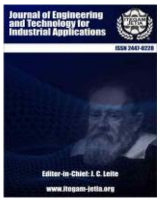
Journal of Engineering and Technology for Industrial Applications

ITEGAM-JETIA

ISSN ONLINE: 2447-0228

Manaus, v.11n.55, p.48-58. September/October, 2025.

DOI: https://doi.org/10.5935/jetia.v11i55.1866



RESEARCH ARTICLE

OPEN ACCESS

REDUCING COMPLEXITY IN VEHICLE RECOGNITION: A COMPARISON OF DIMENSIONALITY REDUCTION METHODS FOR ACOUSTIC FEATURES

Yudhi Darmawan[1], Muladi[2], Heru Wahyu Herwanto[3] and Nur Rachman Supadmana Muda[4]

[1,2,3] Department of Electrical Engineering and Informatics, Universitas Negeri malang, Indonesia.
[4] Politeknik Angkatan Darat, Indonesia.

[1]https://orcid.org/0009-0001-5063-6045, [2]https://orcid.org/0000-0002-2904-5398
[3]https://orcid.org/0009-0000-0796-3596, [4]https://orcid.org/0000-0002-9421-7432

Email: yudhi.darmawan.2305349@students.um.ac.id, muladi@um.ac.id, heru_wh@um.ac.id, nurrudal@gmail.com

ARTICLE INFO

Article History
Received: April 15, 2025
Revised: May 20, 2025
Accepted: September 15, 2025
Published: September 30, 2025

Keywords:
Dimensionality Reduction,
Vehicle Recognition,
Acoustic Features,
Intelligent Transportation Systems.

ABSTRACT

Reducing the complexity of high-dimensional acoustic data is essential for effective vehicle recognition, especially in intelligent transportation systems. This study evaluated six dimension reduction techniques, including Principal Component Analysis (PCA), Kernel PCA, Incremental PCA, Independent Component Analysis (ICA), Truncated Singular Value Decomposition (SVD), and Latent Dirichlet Allocation (LDA), to address the challenges of data redundancy while maintaining relevant features. The dataset includes acoustic signals from seven categories of vehicles extracted using Mel-Frequency Cepstral Coefficients (MFCC), Spectral Centroid, and Spectral Bandwidth. Incremental PCA showed the highest accuracy (0.982) on scenarios with larger training datasets, with effective management of high-dimensional data. ICAs provide optimal performance with fewer components at a higher proportion of test data, demonstrating their efficiency in retaining information. SVD shows stability across all data ratios, confirming its reliability for a wide range of applications. Although LDAs maintain competitive results, their interpretability stands out in certain tasks. These findings emphasize the importance of selecting appropriate dimension reduction methods based on data characteristics and application needs, providing valuable insights to improve the accuracy and efficiency of vehicle recognition systems.

I. INTRODUCTION

The classification of acoustics-based vehicles is increasingly relevant in various applications, such as intelligent transportation, traffic management, and autonomous vehicle systems. Vehicle acoustic analysis offers the advantage of being a non-invasive approach that can be accessed in real-time, even in challenging environmental conditions [1], [2]. Compared to image-based methods, this approach is more efficient for traffic surveillance that requires continuous detection and high reliability [3], [4]. Various methods have been developed to process vehicle acoustic signals, including the use of Mel-Frequency Cepstral Coefficients (MFCC) as a key feature representing the spectral characteristics of the data [5], [6]. In addition, psychoacoustic indicators, such as the Zero Crossing Signature (ZCS), have been used to improve accuracy in vehicle fuel type classification [7], [8]. On the other hand, deep learning models such as CNNs and LSTMs have managed to capture complex non-linear patterns of vehicle signals, which are difficult for conventional techniques to handle [9], [10], [11]. Distributed acoustic sensors have also been implemented to support real-time detection and classification of vehicles with a high degree of accuracy [2], [12], [13].

