

Vulnerable Road Users Classification of Smartphone Sensors Data Using Deep Transfer Learning

Mohammed Elhenawy¹, Huthaifa I. Ashqar², Mahmoud Masoud^{*1}, Mohammed H. Almannaa³, Andry Rakotonirainy¹, Hesham A. Rakha⁴

¹Centre for Accident Research & Road Safety, Queensland University of Technology, QLD 4059, Australia.

²Booz Allen Hamilton, Washington, D.C., 20003 USA

³King Saud University, Saudi Arabia, Riyadh

⁴Virginia Tech, VA 24061 USA

All authors contributed equally to this work.

*Corresponding author: mahmoud.masoud@qut.edu.au, Cell phone:+6141150363, address : 130 Victoria Park Road, Kelvin Grove QLD 4059, Australia

Abstract

As the Autonomous Vehicle (AV) industry is rapidly advancing, classification of non-motorized (vulnerable) road users (VRUs) becomes essential to ensure their safety and to smooth operation of road applications. The typical practice of non-motorized road users' classification usually takes numerous training time and ignores the temporal evolution and behavior of the signal. In this research effort, we attempt to detect VRUs with high accuracy by proposing a novel framework that includes using Deep Transfer Learning, which saves training time and cost, to classify images constructed from Recurrence Quantification Analysis (RQA) that reflect the temporal dynamics and behavior of the signal. Recurrence Plots (RPs) were constructed from low-power smartphone sensors without using GPS data. The resulted RPs were used as inputs for different pre-trained Convolutional Neural Network (CNN) classifiers including constructing 227×227 images to be used for AlexNet and SqueezeNet; and constructing 224×224 images to be used for VGG16 and

VGG19. Results show that the classification accuracy of Convolutional Neural Network Transfer Learning (CNN-TL) reaches 98.70%, 98.62%, 98.71%, and 98.71% for AlexNet, SqueezeNet, VGG16, and VGG19, respectively. The results of the proposed framework outperform other results in the literature (to the best of our knowledge) and show that using CNN-TL is promising for VRUs classification. Because of its relative straightforwardness, ability to be generalized and transferred, and potential high accuracy, we anticipate that this framework might be able to solve various problems related to signal classification.

Keywords: Transportation Mode Classification; Vulnerable Road Users; Recurrence plots; Computer Vision; Image Classification System.

Introduction

Through the ongoing growth of the automated vehicles (AV) trade, the classification of non-motorized road users such as walkers is becoming crucial in developing safety applications for the Cooperative Intelligent Transportation System (C-ITS) to enhance the safety of non-motorized road users [1, 2]. The C-ITS has been widely investigate and used due to its ability to utilize the data and better manage the transportation networks. The C-ITS attempts to advance health, performance and comfort through various connectivity technologies such as vehicle-to-vehicle (V2V). C-ITS shares various forms of information, including knowledge about non-motorized road users, traffic congestion, incidents and road threats [3]. This helps C-ITS to create an integrated person, route, infrastructure, and vehicle network by implementing communications and other transportation technologies. Taking advantage of the new technology and available big datasets can create a fully functional, instant-time, precise, and secure transport facilities [2, 4]. While C-ITS is now taking attention globally, the academic community focused

primarily on motorized road users such as motor vehicles, expressing less interest for non-motorized road users [5]. Major planning challenges faced by deploying AV is intermodal traffic control, where AV regulations and programming should be structured to value human life by reducing the likelihood of crashes and protecting non-motorized road users [6, 7].

The non-motorized road users are known to be one example of people intervening with AVs “humans that do not explicitly interact with the automated vehicle but still affect how the vehicle accomplishes its task by observing or interfering with the actions of the vehicle” [1]. The non-motorized road users, that is typically have lack of a shield protection, can be described as 'vulnerable'. Those were identified in this research effort by the quantity of traffic safety they lack. The lethality of non-motorized road users, in particular walkers and bikers are greater than the norm. This is because of the discrimination factor for non-motorized road users' collisions is small compared to motorized road users [8].

Conversely, the use of smartphones in data analysis has also lately gained attention of academics and policymakers. The smartphone applications (apps) were designed and used successfully in several fields to gather data from smartphones. Researchers can use smartphones in the transport industry to track and gather motion information such as velocity and motion vector from the integrated Global Positioning System (GPS). This information has the potential to identify the travel mode of the individual, that be used in a variety of different ways and may decrease the amount of time and expense of traditional travel surveys substantially.

