles to use the interface's phone features on the screen by car [3]-[5]. These technologies' ultimate purpose is to help users use the necessary functions on the phone in the most convenient way, limiting manipulation, focusing on the journey. The benefits that Android Auto and Apple CarPlay bring are undeniable. However, these two systems are mostly integrated into modern vehicles. There have also been studies focusing on safe driving in the past. These studies are mainly focused on smartphone services, built-in inertial sensors in the car or on the phone [6], [7]. We want to fulfill other applications' insufficiencies; thus, this paper proposes building software installed on all smartphones running Android operating systems to support automobile drivers. Our contributions can be summarized as follows:

The software automatically recognizes the car's in-car status, automatically turns on Bluetooth to support hands-free calling, and turns off Bluetooth to save the battery if it detects the user is no longer riding again. Some essential features have been added, such as automatically read aloud short message service (SMS messages from Gmail or Messenger. Detect the crash and automatically make emergency contact with the relative person with the user's address. It alerts the driver to erratic driving and drives for too long. Driving for a long time can cause fatigue and accidents.

The paper has four sections. After the introduction, we will present the software system's design model using the machine learning algorithm and the accelerometer sensor built into the phone and presents an algorithm to identify the driver's status in section 2. The results and evaluation of the effectiveness of the proposed system are analyzed in section 3. Finally, the conclusions are presented in section 4.

2. MODEL OF THE SOFTWARE DESIGN SYSTEM

2.1. System design

The software system that identifies cars' driving state using machine learning algorithms and a built-in accelerometer is designed with the functional principle diagram, as shown in Figure 1. If the "Driving" status is identified using a machine learning algorithm, the phone will automatically activate Bluetooth to help the driver make the hands-free conversation mode. It automatically receives messages from SMS, Gmail, or Messenger. Suppose a crash is expected to occur with the driver. In that case, the application will confirm with the user by vibrating vigorously and present a voicemail asking if a crash has occurred or not. It avoids issues of false alerts. In 10 seconds, if there is no response by clicking confirmation on the dialog box from the user, the software will default to a crash and automatically turn on global positioning system (GPS) to get information about the driver's current location. It then automatically send notification messages with GPS coordinates and map links to their emergency departments to receive timely help.