Previous research has shown great potential for acoustic analysis in vehicle classification. Hybrid models that combine MFCC with Attention BiLSTM have been shown to improve accuracy in emergency vehicles [3]. In addition, spectrogram-based spectral techniques are used to distinguish vehicle types through traffic noise [5], while acoustics- and radar-based multimodal sensors provide richer data for analysis [13], [14]. Acoustic simulation frameworks also play an important role in accelerating model training without relying on large datasets [15], [16]. While promising, these methods often face challenges due to the high dimensions of the data, which can add to the redundancy of information and hinder the efficiency of analysis. The high dimensions of vehicle acoustic data are one of the main challenges, as the redundancy of the resulting information can increase the risk of overfitting and computational complexity [17], [18]. Synthetic data simulation-based approaches have helped improve the training efficiency of acoustic models, but still face limitations in application to real environments [15], [16]. In addition, deep learning models often require large datasets to achieve optimal performance, which is a constraint in certain applications [19], [20]. This problem is further exacerbated when vehicle acoustic data is not only high-dimensional but also adds complexity to the classification process. This condition can hinder accuracy and efficiency, so solutions are needed to reduce data complexity without losing important information.

In this study, we propose the approach of Reducing Complexity in Vehicle Recognition: A Comparison of Dimensionality Reduction Methods for Acoustic Features to overcome this problem. By comparing various dimension reduction methods such as Principal Component Analysis (PCA), Independent Component Analysis (ICA), Kernel PCA, and other techniques, this study aims to reduce information redundancy while maintaining the most relevant features for vehicle recognition. This approach also evaluates the effectiveness of dimension reduction methods in improving computational efficiency and model accuracy, which is relevant for applications in intelligent transportation systems.

II. THEORETICAL REFERENCE

The introduction of acoustics-based vehicles has attracted widespread attention in the development of intelligent transportation systems and modern traffic surveillance. Acoustic data offers a non-invasive approach that can be used in a variety of environmental conditions, including low-visibility situations. This approach is particularly useful for real-time vehicle detection, especially in scenarios that require continuous surveillance with high reliability. By analyzing the unique acoustic patterns of different types of vehicles, acoustics-based methods have shown great potential in improving the efficiency and accuracy of transportation systems. However, acoustics-based analytics face a major challenge in the form of high data dimensions. Vehicle acoustic data includes a wide range of spectral and temporal features designed to capture vehicle sound characteristics. These high dimensions often result in information redundancy that not only increases the computational load but also increases the risk of overfitting in machine learning models. This is a significant challenge, especially when the dataset used is not large enough to adequately represent the diversity of data, thus reducing the model's ability to generalize under real conditions.

In an effort to overcome these challenges, previous research has developed various data processing techniques. For example, deep learning models such as LSTMs are used to capture non-linear patterns in vehicle acoustic signals [6], [10]. In addition, synthetic data simulations have been applied to improve the efficiency of model training in cases where the original dataset is limited. Nonetheless, these methods often do not fully solve the problem of data redundancy, which remains a major challenge in the introduction of acoustics-based vehicles. The high dimensions of data also have implications for computing efficiency, especially in the development of real-time systems. Big data analysis requires significant computing resources, which are often not ideal for practical applications such as traffic surveillance or autonomous vehicle systems. Thus, an approach is needed that not only simplifies the data structure but also retains relevant information for vehicle recognition. This highlights the importance of dimension reduction techniques as a first step in improving the efficiency of machine learning models for vehicle acoustic analysis.

In this context, our research aims to explore the effectiveness of various dimension reduction methods in addressing the challenges of high-dimensional data. By understanding the role of dimension reduction in filtering relevant information, this study aims to provide a framework that supports the development of more efficient and accurate intelligent transportation systems. Thus, this research is expected to make an important contribution in addressing the technical challenges associated with the introduction of acoustics-based vehicles.

III. MATERIALS AND METHODS

III.1 DATASET

The dataset used in this study consists of two main sources, namely the Vehicle Sounds Dataset [21] and Emergency Vehicle Siren Sounds [22], obtained from Kaggle. These two datasets were chosen because they cover various types of vehicles that are often encountered in transportation and emergency systems. Using acoustic data, this study aims to explore the sound characteristics of vehicles that can be used in a voice-based classification system. This dataset includes seven categories of vehicles, namely ambulances, bics, buses, firetrucks, motorcycles, trains, and trucks, with a varying number of samples for each category. Ambulances, buses, and firetrucks each had 400 sound samples, while the bics category had 300 samples. Meanwhile, motorcycles have 328 samples, trains have 550 samples, and trucks have 240 samples. This variation in the number of samples reflects the availability of records within the dataset, which can affect the class balance in the machine learning process.