A very established practice in non-motorized road users' classification and mode transportation recognition tasks is using the state-of-the-art algorithms for classification by integrating frame-level features over some period as an input. The common approach is to use the traditional statistic techniques such as mean and standard deviation, resulting in a less resolution

dataset and losing very valuable information such as the historical evolution and needs significant training time and expense. In this study, we explore the possibility of using Transfer Learning with Convolutionary Neural Network to classify non-motorized road users with high precision (CNN-TL), which saves training time and cost, to classify images constructed from Recurrence Quantification Analysis (RQA). This approach has the potential to be popular in the transportation mode recognition field due to the potential high accuracy and ease of implementation.

Related work

Scientists have established many methods to effectively differentiate between the modes of transport. Machine learning and artificial intelligence algorithms have shown outstanding performance in creating classification models with high precision, in particular with transportation mode classification. Throughout different experiments, supervised learning models such as Support Vector Machines (SVMs) [9-11], and Random Forests (RFs) [11, 12], Decision Trees [11, 13-18], have all been utilized in different research efforts.

These research efforts have got various levels of accuracy in the classification. There are many variables that influence the precision of detecting modes of transport, as for example the monitoring time, the source of the data, number of modes, and others [12, 19]. A major factor influencing the precision of transportation mode recognition approach, however, is the classifier used in the approach. In most of the research conducted in the past, researchers used only one classification algorithm layer [12, 14, 15]. This is called a conventional framework. On the other side, a few researchers have used more than one classification algorithm layer, which is called , a hierarchical framework [11].

In addition to the number of classification layers, the domain of the extracted features is another important factor that needs to be considered in the transportation mode recognition

approach. The domain of the features can be classified into two categories: time and frequency. Several research efforts were conducted using both as in [9-12, 14, 15, 20] [9, 11, 15]. Both accomplished significant and high accuracy.

Zadeh et al. [21] proposes a geometric approach to detect the risky circumstances that their built-in alert system on smartphones can secure non-motorized road users. This approach can estimate the probability of crash with the use of a fuzzy inference. In addition, a 3D photonic mixer camera was developed to provide pedestrian identification using a sensor device to meet a unique criteria for pedestrian safety in [22]. Anaya et al. [5] have used V2V communications to develop a novel Advanced Driver Assistance Program to prevent collisions between motorcyclists and bikes. A multi-sensor approach was developed to non-motorized road users' security as part of the PROTECTOR project by detecting and identifying non-motorized road users from vehicles in motion in [23]. They explore the impact of using CNN-TL on the precision of the non-motorized road user's classification, which was the first effort to the best of our knowledge in this respect. They aim at precisely detecting non-motorized road users through data obtained from sensors on the smartphone with low power. High level C-ITS protection relies on a specific classification of the non-motorized road users. A binary classifier was introduced to discriminate non-motorized road user's modes (i.e. bicycling, running, and walking) from motorized road user's modes (passenger car and taking bus). A binary classifier is useful in situations where there are higher threats to non-motorized road users. For instance, at an intersection, all subjects' smartphones detect non-motorized road users reports to C-ITS' roadside unit. The C-ITS' roadside also receives messages from vehicles, if any, and then transmits this message onto a warning sign if it detects a potential conflict between non-motorized road users and the vehicle.

We should emphasize that most of the methods proposed in the latest research efforts did

not consider the shortcomings of GPS data such as signal failure or data loss, resulting in unreliable location information. In addition, turning on GPS service in smartphones might quickly drain the battery, thus, this effort attempt to use of collected data of various sensors in a smartphone without GPS information.

Data collection

The dataset was obtained in Blacksburg, VA using a smartphone app by Jahangiri and Rakha [12]. Ten travelers were provided with the app to track their movement in five different modes of transportation, namely: car, bicycle, bus, running, and walking. Data was gathered from four different sensors in the smartphone: The Global Positioning System (GPS), accelerometer, gyroscope, and rotation-vector. Data was warehoused at the maximum viable frequency. Data gathering took place on working days (Mondays to Fridays) and during working hours: from 8:00 AM to 6:00 PM). Several variables have been considered to gather meaningful data that represent natural behaviors. To ensure the sensor positioning has no impact on the data collected, travelers (i.e. participants) were asked to consider holding the smartphone in various positions with no limitations. The data was gathered on various road types, and some periods that indicate congestion conditions that occur in real-life circumstances. The gathering of 30-minute of data per person during the study period was considered appropriate for each mode.

To equate the results of the analysis with results of previous research efforts [11, 12], the selected features extracted from the signals were assumed to have a significant association with the modes of travel for the study. In addition, features that could be derived from the rotation-vector values were omitted or the same purpose. Furthermore, GPS features were ignored in this study allowing this system to be applied in circumstance in which GPS data was unavailable and to relax the issue of battery depleting when GPS service is turned on.