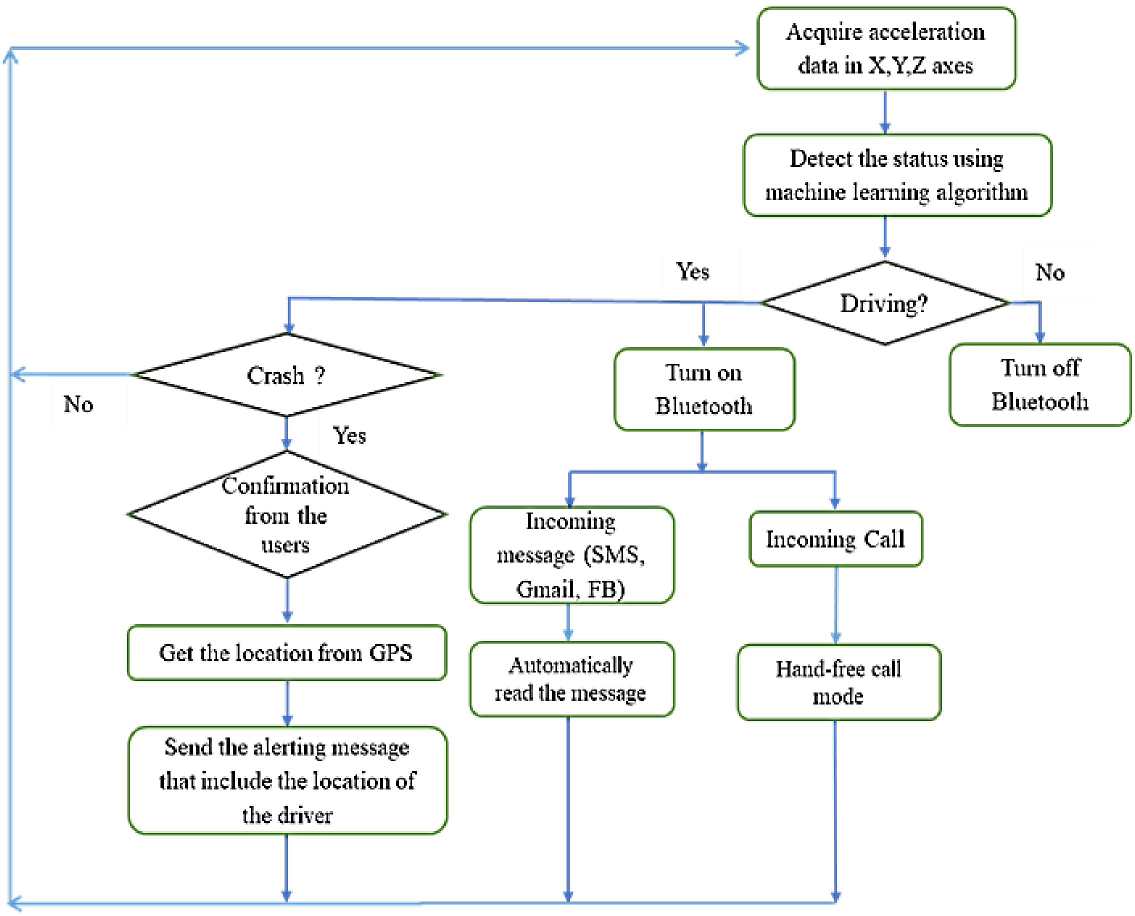


Figure 1. Diagram of the operating principle of the software system

The program can detect the "Driving" status using the acceleration data from the built-in accelerometer. However, we did not provide the raw acceleration data for the machine-learning model. We extract a feature set from the acceleration data. The entire process of determining the status of the driver is shown in Figure 2.

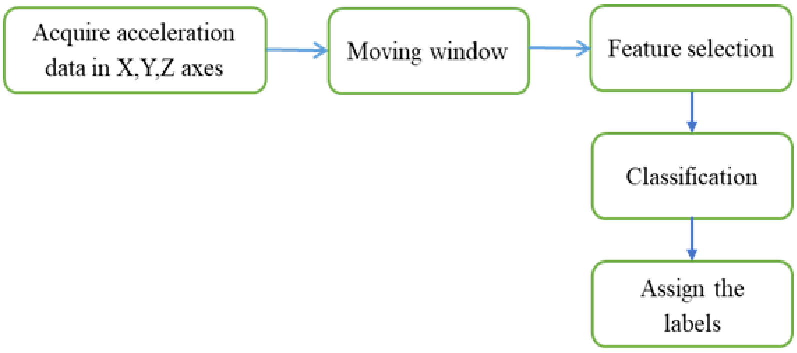


Figure 2. The process of classification of driver status

2.2. Collect data from the accelerometer

The sensor provides 50 samples of the acceleration data every second along the three axes [6]. After conducting data collection from the accelerometer sensor [7], we used a 14-second sliding window to extract the designed features. The monitor data is then labeled in a data set. Table 1 shows an example of observations. The dataset includes 2560 direct observations distributed to the state shown in Table 2.

Table 1. An observation sample along three axes X, Y, and Z

	Driving	
Acceleration in X (g)	Acceleration in Y (g)	Acceleration in Z (g)
- 0.0867	1.2008	9.9403
0.1540	0.7493	9.9128
0.6307	0.1791	9.8924
0.3528	- 0.3694	10.7692
- 0.1083	- 0.2652	10.6853
0.3768	0.2378	9.1378

Table 2. The number of samples observed for each state

Status	Number of samples observed
On Vehicle	581
On Bicycle	308
On Foot	669
Still	899
Tilting	103
Total	2560

2.3. Feature selection

This paper proposes and conducts a feature selection because the classifier can provide better classification results [8]. t-distributed stochastic neighbor embedding (t-SNE) is a non-linear technique to map multidimensional data to lower dimensional space [9]. This study uses the t–SNE technique to map each data point (X, Y, Z) in 3-dimensional space to 2-dimensional space for visualization, easy for data observation. Figure 3 shows the distribution of the training data set without natural selection in 2-dimensional space using the t-SNE technique. The data set does not seem to be stratified when states overlap and overlap. Thus, it is clear that the pretreatment process's influence and natural selection affect the model's classification performance [10], [11].

