Ultrasonic-Based Transportation Mode Detection in Urban Environments

Abhay Joshi, Sai Thejeshwar Sharma, Christian Gentner
Institute of Communications and Navigation, *German Aerospace Center (DLR)*, 82234 Weßling, Germany
{abhay.joshi, sai.sharma, christian.gentner}@dlr.de

***Abstract*—Transportation Mode Detection (TMD) plays a key role in enabling intelligent transportation systems, optimizing mobility services, and supporting energy-efficient urban planning. However, methods that utilize Global Navigation Satellite System (GNSS) signals typically suffer from performance degradation in urban canyons and tunnels due to satellite occlusion and multipath interference, while Inertial Measurement Unit (IMU)-based methods are prone to cumulative errors from sensor drift. Hence, we propose in this paper, a novel ultrasonic sensing framework for classifying urban transportation modes based on vehicle-borne acoustic emissions in the ultrasonic range. To the best of the authors' knowledge, this is the first study to analyze ultrasonic sound for transport mode identification. Using a Pettersson u384 microphone sampling at 384 kHz, we captured audio within the 20–80 kHz frequency band across four representative modes, bus, streetcar, subway, and suburban railway, during several hours of real-world operation in Munich. Each transportation mode exhibits a distinct acoustic signature that can be leveraged to differentiate between modes. We employ a stacked ensemble learning approach combining Random Forest (RF), Support Vector Machines (SVM), Logistic Regression (LR), and Multi-Layer Perceptron (MLP), with a LR meta-learner integrating the predictions. Our method achieves an accuracy of 99.57% in distinguishing between modes. This ultrasonic-based approach provides a robust, privacy-preserving alternative that complements traditional sensing modalities in context-aware mobility systems.**

I. Introduction

Transportation Mode Detection (TMD) plays a critical role in smart city infrastructure, navigation systems, and mobility analytic. By accurately determining whether a person is walking, cycling, driving, or using public transportation, optimizing traffic management, improving route planning, enabling seamless e-ticketing, and enhancing overall urban mobility is possible. While Global Navigation Satellite Systems (GNSSs) provide precise positional data under ideal conditions, the reliability significantly deteriorates in dense urban environments due to signal obstructions and multipath effects, complicating the distinction between similar transportation modes. Furthermore, accurate transport mode identification can improve positioning performance, as knowledge of the travel mode enables more appropriate adjustments within localization algorithms.

Conventional approaches to TMD primarily rely on smartphone-based motion sensors [1], GNSS trajectories [2], cell tower metadata [3], or audible acoustic signatures [4], [5] for mode classification. Among these, motion sensors are often preferred due to their lower energy consumption compared to continuous GNSS sampling.

Accelerometer-based methods are highly sensitive to device placement and user interaction, which can obscure the underlying vehicle dynamics, particularly during low-acceleration or stationary phases, resulting in misclassification errors [6], [7]. While GNSS trajectory analysis provides valuable spatiotemporal features, its effectiveness relies on uninterrupted satellite visibility, which is frequently compromised in urban canyons and unavailable in underground transportation systems. Even when augmented with Geographic Information System (GIS)-based proximity features, GNSS trajectory classifiers struggle to distinguish between vehicular modes with similar spatial patterns [8]. Cell-tower–based methods suffer from limited spatial granularity and irregular sampling intervals due to event-driven handoff mechanisms, which hinder reliable mode inference. Audible acoustic-based TMD is similarly constrained: the high levels of ambient urban noise often mask the spectral features required for accurate vehicle classification, reducing the performance of audio-based models [9]. Furthermore, continuous recording within the speech frequency band raises privacy concerns and may be subject to regulatory constraints. While Richoz et al. [10] achieved improved classification accuracy by fusing inertial, audio, and visual modalities, their approach remains restricted to the audible range, and multimodal fusion lies beyond the scope of this study.

Within this study, we introduced an ultrasonic sensing framework that classifies transportation modes based on the unique acoustic signatures of different vehicles in the ultrasonic frequency range (20–80 kHz). According to the inverse-square law of sound propagation and the frequency-squared dependence of atmospheric absorption [11], ultrasonic energy levels typically observed in public transport cabins, measured at approximately 80 dB Sound Pressure Level (SPL) [12], [13], attenuate to the 50 dB detection threshold within a distance of roughly 5 to 9 meters across the 20–80 kHz frequency band. This propagation behavior localizes ultrasonic signatures to the host vehicle, naturally filtering out distant acoustic interference.