RQA features

Extracting features from the signal is the standard approach in solving mode classification problems in the literature, which can be then used as inputs into the various classification algorithms. The traditional method creates features mainly by using statistics such as the mean, median, and standard deviation values. This process might result in losing the temporal evolution and behavior of the signal, which are valuable information. Extracting features that represent this behavior maximizes the classification precision and accuracy, but it is not yet deeply investigated in the literature. In [24], we proposed extracting features using Recurrence Quantification Analysis (RQA), which we proved that it provided extensive temporal behavior of the obtained signal. RQA is a nonlinear method for analyzing complex dynamic systems by quantifying the recurrence properties of the signal. Eckmann et al. [25] implemented this as a visual tool for finding hidden recurring patterns, un-stationary and systemic shifts. RQA has proved to be a robust method for analyzing dynamic systems, and is capable of quantitatively characterizing the magnitude and complexity of nonlinear, non-stationary and small signals [26-32]. It seems that RQA may result in more subtle-kind of features to the variations in the signal and more robust against the noise in the signal data [30, 31].

In this study we used the extracted features using RQA to create images (we called them RQA images) that could be then used as inputs in a classification algorithm instead of using many numerical features. This has many benefits including the ability of using pretrained deep learning algorithms and representing the various features in one single image, which will save a significant time in computing and reduce the complexity of the system. However, before we introduce the proposed framework, the following is a brief description of how we extracted the features using quantification of patterns that occur in Recurrence Plots (RPs) and more information and details should be found here [24]. Extraction of RQA features involves the setting of three essential parameters: delay (τ) (i.e. lag), phase space dimension (D), and threshold parameter (T). Delay is chosen as the minimum value for the Average Mutual Knowledge (AMI) function. We averaged the collective average information function over all participants and modes in order to calibrate the delay parameter for each channel, as can be seen in

Figure 1 (a). The phase space dimension is calculated using the False Nearest Neighbor (FNN) test, as seen in

Figure 1 (b).