Figure 4 shows a better result so far in terms of the 5-state classification when each state has been clearly separated, leaving only a small amount of data mixed into other classes. Thus, the characteristic selection was conducted and brought outstanding results to classify five states when selecting four characteristics, including average, median, RMS, and range. Table 3 presents a summary of the chosen features from where $X$is the set of acceleration samples, N is the number of samples, and x_i the value of the i[th] sample in the set X.

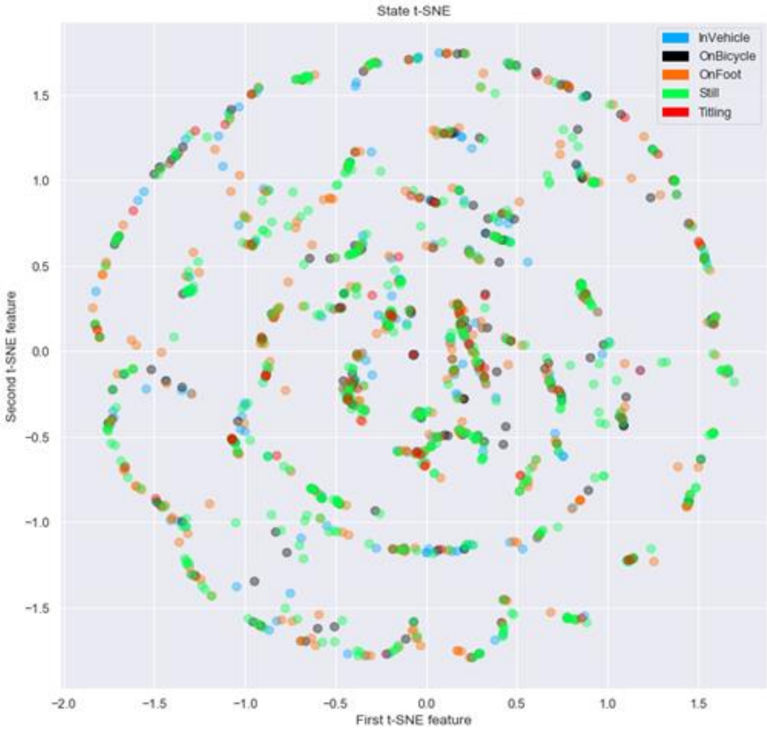


Figure 3. Performing t-SNE collective training data before choosing characteristic

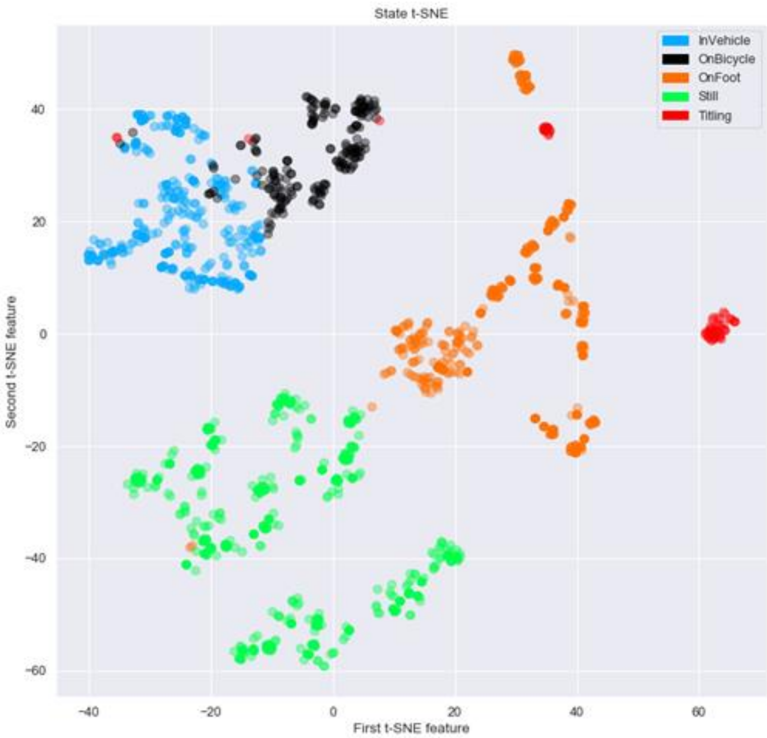


Figure 4. Performing t-SNE using four chosen features

Table 3. The feature formulas

Features	Formula
Mean	Mean $(X) = \frac{1}{N}\sum_{i=1}^{N} x_i$
Median	Median $(X) = \frac{1}{2}\left(x_{\left[\frac{N}{2}\right]} + x_{\left[\frac{N}{2}+1\right]}\right)$
RMS	$RMS_X = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2}$
Range	Range $(X) = [min_{i=1}^{N}\{x_i\}, max_{i=1}^{N}\{x_i\}]$