While airborne ultrasound surveys have been conducted in urban environments, including public transportation facilities [14], [15], the frequency band above 20 kHz remains unexplored mainly for TMD. This range offers relative immunity to urban background noise. It excludes intelligible speech, resulting in a cleaner and more privacy-preserving acoustic channel that, to the best of our knowledge, has yet to be

systematically leveraged for transportation mode classification.

Modern mobile device microphones are typically limited to sampling rates of 44.1–48 kHz, constrained by operating system audio application programming interface (API) that restrict access to true ultrasonic content (>20 kHz). Nevertheless, several research prototypes have demonstrated that near-ultrasonic frequencies can effectively support ranging and proximity-based services on standard smartphones [16], [17]. Dedicated microelectromechanical systems (MEMS) microphones, such as the Knowles/Syntiant SPH0641LU4H-1 and STMicroelectronics IMP23ABSU, offer flat frequency responses extending to 80 kHz, well beyond the audible spectrum. As future handset designs integrate and expose these wideband sensors to developers, native ultrasonic sensing is expected to become a viable option for TMD.

Hence, this paper presents a classification approach that leverages vehicle-generated signatures in the ultrasonic frequency band to distinguish among four prevalent urban transportation modes: bus, suburban railway, subway, and streetcar. To the best of the authors' knowledge, this is the first study to utilize ultrasonic sound for transport mode identification. The proposed method provides a complementary sensing modality for robust mode detection in dense or subterranean environments, where traditional techniques such as GNSS or inertial sensing may be unreliable. Ultrasonic audio data were collected using a Pettersson u384 microphone within the public transport system of Munich, focusing specifically on four modes: bus, streetcar, subway, and suburban railway. Distinct mode-specific patterns are also clearly visible through inspection of the signal in the frequency domain, supporting the feasibility of using ultrasonic characteristics for classification. We analyzed the data using a machine learning approach, employing a stacked ensemble model that combined Random Forest (RF), Support Vector Machines (SVM), Logistic Regression (LR), and Multi-Layer Perceptron (MLP) classifiers. This approach achieved a classification accuracy of 99.57%, demonstrating that ultrasonic sensing is a robust and privacy-preserving alternative for transportation mode detection.

The remainder of this paper is organized as follows. Section II details the ultrasonic sensing platform, data acquisition procedures, signal processing pipeline, and feature extraction techniques. Section III analyses the recorded ultrasonic data, identifies mode-specific acoustic patterns, and evaluates the classification performance and robustness of the proposed machine learning models. Finally, Section IV summarises the findings and outlines potential directions for future research.

II. Experimental Setup and Methodology

The main benchmark datasets for TMD, such as the Sussex–Huawei Locomotion dataset [18], Microsoft GeoLife [19], and the audible-band DCASE challenges, contain inertial sensor data, GNSS information, or conventional audio streams. However, they do not encompass ultrasonic recordings. To address this gap, we have compiled a dedicated urban transportation dataset within the 20–80 kHz frequency range.

Fig. 1 presents the data processing pipeline utilized to convert raw ultrasonic recordings into predictions of transport modes. Each stage of this process is elaborated upon in detail in Sections II-A through II-D.

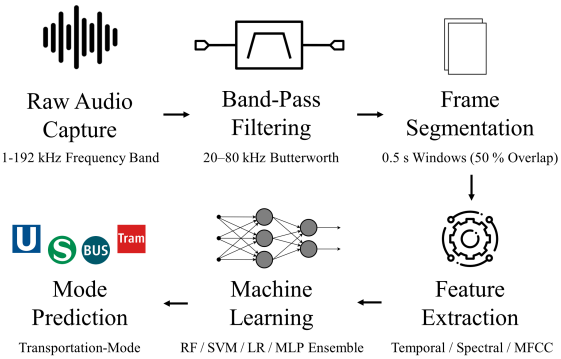


Fig. 1: Data processing pipeline: from ultrasonic audio to transportation mode prediction

A. Data Acquisition and Constraints

We recorded ultrasonic data using the Pettersson u384 Universal Serial Bus (USB) Type-C microphone, which integrates a MEMS-based microphone, anti-aliasing and processing circuitry, providing a calibrated passband of 1–192 kHz when operated at a 384 kHz sampling rate in 16-bit mono configuration. This sampling rate corresponds to a Nyquist frequency of 192kHz, nearly an order of magnitude greater than the 44.1/48 kHz limits typical of consumer-grade audio hardware, ensuring high-fidelity acquisition across the full ultrasonic spectrum. Audio samples were streamed via a USB connection to a laptop capable of sustaining the data throughput and storage demands of continuous field recording. Notably, the u384 is compliant with the USB Audio Class standard and can also interface with smartphones via On-The-Go (OTG)-enabled USB connections.