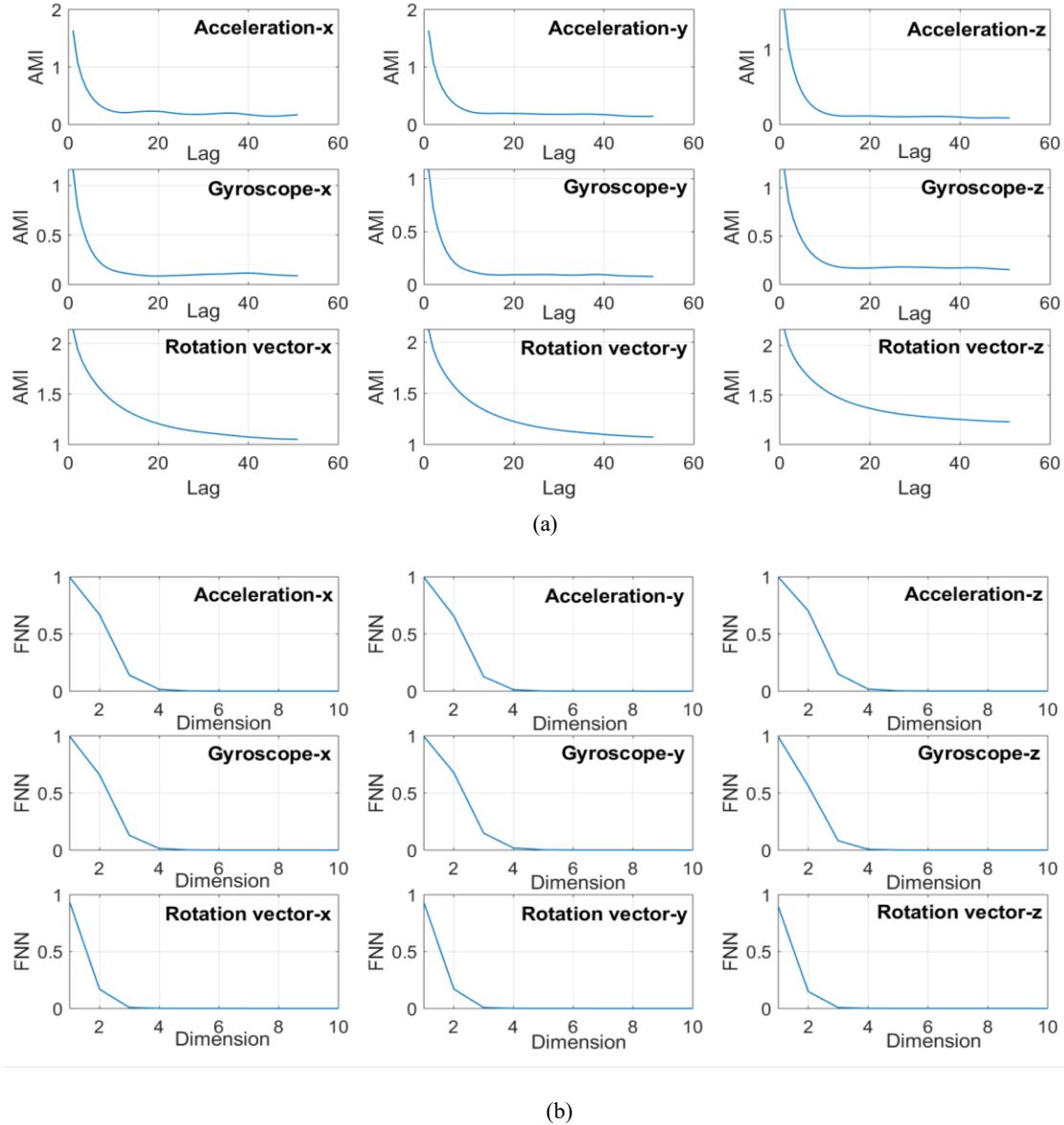


Figure 1: Results of (a) AMI function and (b) FNN test

To calculate the value of T , the space dimension and the delay were used to create the RP and to extract RQA features at various T values. We use the resulted RQA features of each stream from applying RF algorithm as inputs. Consequently, T was calculated for each wave based on the precision of the classification.

Jahangiri and Rakha [12] obtained measurements at a frequency approximately 25 Hz

from the various sensors. As the sensor output samples were not synchronized, a linear interpolation was implemented by the authors to generate continuous signals from the discrete samples. Subsequently, they sampled the designed sensor signals at 100 Hz and divided each sensor's output in each direction (x , y , and z) into 1-s long, non-overlapping windows (t). Using RQA, each point (V_i) in the dimensional space is $V_i = p_i + p_{i+\tau} + p_{i+2\tau} + \dots + p_{i+(D-1)\tau}$, which means that each 1-s window (i.e. 100 sample) results in 70×70 RP for $D = 10$ and $\tau = 4$; and 40×40 RP for $D = 30$ and $\tau = 3$. As a result, six RPs of 70×70 and three RP of 40×40 were extracted to be used in image classification. An example of the extracted RQA images of VRUs and non-VRUs modes are shown in Figure 2.

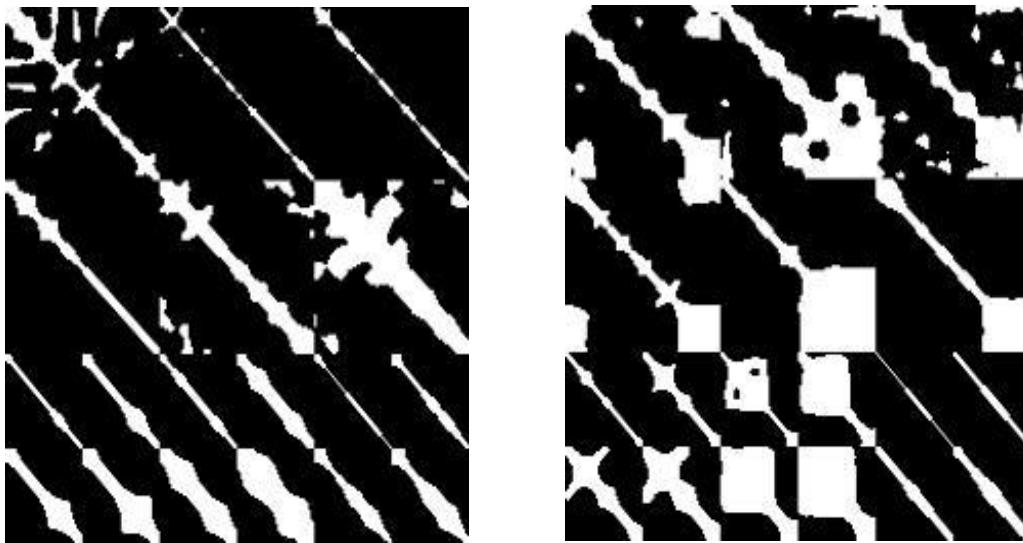


Figure 2: An example of a resulted RQA image of VRU (right) and non-VRU (left)

Methods

Convolutional neural network transfer learning (CNN-TL)

Convolutional neural networks (CNN) is a Deep Learning algorithm that have recently shown outstanding performance in many computer vision applications such as image classification, object classification, and face recognition [33]. In this study, we used CNN as it takes images as inputs, and was proved to be able to process and classify it. Technically, each input image

processes through a series of convolution hidden layers with certain filters to classify it with a defined probabilistic value between 0 and 1. However, because training CNNs needs a relatively huge number of input image data and parameters to be processed, Transfer Learning (TL) was introduced as a pretrained method to expedite training and advance the performance of the CNN models. TL is defined as “a machine learning method where a model developed for a task is reused as the starting point for a model on a second task, which can be used in computer vision and natural language processing aiming to transfer knowledge between related source and target domains” [34, 35]. There are many benefits for using TL, it “overcomes the deficit of training samples for some categories by adapting classifiers trained for other categories and to cope with different data distributions in the source and target domains for the same categories” [33, 34].