2.4. State classification algorithm

Referring to the activity classification service Google activity classification [12], we collect five states' data and design a machine learning model to accurately classify those five states from the original dataset, including on vehicle, on bicycle, on foot, still, tilting. The activities are described in detail in Table 4. Four states (on bicycle, on foot, still, tilting) are grouped to a final state, "No driving." From the input featured data, 60% of the data is for training, and the remaining is for testing. These data sets go through the feature selection step to find the best features for the data set. In this study, we used the gradient boosting decision tree (GBDT) method [13]. GBDT can be used in the classification problem due to its strong generalization ability [14], [15].

Table 4. States description

Status based on [12]	Descriptions based on [12]	Status in our study
On vehicle	The equipment is in the vehicle (car, bus, taxi, truck). The user can be a driver or a passenger.	Driving
On bicycle	The user is on a bicycle.	No driving
On foot	The user is lying, running or walking.	
Still	The device is still in the previous state (stay still, do not move)	
Tilting	The angle of the device relative to gravity is much changed. This state usually occurs when picking up the device from the bottom up (from the table up, from the ground up) or in the user's pocket while learning to sit up.	

3. RESULTS AND DISCUSSION

To clarify the proposed software solution's effectiveness in determining the driver's state using data from the accelerometer sensor built into smartphones and the decision tree (DT) machine learning algorithm, we conducted a review. The overall performance of the DT model reaches 95.1% for all classified states. Tables 5 and 6 list the performance parameters for each state. In particular, the states on vehicle, still are classified accurately to a high level. Overall accuracy (accuracy) is suitable for all classes (greater than 90%). The best precision has been achieved with on vehicle, still, and on foot; the remaining states reach the right level with a slightly lower value. Sensitivity is very high, meaning there are not many negative cases that are misclassified as positive. The demo clip of our application is found at the link https://youtu.be/ZbbJ1nmHhvE. It can be seen that our application software can contribute to improving transportation safety [16]-[25].

Table 5. The confusion matrix

Observed States	Predicted States					Total
	On Vehicle	On Bicycle	On Foot	Still	Tilting	
On vehicle	**226**	1	0	0	0	227
On bicycle	5	**115**	36	0	0	156
On foot	0	0	**230**	2	0	232
Still	0	0	0	**356**	0	356
Tilting	0	6	0	0	**40**	46
Total	231	122	266	358	40	1017

Table 6. Algorithm performance evaluation parameters for each state

Status	Algorithm performance evaluation parameters		
	Accuracy	Precision	Recall
On vehicle	99.4%	99.5%	97.8%
On bicycle	95.3%	73.7%	94.3%
On foot	96.2%	99.1%	86.5%
Still	99.8%	100%	99.4%
Tilting	99.4%	87%	100%

3.1. Evaluate the level of energy consumption

Smartphones from Samsung, Huawei, HTC, LG, or Sony all have to face device battery life problems. Android smartphone batteries rarely exceed 36 hours of use. The recent commercial campaigns of Samsung Galaxy show that Samsung has operated in super energy saving mode. When you activate this mode, your phone will change the color screen to a black and white screen. It also limits the number of apps that you can use. In our study, we worked with the phones in Table 7 to evaluate the energy consumption and proposed some solutions as follows:

- Although the information is also extracted from GPS, but only automatically activates this device when a crash occurs; In other normal operations, GPS does not need to be turned on.
- The sensor's sampling rate in the phone can be reduced (while avoiding unnecessary, constant state transitions).
- Users can start and stop using the software manually to save energy.