Recordings were captured in the ultrasonic range (20–80 kHz), focusing on the spectrum where meaningful urban acoustic signals are present. This range reflects both environmental factors, signal energy above 80 kHz is typically negligible outdoors, and the inherent limitations of MEMS microphones, which exhibit significant sensitivity roll-off beyond this point [14]. The 80 kHz upper limit also reflects physical constraints of MEMS microphones. These sensors typically include a Helmholtz resonator structure, with resonance in the lower ultrasonic range. Beyond this, frequency response drops by 15–20 dB per octave due to diaphragm compliance and port mass. Even ultrasonic-optimized MEMS models maintain flat responses only up to 80 kHz, after which sensitivity declines sharply. At these frequencies, the acoustic wavelength ($\lambda = c/f \approx 4$ mm at 80 kHz) approaches the microphone's aperture size, leading to additional attenuation from diffraction and viscous boundary effects.

The dataset includes several hours of ultrasonic audio, segmented into 30-minute sessions across four main urban transportation modes: bus, streetcar, subway, and suburban

railway. Data were collected on various routes and at different times of day, covering underground, elevated, and surface segments, as well as diverse road conditions. Each session is timestamped and labeled with transportation mode, route, and operational context, making the dataset well-suited for robust classification tasks in real-world urban environments. The resulting dataset contains a class distribution that reflects typical urban transportation patterns, with suburban railway representing 18.82%, bus 19.95%, subway 18.21%, and streetcar 19.59%, along with silent background (only representative) comprising 23.42% of the total samples.

B. Signal Processing Pipeline

We analyzed the Signals in their unmodified form and the preprocessing step applied an eighth-order zero-phase Butterworth bandpass filter with cutoff frequencies at 20–80 kHz. This filtering effectively attenuates low-frequency urban noise (e.g., speech, engine hum) and high-frequency artifacts, thereby isolating the ultrasonic band relevant for analysis and classification.

We extracted Spectral features using Short-Time Fourier Transform (STFT) with 5 ms Hamming windows and 90% overlap. A minimum Fast Fourier Transform (FFT) length of 8192 points was used to ensure high spectral resolution, enabling the separation of closely spaced frequency components that are characteristic of different mechanical systems. This configuration supports the visualization of transient events that vary across transportation modes.

Temporal segmentation was implemented using a sliding window approach, dividing each recording into 0.5-second frames with 50% overlap, corresponding to a 0.25-second advancement. This configuration offers an optimal balance between temporal granularity and spectral stability, capturing fast acoustic transitions while preserving sufficient context for robust frequency-domain analysis. At a sampling rate of 384 kHz, each frame contains 192,000 samples, ensuring statistical reliability for downstream feature extraction.

To characterize time-averaged spectral energy distributions, Power Spectral Density (PSD) was estimated using Welch's method with 1-second windows and 50% overlap. These PSD estimates provided noise-robust spectral representations by averaging out transient fluctuations. Empirically, many of the discriminative features for TMD were prominently visible in the PSD plots. Spectrograms were rendered using logarithmic magnitude scaling and normalized colormaps to emphasize dynamic spectral content. These visualizations served as qualitative diagnostics and informed manual inspection of signatures.

C. Feature Extraction

To maintain consistency with the signal processing stage (cf. Section II-B), each 0.5-second frame (192,000 samples) was filtered using an eighth-order zero-phase Butterworth bandpass filter spanning 20-80 kHz. Feature extraction was performed on each filtered frame, resulting in a 135-dimensional feature vector that includes both handcrafted features and those derived from ultrasonic Mel-Frequency Cepstral Coefficients (MFCC). These features are organized into five categories:

- **Temporal Features**: These features capture the dynamics of energy and the shape of the signal through statistical and energy-based descriptors.
- **Spectral Features**: They represent aspects of the spectral shape, roll-off, and concentration within the ultrasonic range.
- **Ultrasonic Band Features**: These quantify the band-wise spectral power across the 20–80 kHz range using fixed 10 kHz intervals.
- **MFCCs and Their Derivatives**: These features describe the ultrasonic spectral envelope utilizing MFCCs and their temporal derivatives.
- **Mel Band Energies**: They capture the energy distribution across the subbands of the ultrasonic Mel scale.