In this study, we applied Convolutional Neural Network Transfer Learning (CNN-TL) to classify the resulted RQA images using: 1) AlexNet, which contains five convolutional, three fully connected, max-pooling, and dropout layers [36]; 2) SqueezeNet, which contains two convolution layers, eight Fire Modules, and max-pooling layers [37]; 3) VGG16 and 4) VGG19 [38], both of them contain three convolutional layers, max-pooling, and two fully-connected layers. However, the “16” and “19” stand for the number of weight layers in the network.

Proposed Framework and Results

In order to use CNN-TL for classifying VRU and non-VRU, we proposed the framework shown in Figure 3: . Following to extracting RPs using RQA analysis, we resized and concatenated the nine resulted RPs to construct 227×227 images to be used for AlexNet and SqueezeNet, and 224×244 images to be used for VGG16 and VGG19. For each CNN method, we used 47.5% of the images for training the pre-trained deep neural network using transfer learning, 2.5% of the images were used for validation, and the remaining 50% were used for testing. Consequently, as a key

advantage of the proposed framework, we used RF with varying number of trees from 10 to 200 to capture the temporal dependencies between the consecutive non-overlapping windows (t) of 1-s width and return the probability of a window/image being VRU. As this type of neural network fails to model the time dependency, RF aims to model this temporal relationship using the concatenating VRU probability of consecutive windows to form a vector of probabilities. In this study, we choose 3, 5 and 7 consecutive windows, which corresponds to 3, 5 and 7 seconds, respectively.

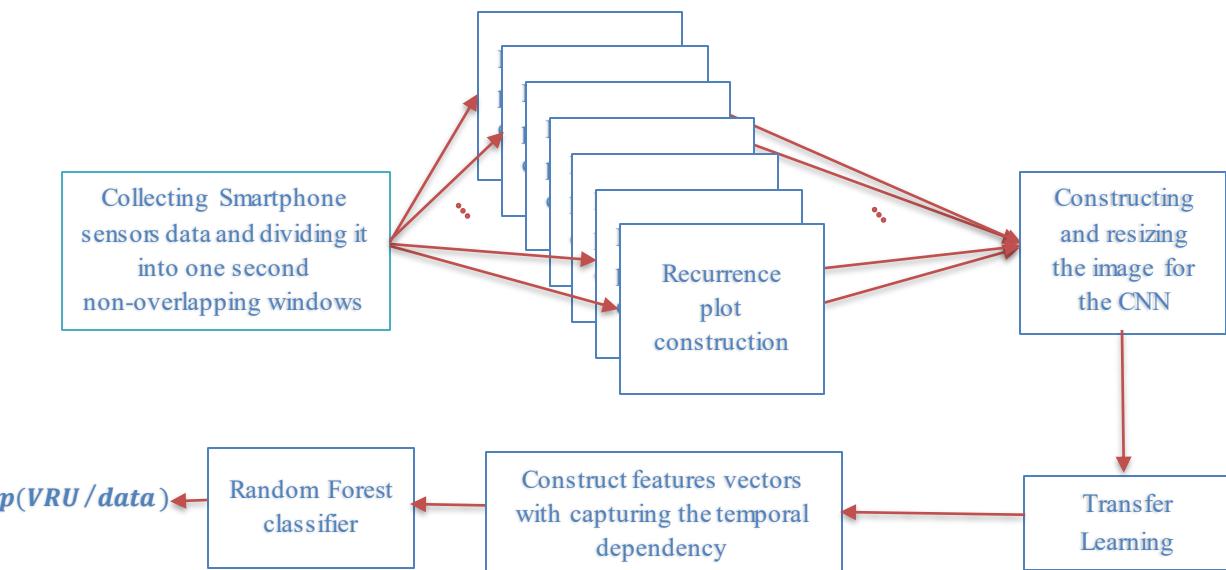


Figure 3: Proposed framework for classifying VRU using CNN-TL and RF

CNN-TL were trained and tested as a binary classifier (i.e. classifying whether the class is a VRU or a non-VRU). RQA images resulted from analyzing data collected using different smartphone sensors, namely: accelerometer, gyroscope, and rotation-vector. As Figure 3 shows, the classification results reaches the highest accuracy of 98.70%, 98.62%, 98.71% and 98.71% using only 7 consecutive non-overlapping windows for AlexNet, SqueezeNet, VGG16, and VGG19, respectively. Figure 3 shows the results of the different CNN-TL methods.

