Transportation Mode Detection Using Cumulative Acoustic Sensing and Analysis

Dinesh Vij[*] and *Naveen Aggarwal*[†]

UIET, Panjab University, Chandigarh
[*]vijdinesh@gmail.com
[†]navagg@gmail.com

1. BASELINE EVALUATION RESULTS

We extracted 13-dimensional (13-d) MFCC features in the preliminary evaluation for capturing the distinct characteristics of different transportation modes. To improve the detection accuracy and robustness, parameter tuning of these features was performed and extracted features were then processed by applying normalization. For modeling the feature vectors using a classifier and performing transportation mode classification, we adhered to a standard SVM-based method. The confusion matrices calculated using MFCC features for module 1 and 2 are presented in Table 1 and 2 respectively. Values given in the matrices denote the percentage of classified test instances.

Table 1: Confusion matrix using MFCC features (module 1).

Transportation Mode/ Transportation Mode	Airplane	On-Road	Train
Airplane	98.25	1.75	0
On-Road	1.75	91.23	7.02
Train	4.39	5.26	90.35

Table 2: Confusion matrix using MFCC features (module 2).

Transportation Mode/ Transportation Mode	Auto-Rickshaw	Bus	Car	Pedestrian
Auto-Rickshaw	81.58	10.53	7.24	0.66
Bus	18.42	75	5.92	0.66
Car	7.24	4.61	86.18	1.97
Pedestrian	3.29	0.66	1.32	94.74

2. WAVELET PACKET TRANSFORM (WPT) FEATURES

We decomposed the wavelet packet tree down to different levels to derive an effective feature set. From the wavelet packets, which are the last level nodes of the full tree of subspaces generated without node selection, we computed log root mean square features. The confusion matrices calculated using WPT for module 1 and 2 are presented in Table 3 and 4 respectively. Values given in the matrices denote the percentage of classified test instances.

Table 3: Confusion matrix using WPT features (module 1).

Transportation Mode/ Transportation Mode	Airplane	On-Road	Train
Airplane	100	0	0
On-Road	0	100	0
Train	0	0	100

Table 4: Confusion matrix using WPT features (module 2).

Transportation Mode/ Transportation Mode	Auto-Rickshaw	Bus	Car	Pedestrian
Auto-Rickshaw	92.76	6.58	0.66	0
Bus	3.95	94.08	1.97	0
Car	1.32	1.32	97.37	0
Pedestrian	0	0	0	100

It is apparent from the matrices that some confusions still remain between *'auto-rickshaw,'* *'bus,'* and *'car'* transportation modes, owing to the resemblance in the acoustic signature of these categories. However, because of the open recording environment and absence of vehicle-specific noise, the *'pedestrian'* transportation mode is recognized with 100% accuracy on using WPT.

3. IMPROVEMENT WITH DEEP LEARNING

To determine the robustness and effectiveness of the proposed deep learning based technique, we tested it on two datasets: 'Normal Dataset' and 'Noisy Dataset.' The results of the evaluation are given in the following sub-sections:

3.1.1. Testing on Normal Dataset

To analyze the evaluation results on the normal dataset in detail, the confusion matrices calculated using DNN for module 1 and 2 are presented in Table 5 and 6 respectively. Values given in the matrices denote the percentage of classified test instances.

Table 5: Confusion matrix using Deep Learning (module 1).

Transportation Mode/ Transportation Mode	Airplane	On-Road	Train
Airplane	100	0	0
On-Road	0	100	0
Train	0	0	100

Table 6: Confusion matrix using Deep Learning (module 2).

Transportation Mode/ Transportation Mode	Auto-Rickshaw	Bus	Car	Pedestrian
Auto-Rickshaw	95.4	2.63	1.97	0
Bus	7.24	92.76	0	0
Car	0.66	0.66	98.02	0.66
Pedestrian	0	0	0.66	99.34

The predictions given in Table 6 show that in several cases, the *'bus'* mode is wrongly predicted as *'auto-rickshaw'* and vice versa, because of the similar acoustic signature of these modes (owing to the similar kind of movement of these public transports on the city roads).

3.1.2. Testing on Noisy Dataset

Afterward evaluating the proposed methods on the normal dataset, we evaluated them on the noisy dataset to validate their robustness and effectiveness. By varying the training set, two sorts of evaluations were carried out on this dataset: (1) classification model trained using the normal dataset, (2) classification model trained using the noisy dataset. Regarding the evaluation (1), testing using the noisy dataset yielded a bit lower classification accuracy of 88.10% (WPT) vs. 96.03% (DNN) for module 1, and 83.04% (WPT) vs. 91.67% (DNN) for module 2. In the evaluation (2), when the classification model was trained using the noisy dataset, testing on the noisy dataset provided a better classification accuracy of 97.62% (WPT) vs. 97.62% (DNN) for module 1, and 92.86% (WPT) vs. 93.75% (DNN) for module 2. The results are shown in Table 7.

Table 7: Testing on Noisy Dataset - Classification accuracy (in %).