The MFCCs were calculated using a 13-coefficient Mel filter bank with 24 bands, covering the frequency range of 20–80 kHz. Table I provides a summary of all feature categories: the *Sub-features* column details each type of measurement, while the *Count* column presents the total number of feature dimensions for each category.

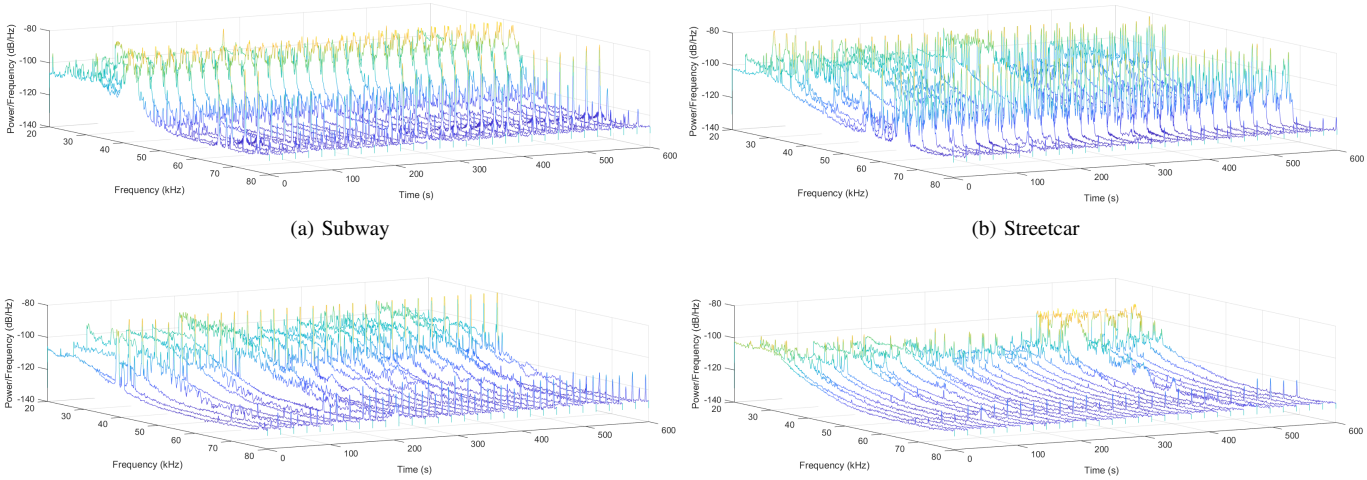
TABLE I: Overview of Extracted Features by Category

Category	**Sub-features**	**Count**
Temporal Features	Mean, Std, Skewness, Kurtosis, Root Mean Square (RMS), Peak Amplitude, Crest Factor, Zero-Crossing Rate (ZCR), Energy, Abs Mean, Sub-window RMS/Peak	18
Spectral Features	Spectral Centroid, Bandwidth, Flatness, Roll-off (25%, 50%, 75%, 90%), Peak Frequency, Peak Magnitude	9
Ultrasonic-Band Features	Energy in 6 frequency bands (20–30 kHz to 70–80 kHz)	6
MFCC and Derivatives	MFCCs, Delta, Delta-Delta (each with Mean and Std)	78
Mel Band Energies	Mean energy across 24 ultrasonic Mel bands	24

To prepare the data for classification, we implemented a series of preprocessing steps. All features underwent standardization through z-score normalization, which adjusted each feature to have a mean of zero and a unit variance. This normalization significantly reduces the impact of varying feature scales, an essential consideration for regularized and distance-based models. To ensure representative class distributions, the dataset was stratified by transportation mode prior to splitting. An 80%–20% division was employed to separate the data into training and test sets. The resulting feature matrix and labels were then utilized as input for the classification framework described in the subsequent section.

D. Classification Framework

We implemented a modular classification pipeline in Python using `scikit-learn`, evaluating five models trained on the extracted ultrasonic features: SVM, LR, RF, MLP, and a stacking ensemble.