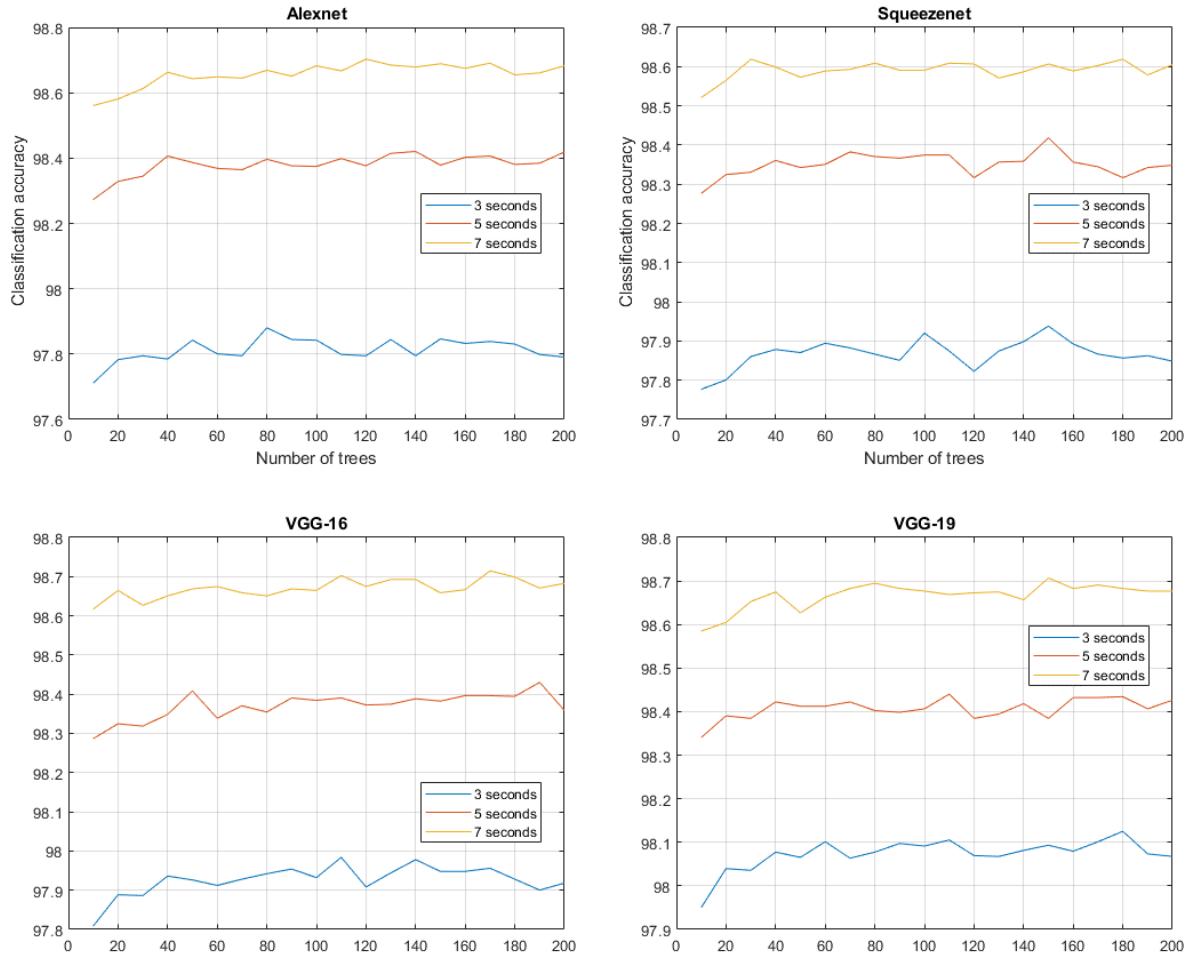


Figure 3: Classification results of VRU and non-VRU using AlexNet, SqueezeNet, VGG16, and VGG19.