Module/ Scenario	Transportation Mode	Model trained using normal dataset		Model trained using noisy dataset	
		Hand-engineered features (WPT + SVM)	Deep Learning	Hand-engineered features (WPT + SVM)	Deep Learning
Module 1	Airplane	96.43	98.81	100	100
	On-Road	71.43	92.86	96.43	96.43
	Train	96.43	96.43	96.43	96.43
	Overall	**88.10**	**96.03**	**97.62**	**97.62**
Module 2	Auto-Rickshaw	88.10	86.90	89.29	92.86
	Bus	82.14	91.67	89.29	85.71
	Car	83.33	96.43	92.86	96.43
	Pedestrian	78.57	91.67	100	100
	Overall	**83.04**	**91.67**	**92.86**	**93.75**

From the rightmost column of the Table, it is evident that in the *'airplane'* mode, there is no effect of speech noise on the mode detection accuracy due to the loud engine and air turbulence noise. Whereas, in *'train'* and *'on-road'* transportation modes, detection accuracy slightly decreases when speech or other unwanted noises are present. Further, for the four subclasses of *'on-road'* transportation mode, it was observed that while commuting in *'auto-rickshaw,' 'bus,'* or *'car,'* detection accuracy slightly decreases in the presence of unwanted noise signals. However, in the *'pedestrian'* mode, speech or other unwanted noises have no effect on the mode detection accuracy, due to the open recording environment.

Furthermore, it was observed that the proposed deep learning based approach has more generalization power than the best performing hand-engineered feature set (WPT). It is so because even when the classification model is trained on a different dataset, testing on the noisy dataset gives far better results when DNN is used instead of WPT features, which demonstrates the effectiveness of the proposed DNN based approach.

4. PERFORMANCE COMPARISON WITH OTHER APPROACHES

An identical dataset to the proposed one could not be established in the literature; therefore our work cannot be directly compared with the previous ones. However, to highlight the detection accuracy achieved by the approaches based on the use of other smartphone sensors, we have added some quantitative results and compared our proposed work with few other approaches. A brief description of those studies and their comparison with the proposed work is provided in the Table 8.

Table 8: Transportation mode detection on mobile phones.

Author [Reference]	**Classes**	**Sensors**	**Detection Accuracy**
Assemi et al. [1]	Walk/run, bicycle, car, bus, train, plane	GPS logs	97%
Hemminki et al. [2]	Stationary, Walk, Bus, Train, Metro, Tram, Car	Accelerometer	84.9%
Reddy et al. [3]	Stationary, walking, running, biking, motorized transport	Accelerometer, GPS	93.6%
Shin et al. [4]	Walk, train/tram, car, bus	Accelerometer, GSM	82%
Fang et al. [5]	Still, walking, running, biking, vehicle	Accelerometer, magnetometer, gyroscope	95%
Garg et al. [6]	Bus, car, bike, auto-rickshaw	Accelerometer, gyroscope, GPS, orientation, light, magnetometer, microphone	92.88%
Singla et al. [7]	Boarding and deboarding in cabs	Barometer	83.33%
Proposed Work	**Module 1:** Airplane, On-Road, Train. **Module 2:** Four sub-categories of On-Road, i.e., Auto-rickshaw, Bus, Car, Pedestrian.	Microphone	**Module 1:** 100% (standard dataset), 97.62% (noisy dataset). **Module 2:** 96.38% (standard dataset), 93.75% (noisy dataset).

REFERENCES

1. Assemi B, Safi H, Mesbah M, Ferreira L. Developing and Validating a Statistical Model for Travel Mode Identification on Smartphones. IEEE Trans Intell Transp Syst 2016; 17(7):1920– 1931. doi:10.1109/TITS.2016.2516252.

2. Hemminki S, Nurmi P, Tarkoma S. Accelerometer-based Transportation Mode Detection on Smartphones. Proc. 11th ACM Conf. Embed. Networked Sens. Syst. (SenSys'13), New York, NY, USA: ACM; 2013, p. 13:1–13:14. doi:10.1145/2517351.2517367.
3. Reddy S, Mun M, Burke J, Estrin D, Hansen M, Srivastava M. Using Mobile Phones to Determine Transportation Modes. ACM Trans Sen Netw 2010; 6(2):13:1–13:27. doi:10.1145/1689239.1689243.
4. Shin D, Aliaga D, Tunçer B, Arisona SM, Kim S, Zünd D, Schmitt G. Urban sensing: Using smartphones for transportation mode classification. Comput Environ Urban Syst 2015; 53:76–86. doi:https://doi.org/10.1016/j.compenvurbsys.2014.07.011.
5. Fang S, Fei Y, Xu Z, Tsao Y. Learning Transportation Modes From Smartphone Sensors Based on Deep Neural Network. IEEE Sens J 2017; 17(18):6111–6118. doi:10.1109/JSEN.2017.2737825.
6. Garg S, Singh P, Ramanathan P, Sen R. VividhaVahana: Smartphone Based Vehicle Classification and Its Applications in Developing Region. Proc. 11th Int. Conf. Mob. Ubiquitous Syst.: Comput. Netw. Serv. (MOBIQUITOUS'14), ICST, Brussels, Belgium: ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering); 2014, p. 364–373. doi:10.4108/icst.mobiquitous.2014.257982.
7. Singla S, Jain S, Aggarwal N, Raman B, Bansal D, Ramakrishnan KK. BaroTrack: Low Cost Tracking of Commuter on Road. Proc. 7th ACM Int. Work. Mobility, Interf. Middlew. Manag. HetNets (MobiMWareHN'17), New York, NY, USA: ACM; 2017, p. 2:1–2:6. doi:10.1145/3083201.3083202.