(a) Subway (b) Streetcar

(c) Bus (d) Suburban Railway

Fig. 2: PSD (20–80 kHz) visualizations over a 10-minute window for four urban transportation modes: (a) subway, (b) streetcar, (c) bus and (d) suburban railway. Distinct spectral signatures are observable for each mode, reflecting characteristic ultrasonic emission patterns.

To optimize model performance, each base classifier underwent a grid search over predefined hyperparameter spaces, utilizing 5-fold stratified cross-validation. For instance, the SVM was tuned over kernel type, regularization strength, and gamma scale; LR employed L2 regularization with the LBFGS solver; and the RF classifier was fine-tuned for the number of estimators, tree depth, and split constraints. The MLP was configured with two hidden layers (100 and 50 units), Rectified Linear Unit (ReLU) activation, and adaptive learning rate scheduling.

In addition to individual models, we evaluated a stacking ensemble that combined the four base classifiers. A LR meta-classifier was trained on their outputs to learn a higher-level decision function, enhancing generalization by leveraging complementary prediction patterns from the base learners. Thereby, model selection was driven by cross-validation accuracy and validated through overall generalization performance on a held-out test set. The final evaluation metrics encompasses overall accuracy, class-wise performance as indicated by confusion matrices, and classification reports.

III. Results and Analysis

The following analysis presents key findings from the ultrasonic recordings, beginning with spectral distinctions across transportation modes and followed by classification results.

A. Spectral Characteristics

We measured baseline ambient noise in a quiet outdoor setting and observed a flat PSD across the 20-80 kHz range, indicating minimal environmental or electronic interference. The only notable deviation was a rare, low-amplitude peak near 30 kHz, matching the microphone's intrinsic mechanical resonance.

Fig. 2 illustrates the averaged power spectral density (20–80 kHz) over a five-minute interval for four urban transportation modes. The x-axis represents frequency in kHz, while the y-axis indicates PSD in dB/Hz (as detailed in Section II-B). Each subplot features a unique ultrasonic signature, summarized as follows:

- **Subway** (Fig. 2a): Denoted by a wideband comb of narrow, evenly spaced peaks centered around 25 kHz and 31–33 kHz, along with a significant characteristic energy band between 40–45 kHz. Additional spectral elements include a comb near 50 kHz and a strong resonant peak around 68–70 kHz, suggesting the presence of multiple periodic mechanical sources.
- **Streetcar** (Fig. 2b): Exhibits modest peaks near 25 kHz, scattered spikes between 30–40 kHz, and a characteristic energy band between 52–60 kHz. Weaker harmonics are also observed around 65–68 kHz, indicative of potential secondary resonances.
- **Bus** (Fig. 2c): Displays intermittent narrow spikes at 30 kHz, a characteristic energy band between 37–40 kHz, and low-amplitude scattered peaks in the 75–80 kHz range, creating a distinctive ultrasonic profile.
- **Suburban railway** (Fig. 2d): Shows a sparse distribution of low-amplitude comb harmonics, with localized peaks between 32–34 kHz and 48–50 kHz. The overall flatter spectral signature indicates a quieter operational mode with minimal high-frequency mechanical emissions.

Temporal consistency analysis showed that the spectral patterns remained relatively stable across recordings, even with variations in speed, load, and environment. This stability underscores their reliability for classifying transport modes in real-world scenarios.

B. Classification Performance

Table II provides the per-class precision, recall, and F1-score for our stacking-ensemble model. Overall, the classifier

achieves an accuracy of 99.57%, indicating robust performance across all transportation categories.

TABLE II: Stacking Ensemble Model Classification Result

Class	Precision	Recall	F1-Score
Bus	1.00	0.99	1.00
Suburban Railway	0.99	0.99	0.99
Silent	0.99	0.99	0.99
Streetcar	1.00	1.00	1.00
Subway	1.00	1.00	1.00
Accuracy			1.00
Macro avg	1.00	1.00	1.00
Weighted avg	1.00	1.00	1.00

Fig. 3 presents the normalized confusion matrix, where nearly all of the samples fall along the diagonal. Only 0.6% of the Bus samples are misclassified as Suburban Railway, and 0.3% of Suburban Railway samples are misidentified as Bus, indicating a very low incidence of cross-class errors.