All data in this dataset is in the form of audio files in .wav format obtained from various environmental conditions. The available recordings reflect differences in sound intensity, frequency patterns, and the potential for background noise. Therefore, this dataset presents its own challenges in the acoustic analysis process, especially in identifying unique patterns for each type of vehicle. The acoustic characteristics of this dataset allow for further exploration of sound patterns that can be used in the classification of sound-based vehicles. With various variations in data, this study can evaluate the effectiveness of the method used in distinguishing sounds between vehicle classes based on their acoustic features.

III.2 FEATURE EXTRACTION

The feature extraction process aims to convert the audio signal data into numerical representations that can be used in further analysis. In this study, three main features were used, namely Mel-Frequency Cepstral Coefficients (MFCC), Spectral Centroid, and Spectral Bandwidth. Each of these features has a specific role in representing the acoustic characteristics of the vehicle's sound signal.

III.2.1 Mel-Frequency Cepstral Coefficients (MFCC).

MFCC is the most commonly used feature in voice-based pattern recognition, including vehicle acoustic analysis. The MFCC calculation process begins by applying Fast Fourier Transform (FFT) to transform the signal from the time domain to the frequency domain. FFT is calculated using Equation 1.

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-j2\pi kn/N} \quad (1)$$

where $X(k)$ is a representation of the signal in the frequency domain, $x(n)$ is the signal in the time domain, and N is the number of samples in the analysis window. Furthermore, the frequency resulting from the FFT transformation is mapped into the Mel scale, which reflects the human perception of sound. The Mel Scale is calculated using Equation 2.

$$m = 2595\, log_{10}(1 + \frac{f}{7000}) \quad (2)$$

where f is the frequency in Hz, and m is the frequency in the Mel scale. The final step is to take the logarithm of the Mel bank filter energy and apply the Discrete Cosinus Transform (DCT) to eliminate the correlation between the coefficients. DCT is calculated by Equation 3.

$$C_n = \sum_{k=1}^{K} log(E_k)\, cos[n\frac{\pi}{K}(k - 0.5)] \quad (3)$$

where C_n is the MFCC coefficient to- n, E_k adalah energi dalam filter bank ke- k, dan K adalah jumlah filter bank. Dalam penelitian ini, digunakan 13 koefisien MFCC untuk merepresentasikan pola frekuensi utama dari sinyal akustik kendaraan.

III.2.2 SPECTRAL CENTROID

Spectral centroid is a metric used to measure the spectral center of a sound signal. This feature is often associated with the perception of sound "brightness", where the sound with the dominant high frequency will have a greater spectral centroid value. Spectral centroid is calculated with Equation 4.

$$SC = \frac{\sum_{k=1}^{N} f_k S_k}{\sum_{k=1}^{N} S_k} \quad (4)$$

where SC is a spectral centroid, f_k is the frequency of the bin ke-k, S_k is the magnitude of the spectrum on the bin to-k, and N is the total number of frequency bins. Higher spectral centroid values indicate a concentrated distribution of energy at high frequencies, while lower values indicate a predominance of low frequencies.

III.2.3 SPECTRAL BANDWIDTH

Spectral bandwidth measures how dispersed the spectral energy of a signal is around a spectral centroid. This feature is important in distinguishing vehicles based on the complexity of the sound produced. Vehicles such as buses or trucks tend to have a wider spectral bandwidth than motorcycles because they produce more complex sounds. Spectral bandwidth is calculated by Equation 5.

$$SBW = \sqrt{\frac{\sum_{k=1}^{N} S_k (f_k - SC)^2}{\sum_{k=1}^{N} S_k}} \quad (5)$$

where SBW is spectral bandwidth, SC is spectral centroid, f_k is the frequency of bin to-k, S_k is the spectral magnitude at the kth bin, and N is the total number of frequency bins. Higher spectral bandwidth values indicate a more dispersed distribution of energy in the frequency domain, while lower values indicate energy concentrations over a narrow frequency range.