Conclusion

As the AV industry is rapidly advancing, non-motorized road users' (i.e. VRUs) classification has become key to enhancing their safety in the road. In this research, by investigating the use of a novel CNN-TL image classification framework, we investigated the impact of extracted RPs, which captures the temporal evolution, on the precision of non-motorized road users' classification. We extracted RPs using data from smartphone sensors like gyroscope, accelerometer, and rotation vector, without GPS data (we assumed they might have some possible

issues such as quickly depleting the smartphone's battery if the service is turned on). We proposed a framework consisting of CNN-TL as a pretrained algorithm to reduce training time and increase the classification accuracy. We also applied the RF algorithm to capture the temporal relationships between non-overlapping windows. We applied different CNNs including AlexNet, SqueezeNet, VGG16, and VGG19. The classification accuracy reached 98.70%, 98.62%, 98.71%, and 98.71% using 7 consecutive windows for AlexNet, SqueezeNet, VGG16, and VGG19, respectively.

Results of the proposed framework proved that the proposed framework is promising, and it outperformed the results in the literature. Our experimental results show that using CNN-TL applied to extracted RQA images has a significant discriminating ability for VRUs classification, which seems to be not captured using other classification algorithms. Unlike other methods, images resulted from RQA would relax the assumptions about linearity, multicollinearity, or stationarity of the data that would be required using other features. Because of its relative straightforwardness, ability to be generalized and transferred, and potential high accuracy, we anticipate that this framework might be able to solve various problems related to signal classification and would become a popular choice in the future.

Data availability

The dataset used to support the findings of this study is owned by Virginia Tech Transportation Institute (VTTI), <https://www.vtti.vt.edu/index.html>, and available upon request.

References

- [1] K. Saleh, M. Hossny, and S. Nahavandi, "Towards trusted autonomous vehicles from vulnerable road users perspective," in *2017 Annual IEEE International Systems Conference (SysCon)*, 2017, pp. 1-7.
- [2] W. Höfs *et al.*, "International deployment of cooperative intelligent transportation systems: bilateral efforts of the European Commission and United States Department of Transportation," U.S. Department of Transportation. Research and Innovative Technology Administration (RITA), 2012.
- [3] M. Alam, J. Ferreira, and J. Fonseca, "Introduction to intelligent transportation systems," pp. 1-17: Springer.
- [4] M. Elhenawy, A. Bond, and A. Rakotonirainy, "C-ITS Safety Evaluation Methodology based on Cooperative Awareness Messages," in *21st International Conference on Intelligent Transportation Systems (ITSC)*, 2018, pp. 2471-2477: IEEE.
- [5] J. J. Anaya, E. Talavera, D. Giménez, N. Gómez, J. Felipe, and J. E. Naranjo, "Vulnerable road users detection using v2x communications," in *IEEE 18th International Conference on Intelligent Transportation Systems*, 2015, pp. 107-112: IEEE.
- [6] T. Litman, *Autonomous vehicle implementation predictions*. Victoria Transport Policy Institute Victoria, Canada, 2017.
- [7] E. Papa and A. Ferreira, "Sustainable accessibility and the implementation of automated vehicles: Identifying critical decisions," *Urban Science*, vol. 2, no. 1, p. 5, 2018.
- [8] SWOV, "SWOV Fact sheet: Vulnerable road users," SWOV Institute for Road Safety Research, Leidschendam, the Netherlands, 2012.
- [9] B. Nham, K. Siangliulue, and S. Yeung, "Predicting mode of transport from iphone accelerometer data," Tech. report, Stanford Univ2008.
- [10] T. Nick, E. Coersmeier, J. Geldmacher, and J. Goetze, "Classifying means of transportation using mobile sensor data," in *Neural Networks (IJCNN), The 2010 International Joint Conference on*, 2010, pp. 1-6: IEEE.
- [11] H. I. Ashqar, M. H. Almannaa, M. Elhenawy, H. A. Rakha, and L. House, "Smartphone Transportation Mode Recognition Using a Hierarchical Machine Learning Classifier and Pooled Features From Time and Frequency Domains," *IEEE Transactions on Intelligent Transportation Systems*, no. 99, pp. 1-9, 2018.
- [12] A. Jahangiri and H. A. Rakha, "Applying machine learning techniques to transportation mode recognition using mobile phone sensor data," *IEEE transactions on intelligent transportation systems*, vol. 16, no. 5, pp. 2406-2417, 2015.
- [13] X. Yu *et al.*, "Transportation activity analysis using smartphones," in *Consumer Communications and Networking Conference (CCNC), 2012 IEEE*, 2012, pp. 60-61.
- [14] L. Stenneth, O. Wolfson, P. S. Yu, and B. Xu, "Transportation mode detection using mobile phones and GIS information," in *19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, United states, 2011, pp. 54-63.
- [15] S. Reddy, M. Mun, J. Burke, D. Estrin, M. Hansen, and M. Srivastava, "Using mobile phones to determine transportation modes," *ACM Transactions on Sensor Networks (TOSN)*, vol. 6, no. 2, p. 13, 2010.
- [16] V. Manzoni, D. Maniloff, K. Kloeckl, and C. Ratti, "Transportation mode identification and real-time CO₂ emission estimation using smartphones," Technical report, Massachusetts Institute of Technology, Cambridge2010.