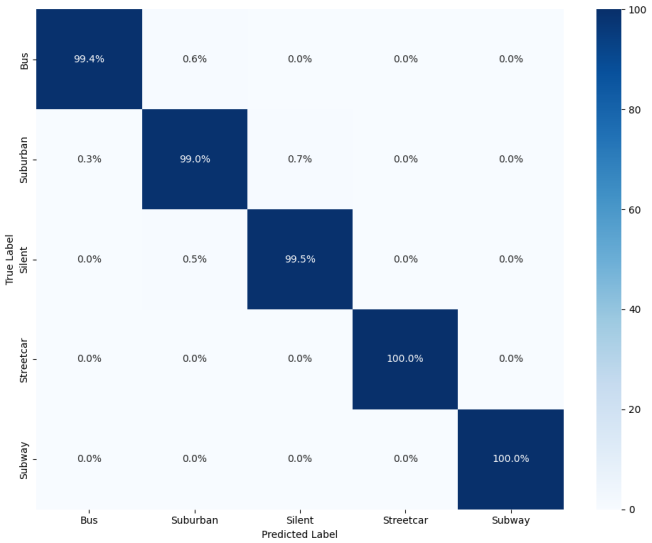


Fig. 3: Confusion Matrix of the Stacking Ensemble Model

C. Evaluation Trip

We conducted an evaluation trip covering 35.79 km over approximately 78 minutes. The multimodal route included the following consecutive segments: a 14-minute bus ride, a 5-minute transfer to the suburban railway, a 34-minute suburban railway segment, a 6-minute transfer to the subway, a 7-minute subway ride, a 9-minute transfer to the streetcar, and a final 3-minute streetcar ride. Fig. 4 shows the complete trajectory, rendered in gpx.studio atop OpenStreetMap, with each transportation mode marked in a distinct color. Please note that for the evaluation trip, only the actual public transport rides were considered for classification; transfer segments were excluded. This study focuses on the detection of TMD. Detecting transition phases between modes may require additional sensors or specialized algorithms, which are beyond the scope of this work.

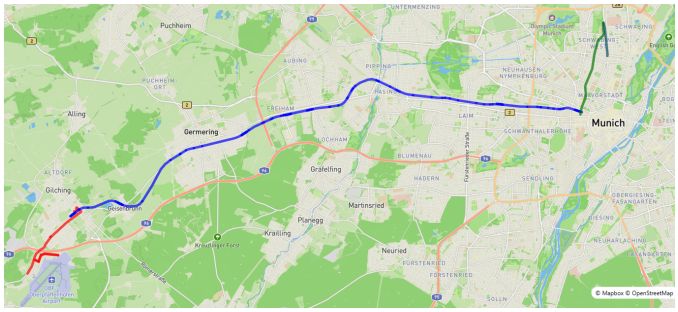


Fig. 4: Multimodal trajectory (35.79 km) in Munich, color-coded by transport mode: bus (red), suburban railway (blue), subway (green), and streetcar (teal).

We perform inference with a simple voting mechanism to smooth the predicted transport mode labels. This involves taking a vote on consecutive 10 frames i.e., 5 second window at each time step to determine the mode. We thereby balance sensitivity to genuine mode transitions against suppression of spurious fluctuations, yielding more consistent sequential inferences than fixed-window smoothing.

The results are illustrated in Fig. 5, which compares the ground truth labels with the detected labels following the application of the adaptive smoothing filter. We achieved an overall classification accuracy of 97%, with per-mode accuracy of 83.6% for bus, 100% for suburban railway, 99.8% for subway, and 98.9% for streetcar. Notably, we observed that the residual misclassifications within the bus category are primarily assigned to the suburban railway class, likely due to the varying acoustic signatures of different bus vehicles. Incorporating data from complementary sensors could further reduce these fluctuations and enhance the robustness of the classification.

IV. Conclusion and Future Work

In this work, we introduced a novel framework for transportation mode detection using vehicle-borne ultrasonic emissions within the 20–80 kHz frequency range. By analyzing distinct acoustic signatures captured from real-world operations in Munich across four representative urban modes, i.e. bus, streetcar, subway, and suburban railway. We demonstrated that ultrasonic sensing can effectively distinguish between transportation modes with high reliability. Leveraging a stacked ensemble learning approach, our method achieved a classification accuracy of 99.57%, underscoring the discriminative power of ultrasonic features and their potential as a robust, privacy-preserving alternative to traditional GNSS- and Inertial Measurement Unit (IMU)-based systems.