III.3 PREPROCESSING

The preprocessing stage is carried out to ensure that the data extracted from the audio recording has a uniform format and is ready to be used in analysis. The first step is to normalize the feature using the StandardScaler method, which makes the data have a mean of zero and a standard deviation of one. This normalization is important to ensure consistent data scales and support the stability of machine learning algorithms. Furthermore, the negative value on the data is changed to zero to meet the requirements of certain methods, such as Latent Dirichlet Allocation (LDA), which can only work with non-negative data. Also, because the length of the audio file varies, padding is done to add a zero value at the end of the feature, so all recordings have the same dimensions. The maximum length is determined based on the longest record in the dataset. This preprocessing process ensures that the data is free from dimensional and scale inconsistencies, so it is ready to be applied to dimensional reduction and classification methods efficiently.

III.4 DIMENSIONALITY REDUCTION TECHNIQUES

Dimension reduction is a technique used to reduce the number of variables in a dataset while still retaining relevant information. In this study, six dimensional reduction methods were compared to determine their effectiveness in vehicle acoustic analysis, namely Principal Component Analysis (PCA), Kernel PCA, Incremental PCA, Independent Component Analysis (ICA), Truncated Singular Value Decomposition (TSVD), and Latent Dirichlet Allocation (LDA). Each method has different characteristics and approaches to handling high-dimensional data.

III.4.1 PRINCIPAL COMPONENT ANALYSIS (PCA)

PCA is an eigen-decomposition-based dimensionality reduction method that aims to find the key components that capture the greatest variation in a dataset. PCA works by inventing linear transformations that project data into a new space with a smaller number of dimensions but still retain as much information as possible from the original data. The main steps in PCA involve data normalization, calculation of covariance matrices, decomposition of eigen, and transformation to the main component space. 1 detailed algorithm of PCA.

Algorithm 1: *steps from the PCA.*

Input	:	Dataset X with dimension m×n (m samples, n features)
Output	:	Dataset X_{PCA} with dimensions $m \times k$, where $k<n$
Step	1	Data Normalization: Standardize each feature using $X_{norm} = \frac{X-\mu}{\sigma}$
	2	Calculate the Covariance Matrix: $C = \frac{1}{m} X_{norm}^T X_{norm}$
	3	Hitung eigenvalue dan eigenvector: $CW = \lambda W$
	4	Choose k Eigenvektor With Eigenvalue Biggest
	5	Transformasi Data: $X_{PCA} = X_{norm} W_k$

Source: Authors, (2025).

III.4.2 KERNEL PRINCIPAL COMPONENT ANALYSIS (KERNEL PCA)

The PCA kernel is an extension of PCA that allows the detection of non-linear patterns in data by using kernel functions to map data to a high-dimensional feature space before applying PCA. This allows the identification of non-linear relationships that cannot be found with standard PCA. Algorithm 2 is a process of the PCA Kernel

Algorithm 2: *Kernel Process* PCA.

Input	:	Dataset X with dimensions $m \times n$, kernel functions $K(x_i, x_j)$
Output	:	Dataset X_{kPCA} In the new feature space with dimension k
Step	1	Implement Kernel Function: Calculate the kernel matrix $K_{ij} = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right)$
	2	Calculate the Eigenvalue and Eigenvector of the Kernel
	3	Select k Eigenvector with the Largest Eigenvalue
	4	Transformasi Data: $X_{kPCA} = KW_k$

Source: Authors, (2025).

III.4.3 INCREMENTAL PRINCIPAL COMPONENT ANALYSIS

Incremental PCA is a variation of PCA that allows the process of gradual reduction of dimensions in small batches, making it more efficient for large datasets. This technique updates the main component each time a new batch is processed without the need to store the entire dataset in memory. Algorithm 3 is the process of Incremental PCA.

Algorithm 3: *process* Incremental PCA.