- [17] Y. Zheng, L. Liu, L. Wang, and X. Xie, "Learning transportation mode from raw gps data for geographic applications on the web," presented at the Proceedings of the 17th international conference on World Wide Web, Beijing, China, 2008.
- [18] P. Widhalm, P. Nitsche, and N. Brandie, "Transport mode detection with realistic Smartphone sensor data," in *2012 21st International Conference on Pattern Recognition (ICPR 2012), 11-15 Nov. 2012*, Piscataway, NJ, USA, 2012, pp. 573-6: IEEE.
- [19] M. Elhenawy, A. Jahangiri, and H. A. Rakha, "Smartphone Transportation Mode Recognition using a Hierarchical Machine Learning Classifier," presented at the 23rd ITS World Congress, Melbourne, Australia, 2016.
- [20] F. Biljecki, H. Ledoux, and P. Van Oosterom, "Transportation mode-based segmentation and classification of movement trajectories," *International Journal of Geographical Information Science*, vol. 27, no. 2, pp. 385-407, 2013.
- [21] R. B. Zadeh, M. Ghatee, and H. R. Eftekhari, "Three-phases smartphone-based warning system to protect vulnerable road users under fuzzy conditions," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 7, pp. 2086-2098, 2018.
- [22] B. Fardi, J. Dousa, G. Wanielik, B. Elias, and A. Barke, "Obstacle detection and pedestrian recognition using a 3D PMD camera," pp. 225-230: IEEE.
- [23] D. M. Gavrila, M. Kunert, and U. Lages, "A multi-sensor approach for the protection of vulnerable traffic participants the PROTECTOR project," in *Proceedings of the 18th IEEE Instrumentation and Measurement Technology Conference*, 2001, vol. 3, pp. 2044-2048: IEEE.
- [24] H. I. Ashqar, M. Elhenawy, M. Masoud, A. Rakotonirainy, and H. A. Rakha, "Vulnerable Road User Detection Using Smartphone Sensors and Recurrence Quantification Analysis," in *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, 2019, pp. 1054-1059.
- [25] J. Eckmann, S. O. Kamphorst, and D. Ruelle, "Recurrence plots of dynamical systems," *World Scientific Series on Nonlinear Science Series A*, vol. 16, pp. 441-446, 1995.
- [26] L. Librizzi, F. Noè, A. Vezzani, M. De Curtis, and T. Ravizza, "Seizure - induced brain - borne inflammation sustains seizure recurrence and blood-brain barrier damage," *Annals of neurology*, vol. 72, no. 1, pp. 82-90, 2012.
- [27] G. Roma, W. Nogueira, P. Herrera, and R. de Boronat, "Recurrence quantification analysis features for auditory scene classification," *IEEE AASP challenge on detection and classification of acoustic scenes and events*, vol. 2, 2013.
- [28] L. Silva, J. R. Vaz, M. A. Castro, P. Serranho, J. Cabri, and P. Pezarat-Correia, "Recurrence quantification analysis and support vector machines for golf handicap and low back pain EMG classification," *Journal of Electromyography and Kinesiology*, vol. 25, no. 4, pp. 637-647, 2015.
- [29] C. L. Webber Jr and N. Marwan, "Recurrence quantification analysis," *Theory and Best Practices*, 2015.
- [30] H. Yang, "Multiscale recurrence quantification analysis of spatial cardiac vectorcardiogram signals," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 2, pp. 339-347, 2011.
- [31] J. P. Zbilut, N. Thomasson, and C. L. Webber, "Recurrence quantification analysis as a tool for nonlinear exploration of nonstationary cardiac signals," *Medical engineering & physics*, vol. 24, no. 1, pp. 53-60, 2002.
- [32] U. R. Acharya, S. V. Sree, G. Swapna, R. J. Martis, and J. S. Suri, "Automated EEG

analysis of epilepsy: a review," *Knowledge-Based Systems*, vol. 45, pp. 147-165, 2013.

[33] H.-C. Shin *et al.*, "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning," vol. 35, no. 5, pp. 1285-1298, 2016.

[34] S. J. Pan and Q. Yang, "A Survey on Transfer Learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345-1359, 2010.

[35] K. Weiss, T. M. Khoshgoftaar, and D. Wang, "A survey of transfer learning," *Journal of Big Data*, vol. 3, no. 1, p. 9, 2016/05/28 2016.

[36] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097-1105.

[37] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 0.5 MB model size," *arXiv preprint arXiv:1602.07360*, 2016.

[38] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv*, 2014.