Table 7. Algorithm performance evaluation parameters for each state

No.	Model	Consumption when using the software (with GPS), calculated by the decreasing percentage per day (%)	Consumption when using the software (not using GPS), calculated by the decreasing rate per day (%)
1	Oppo A37	35.6	25.0
2	Oppo A71	34.7	23.1
3	Oppo A83	32.5	22.4
4	Oppo F5	30.3	21.6
5	Samsung J7	35.3	25.5
6	Samsung J7 Plus	34.5	23.8
7	Samsung A8	32.6	22.7
8	Samsung Galaxy A9 Pro	30.8	21.9
9	Samsung S6	27.3	20.1
10	Samsung S7	26.2	20.3
11	Samsung Galaxy S8	25.4	20.2
12	Xiaomi Note 5A	35.1	25.4
13	Xiaomi 4X	34.0	23.6
14	Xiaomi Note 4	32.3	22.5
15	Xiaomi Mi A1	30.6	21.7
16	Sony Xperia L2	35.7	25.6
17	Sony Xperia M5	34.8	23.8
18	Sony Xperia XA1	32.1	22.9
19	Sony Xperia XZ1	30.2	21.7
20	Lenovo K6 Power	32.3	22.4
21	Lenovo K6 Note	30.5	21.3

4. CONCLUSION

This paper built an Android application to assist the driver of safe cars. We focused on studying the specific features based on the acceleration data. We used the GDBT algorithm to classify data into five states: on bicycle, on foot, still, on vehicle and tilting. Our machine learning model in the proposal makes the overall model accuracy reach 96.5%. This result outperformed the overall accuracy of 93% in the previous work.

REFERENCES

[1] L. T. Truong, L. M. Kieu, and T. A. Vu, "Spatiotemporal and random parameter panel data models of traffic crash fatalities in Vietnam," *Accident Analysis & Prevention,* vol. 94, pp. 153-161, 2016, doi: 10.1016/j.aap.2016.05.028.

[2] Q. N. La, A. H. Lee, L. B. Meuleners, and D. V. Duong, "Prevalence and factors associated with road traffic crash among taxi drivers in Hanoi, Vietnam," *Accident Analysis & Prevention*, vol. 50, pp. 451-455, 2013, doi: 10.1016/j.aap.2012.05.022.

[3] M. Shields, "Assessing cognitive demand testing methods for voice-based infotainment systems," *2020 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*, 2020, pp. 1-2, doi: 10.1109/VL/HCC50065.2020.9127259.

[4] K. T. Nguyen, T. V. Pham, V. D. Nguyen, L. T. Do, A.-V. Tran, and D.-T. Tran, "Development of a smartphone application for safe car driving using google api and built-in sensor," *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 14, no. 2, pp. 178-195, 2020, doi: 10.3991/ijim.v14i02.11118.

[5] D. L. Strayer *et al.,* "Visual and cognitive demands of carplay, android auto, and five native infotainment systems," *Human factors,* vol. 61, no. 8, pp. 1371-1386, 2019, doi: 10.1177/0018720819836575.

[6] T. D. Tan, L. M. Ha, N. T. Long, H. H. Tue, and N. P. Thuy, "Novel MEMS INS/GPS integration scheme using parallel Kalman filters," *2008 IEEE/SICE International Symposium on System Integration*, 2008, pp. 72-76, doi: 10.1109/SI.2008.4770429.

[7] T. D. Tran, D. V. Dao, T. T. Bui, L. T. Nguyen, T. P. Nguyen, and S. Susumu, "Optimum design considerations for a 3-DOF micro accelerometer using nanoscale piezoresistors," *2008 3rd IEEE International Conference on Nano/Micro Engineered and Molecular Systems*, 2008, pp. 770-773, doi: 10.1109/NEMS.2008.4484440.

[8] S. Uyun and E. Sulistyowati, "Feature selection for multiple water quality status: integrated bootstrapping and SMOTE approach in imbalance classes," *International Journal of Electrical and Computer Engineering (IJECE),* vol. 10, no. 4, pp. 4331, 2020, doi: 10.11591/ijece.v10i4.pp4331-4339.