Input	:	Dataset X Large, batch size B
Output	:	Dataset X_{IPCA} with smaller dimensions
Step	1	Take Batch Data: X_t dari dataset
	2	Normalisasi Data Batch
	3	Update Komponen PCA dengan Batch New: $X_{IPCA}^{(t)} = X_{IPCA}^{(t-1)} + \eta(X_t - X_{IPCA}^{(t-1)})$
	4	Iterasi Until all data is processed

Source: Authors, (2025).

III.4.4 INDEPENDENT COMPONENT ANALYSIS (ICA)

The ICA aims to find statistically independent components in the data. This technique is often used in signal processing, such as separating sound sources from recordings that contain multiple sources. Algorithm 4 is an ICA process.

Algorithm 4: Independent Component Analysis (ICA).

Input	:	Dataset X which is a linear mixture of the source S
Output	:	Independent source S
Step	1	Standarisasi Data
	2	Estimation of the Mixing Matrix: $X = AS$
	3	Find the W Separator Matrix by maximizing statistical independence
	4	Separate Sources: $S = WX$

Source: Authors, (2025).

III.4.5 TRUNCATED SINGULAR VALUE DECOMPOSITION (TSVD)

TSVD is a matrix decomposition-based dimension reduction technique that retains only the largest k singular value to simplify the data. Algorithm 5 is the process of TSDV

Algorithm 5: Proses Truncated Singular Value Decomposition.

Input	:	Dataset X With dimensi $m \times n$
Output	:	Dataset X_{TSVD} with smaller dimensions
Step	1	Do Dekomposisi SVD: $X = U\Sigma V^T$
	2	Select k Largest Singular Value
	3	Rekonstruksi Data: $X_{TSVD} = U_k \Sigma_k V_k^T$

Source: Authors, (2025).

III.4.6 LATENT DIRICHLET ALLOCATION (LDA)

LDA is a probabilistic-based method used to identify latent distributions in datasets, often applied in text analysis but can also be used in feature-based classification. Algorithm 6 is the process of the Latent Dirichlet Allocation (LDA) method.

Algorithm 6: Proses Metode Latent Dirichlet Allocation (LDA).

Input	:	Dataset X which represents the distribution of features
Output	:	Latent distribution in datasets
Step	1	Determine the Number of Latent Components k
	2	Calculate Conditional Probability: $P(w\|z) = \frac{n_{w,z}+\beta}{\sum_{w`}(n_{w`,z}+\beta)}$
	3	Iteration to Update Latent Parameters

Source: Authors, (2025).

III.5 CLASSIFICATION

Classification is the final stage in data analysis to predict classes from unknown data based on previously extracted and reduced features. In this study, the Random Forest Classification algorithm is used, which is an ensemble learning-based method consisting of a set of decision trees. Random Forest's advantages include its resistance to overfitting and its ability to handle high-dimensional datasets.

The Random Forest algorithm works by building multiple decision trees from a random subset of datasets (bagging) and combining the prediction results from each tree to produce the final prediction. Algorithm 7 is a random forest classification step.

Algorithm 7: *Steps* random forest classification.

Input	:	Dataset tereduksi X, Class labels y
Output	:	Class predictions for new data
Step	1	Create n random subsets of datasets (X,y) using bootstrap techniques.
	2	For each subset, build a decision tree based on the Gini Index split criteria
	3	Use the results from each tree to predict the test data class.
	4	Take the final prediction result based on the majority of votes from the entire tree (voting).

Source: Authors, (2025).

III.6 EVALUATION

The model evaluation in this study was carried out to assess the performance of various dimensional reduction and classification methods in recognizing vehicle types based on acoustic data. Four main metrics are used in the evaluation, namely accuracy, precision, recall, and F1-score. These four metrics were chosen because they can provide a comprehensive view of the model's ability to make accurate and consistent predictions. Accuracy measures the proportion of correct predictions compared to the total number of samples. Equation 6 how to calculate accuracy. Precision measures how accurate the model is in predicting a particular type of vehicle, such as an ambulance, compared to the total predictions generated as an ambulance. Equation 7 how to calculate precision. Recall measures the model's ability to capture all samples that actually come from a particular vehicle class. Equation 8 is a way to calculate recalls. F1-score is a harmonious average of precision and recall. Equation 9 is a way to calculate F1-Score.