Our findings highlight the feasibility and effectiveness of using ultrasonic audio as a complementary sensing modality in intelligent transportation systems. However, future work must address challenges related to sensor variability and environmental influences. Specifically, differences in ultrasonic amplitude across vehicle generations and the non-linear frequency responses of MEMS microphones necessitate robust

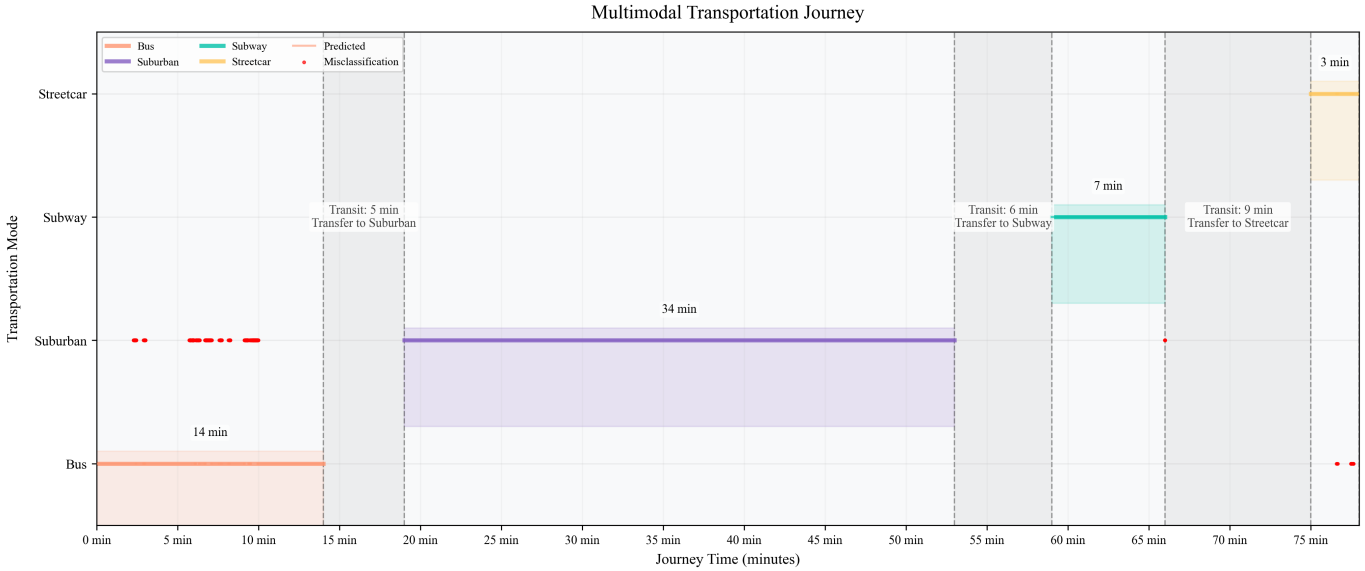


Fig. 5: Comparison of ground truth and detected transportation modes after simple smoothing.

calibration and equalization strategies. Initial tests with alternative MEMS sensors like the IMP23ABSU further underline the importance of careful sensor selection to maintain data consistency and performance.

References

[1] L. Wang, M. Ciliberto, H. Gjoreski, P. Lago, K. Murao, T. Okita, and D. Roggen, "Summary of shl challenge 2024: Motion sensor-based locomotion and transportation mode recognition in missing data scenarios," in *Companion of the 2024 on ACM International Joint Conference on Pervasive and Ubiquitous Computing*, ser. UbiComp '24. New York, NY, USA: Association for Computing Machinery, 2024, p. 555–562. [Online]. Available: https://doi.org/10.1145/3675094.3678456

[2] T. Fourez, N. Verstaevel, F. Migeon, F. Schettini, and F. Amblard, "Transport mode detection on gnss and accelerometer data: a temporality based workflow," *Transportation Research Procedia*, vol. 82, pp. 3454–3471, 2025, world Conference on Transport Research - WCTR 2023 Montreal 17-21 July 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352146524003818

[3] S. Mostafa, M. Youssef, and K. A. Harras, "Modesense: Ubiquitous and accurate transportation mode detection using serving cell tower information," in *Proceedings of the 32nd ACM International Conference on Advances in Geographic Information Systems*, ser. SIGSPATIAL '24. New York, NY, USA: Association for Computing Machinery, 2024, p. 184–195. [Online]. Available: https://doi.org/10.1145/3678717.3691250

[4] S. Lee, J. Lee, and K. Lee, "Vehiclesense: A reliable sound-based transportation mode recognition system for smartphones," in *2017 IEEE 18th International Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, 2017, pp. 1–9.