[9] M. Wattenberg, F. Viégas, and I. Johnson, "How to use t-SNE effectively," *Distill*, vol. 1, no. 10, 2016, doi: 10.23915/distill.00002.
[10] A. A.-H. Hassan, W. Md Shah, M. F. I. Othman, and H. A. H. Hassan, "Evaluate the performance of K-Means and the fuzzy C-Means algorithms to formation balanced clusters in wireless sensor networks," *International Journal of Electrical & Computer Engineering (IJECE)*, vol. 10, no. 2, pp. 1515-1523, 2020, doi: 10.11591/ijece.v10i2.pp1515-1523.
[11] P. V. Thanh *et al.*, "Development of a realtime, simple and high-accuracy fall detection system for elderly using 3-DOF accelerometers," *Arabian Journal for Science and Engineering*, vol. 44, no. 4, pp. 3329-3342, 2019, doi: 10.1007/s13369-018-3496-4.
[12] M. Obuchi, W. Sasaki, T. Okoshi, J. Nakazawa, and H. Tokuda, "Investigating interruptibility at activity breakpoints using smartphone activity classification API," *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*, 2016, pp. 1602-1607, doi: 10.1145/2968219.2968556.
[13] J. H. Friedman, "Stochastic Gradient boosting," *Computational Statistics & Data Analysis*, vol. 38, no. 4, pp. 367-378, 2002, doi: 10.1016/S0167-9473(01)00065-2.
[14] R. Sun, G. Wang, W. Zhang, L. T. Hsu, and W. Y. Ochieng, "A gradient boosting decision tree based GPS signal reception classification algorithm," *Applied Soft Computing*, vol. 86, 2020, Art. no. 105942, doi: 10.1016/j.asoc.2019.105942.
[15] J. Hu and J. Min, "Automated detection of driver fatigue based on EEG signals using gradient boosting decision tree model," *Cognitive neurodynamics*, vol. 12, no. 4, pp. 431-440, 2018, doi: 10.1007/s11571-018-9485-1.
[16] R. K. Megalingam, R. N. Nair, and S. M. Prakhya, "Wireless vehicular accident detection and reporting system," *2010 International Conference on Mechanical and Electrical Technology*, Singapore, 2010, pp. 636-640, doi: 10.1109/ICMET.2010.5598437.
[17] J. Zaldivar, C. T. Calafate, J. C. Cano, and P. Manzoni, "Providing accident detection in vehicular networks through OBD-II devices and Android-based smartphones," *2011 IEEE 36th Conference on Local Computer Networks*, Bonn, 2011, pp. 813-819, doi: 10.1109/LCN.2011.6115556.
[18] P. Parmar and A. M. Sapkal, "Real time detection and reporting of vehicle collision," *2017 International Conference on Trends in Electronics and Informatics (ICEI)*, Tirunelveli, 2017, pp. 1029-1034, 2017, doi: 10.1109/ICOEI.2017.8300863.
[19] I. A. Zualkernan, F. Aloul, F. Basheer, G. Khera, and S. Srinivasan, "Intelligent accident detection classification using mobile phones," *2018 International Conference on Information Networking (ICOIN)*, Chiang Mai, pp. 504-509, 2018, doi: 10.1109/ICOIN.2018.8343170.
[20] W. Jules, T. Chris, T. Hamilton, D. Brian, and S. Douglas, "WreckWatch: Automatic Traffic accident detection and notification with smartphones," *Mobile Networks and Applications*, pp. 285-303, 2011, doi: 10.1007/s11036-011-0304-8.
[21] T. V. Nguyen, V. D. Nguyen, T.-T. Nguyen, P. C. Phi Khanh, T.-A. Nguyen, and D.-Ta. Tran, "Motorsafe: An android application for motorbike drivers using decision tree algorithm," *International Journal of Interactive Mobile Technologies (iJIM),* vol. 14, no. 02, pp. 119-129, 2020, doi: 10.3991/ijim.v14i02.10742.
[22] P. V. Thanh, T.-A. Nguyen, N. T. Duc, N. D. Anh, and T. D. Tan, "Development of a real time supported program for motorbike drivers using smartphone built-in sensors," *International Journal of Engineering and Technology (IJET)*, vol. 9, no. 2, pp. 871-881, 2017, doi: 10.21817/ijet/2017/v9i2/170902236.
[23] C. C. Duong *et al.*, "Smartphone-based sensing for intelligent inland waterway transportation," *International Journal of Interactive Mobile Technologies (iJIM),* vol. 14, no. 18, pp. 195-204, 2020, doi: 10.3991/ijim.v14i18.16449.
[24] K. T. Nguyen, Q. T. Vu, V. D. Nguyen, L. Do, and D.-T. Tran, "BoatSafe: A smartphone application for safe inland waterway transportation," *International Journal of Simulation: Systems, Science and Technology*, vol. 20, no. 4, pp. 10.1-10.8, 2019, doi: 10.5013/IJSSST.a.20.04.10.
[25] D. Tian, C. Zhang, X. Duan, and X. Wang, "An automatic car accident detection method based on cooperative vehicle infrastructure systems," *IEEE Access***,** vol. 7, pp. 127453-127463, 2019, doi: 10.1109/ACCESS.2019.2939532.

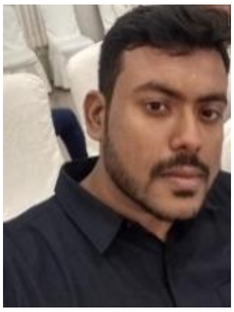
BIOGRAPHIES OF AUTHORS

Thi Thu Nguyen was born in Ha Noi, Vietnam, in 1977. She received her E.Eng from the University of Transport and Communications, Vietnam, in 2000. She received her M.Eng. in 2005 and Dr.Eng. in in 2018. Dr. Thu is now a lecturer at Hanoi University of Industry, Hanoi, Vietnam. She can be contacted at email: thunt@haui.edu.vn.

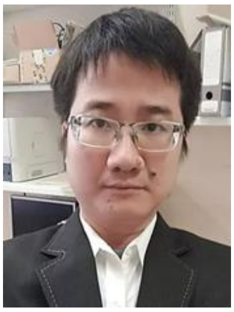
Phuc Thinh Doan received the B.S. degrees in 2007, the M.S. degree in 2011, and the Ph.D. degree from Pukyong National University. Dr. Phuc is now a lecturer at Nguyen Tat Thanh University, Vietnam. He can be contacted at email: dpthinh@ntt.edu.vn.

Anh-Ngoc Le is a Vice Dean of Electronics and Telecommunications Faculty, Electric Power University. His general research interests are embedded and intelligence systems, communication networks, the Internet of things, AI, and Big data analysis. On these topics, he published more than 35 papers in international journals and conference proceedings. He served as a TPC member, session chair, and reviewer of international conferences and journals. He can be contacted at email: anhngoc@epu.edu.vn.

Kolla Bhanu Prakash is working as Professor in the Computer Science Engineering Department, K L University Vijayawada, Andhra Pradesh, India. Dr. Kolla Bhanu Prakash has 13+ years of experience working in academia, research, teaching, academic administration. His current research interests include Media Mining, Image Processing, Data Mining and Natural Language Processing. He was the recipient of Best Speaker award during M.Sc. He has authored over 26 research papers in various national and international journals and conferences. His publications are indexed in Scopus, DBLP, Thomson Reuters. He can be contacted at email: drkbp1981@gmail.com.



Subrata Chowdhury is an Assistant professor in the SVCET Engineering College and Technology Chittoor. He has been the author of books and is the editor for the book series for the reputed international publishers. He has been awarded national and international awards. He is associated with international journals and conferences as the speaker and the reviewer. He has been the guest speaker for many workshops and seminars. He has been the reviewer for many journals. His area of expertise is IoT Healthcare, Blockchain, Machine learning. He can be contacted at email: subrata895@gmail.com.



Duc-Nghia Tran is a scientist at the Institute of Information Technology (IOIT), Vietnam Academy of Science and Technology (VAST). He received a Ph.D. degree from Sorbonne Paris Cité (France). His research interests are mathematics and signal processing, Electron Paramagnetic Resonance (EPR), parameter estimation, data analysis. He can be contacted at email: nghiatd@ioit.ac.vn.

Duc-Tan Tran is an Associate professor and Vice Dean of Faculty of Electrical and Electronic Engineering (FEEE), Phenikaa University. From August 2016 to May 2019, he was an Associate professor and Vice Dean of Electronics and Telecommunication Faculty, VNU University of Engineering and Technology. He has published over 150 research papers. His publications received the "Best Paper Award" at the 9th International Conference on Multimedia and Ubiquitous Engineering (MUE-15), and International Conference on Green and Human Information Technology (ICGHIT-2015). He was the recipient of the award for the excellent young researcher from Vietnam National University in 2008, Hanoi and the third prize in the contest "Vietnamese Talents" in 2008. His main research interests include the representation, processing, analysis, and communication of information embedded in signals and datasets. He serves as a technical committee program member, track chair, session chair and reviewer of many international conferences and journals. He can be contacted at email: tan.tranduc@phenikaa-uni.edu.vn; tantdvnu@gmail.com.