[5] ——, "Deepvehiclesense: An energy-efficient transportation mode recognition leveraging staged deep learning over sound samples," *IEEE Transactions on Mobile Computing*, vol. 22, no. 6, pp. 3270–3286, 2023.

[6] S. Hemminki, P. Nurmi, and S. Tarkoma, "Accelerometer-based transportation mode detection on smartphones," in *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems*, ser. SenSys '13. New York, NY, USA: Association for Computing Machinery, 2013. [Online]. Available: https://doi.org/10.1145/2517351.2517367

[7] A. Karite and C. Gentner, "Real-time detection of transport modes and movement states via smartphone data," in *Accepted: IEEE/ION Position, Location and Navigation Symposium (PLANS) 2025*, 2025.

[8] J. Li, X. Pei, X. Wang, D. Yao, Y. Zhang, and Y. Yue, "Transportation mode identification with gps trajectory data and gis information," *Tsinghua Science and Technology*, vol. 26, no. 4, pp. 403–416, 2021.

[9] L. Wang and D. Roggen, "Sound-based transportation mode recognition with smartphones," in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019, pp. 930–934.

[10] S. Richoz, L. Wang, P. Birch, and D. Roggen, "Transportation mode recognition fusing wearable motion, sound, and vision sensors," *IEEE Sensors Journal*, vol. 20, no. 16, pp. 9314–9328, 2020.

[11] H. E. Bass, H. Bauer, and L. B. Evans, "Atmospheric absorption of sound: Analytical expressions," *The Journal of the Acoustical Society of America*, vol. 52, no. 3B, pp. 821–825, 09 1972. [Online]. Available: https://doi.org/10.1121/1.1913183

[12] A. HARDY, "Measurement and assessment of noise within passenger trains," *Journal of Sound and Vibration*, vol. 231, no. 3, pp. 819–829, 2000. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0022460X99925658

[13] N. K. Mondal, M. Dey, and J. K. Datta, "Vulnerability of bus and truck drivers affected from vehicle engine noise," *International Journal of Sustainable Built Environment*, vol. 3, no. 2, pp. 199–206, 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2212609014000508

[14] M. Grimshaw-Aagaard and B. Bemman, "Ultrasonics and urban greening: an exploratory study on ultrasound presence in urban spaces," *Personal Ubiquitous Comput.*, vol. 28, no. 5, p. 677–692, Apr. 2024. [Online]. Available: https://doi.org/10.1007/s00779-024-01798-5

[15] T. G. Leighton, "Are some people suffering as a result of increasing mass exposure of the public to ultrasound in air?" Jan 2016. [Online]. Available: https://10.1098/rspa.2015.0624

[16] P. Lazik, N. Rajagopal, B. Sinopoli, and A. Rowe, "Ultrasonic time synchronization and ranging on smartphones," in *21st IEEE Real-Time and Embedded Technology and Applications Symposium*, 2015, pp. 108–118.

[17] J. Meklenburg, M. Specter, M. Wentz, H. Balakrishnan, A. Chandrakasan, J. Cohn, G. Hatke, L. Ivers, R. Rivest, G. J. Sussman, and D. Weitzner, "Sonicpact: An ultrasonic ranging method for the private automated contact tracing (pact) protocol," 2020. [Online]. Available: https://arxiv.org/abs/2012.04770

[18] L. Wang, H. Gjoreski, M. Ciliberto, S. Mekki, S. Valentin, and D. Roggen, "Enabling reproducible research in sensor-based transportation mode recognition with the sussex-huawei dataset," *IEEE Access*, vol. 7, pp. 10 870–10 891, 2019.

[19] Y. Zheng, H. Fu, X. Xie, W.-Y. Ma, and Q. Li, *Geolife GPS trajectory dataset - User Guide*, geolife gps trajectories 1.1 ed., July 2011. [Online]. Available: https://www.microsoft.com/en-us/research/publication/geolife-gps-trajectory-dataset-user-guide/