$$Accuracy = \frac{number\ of\ corrent\ predictions}{Total\ number\ of\ sampel} \tag{6}$$

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \tag{7}$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \tag{8}$$

$$F1 - score = 2\ \frac{Precision\ .Recall}{Precision + Recall} \tag{9}$$

IV. RESULTS AND DISCUSSIONS

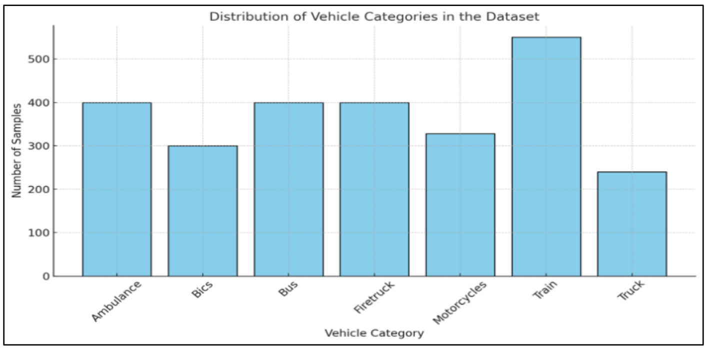


Figure 1: Distribution Of Vehicle Categories In The Dataset.
Source: Authors, (2025).

The dataset used in this study shows an uneven distribution among vehicle categories. The Train category has the highest number of samples, which is 550, while the Truck category has the lowest sample number, which is 240. This imbalance in the amount of data has the potential to create bias in classification results, where the model tends to provide higher accuracy for categories with a larger amount of data. The audio recordings in this dataset are also taken from various environmental conditions, which also adds to the challenges in the classification process. Background noise, such as traffic noise, wind, or other environmental sounds, can obscure the unique acoustic patterns of each vehicle category. This challenge becomes even more significant for categories with fewer sample counts, such as Truck and Bics, as limited data makes their voice pattern representation less than optimal in machine learning models. The distribution of the results of the dataset feature attraction is shown in Figure 2-6, and Table 1.

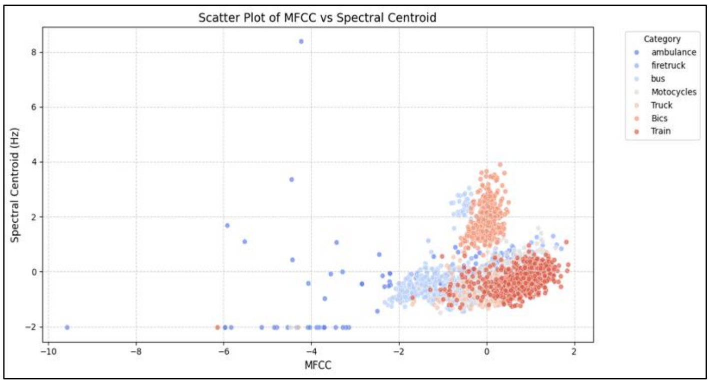


Figure 2 Scatter Plot of MFCC vs Spectral Centroid.
Source: Authors, (2025).

Figure 2 shows the relationship between MFCCs (Mel-Frequency Cepstral Coefficients) and Spectral Centroids for each vehicle category. MFCC represents the primary frequency pattern of acoustic signals, while Spectral Centroid depicts spectral energy centers, which are often associated with the perception of sound "brightness". In Figure 2, it can be seen that vehicle categories such as Train and Bics have a distribution pattern that tends to be separate from other categories, indicating that these features are quite informative in distinguishing several types of vehicles. In contrast, categories such as Ambulance, Bus, and Firetruck have more overlapping distribution patterns, reflecting the similarity in their acoustic characteristics. The distribution of the data also shows the existence of several outliers, especially for categories with a smaller number of samples, such as trucks. This can be a challenge in the classification process, as overlap between categories can lead to prediction errors. However, the combination of MFCC and Spectral Centroid still provides indications of a unique pattern for certain types of vehicles, especially for categories with high-frequency dominant sounds such as Trains.

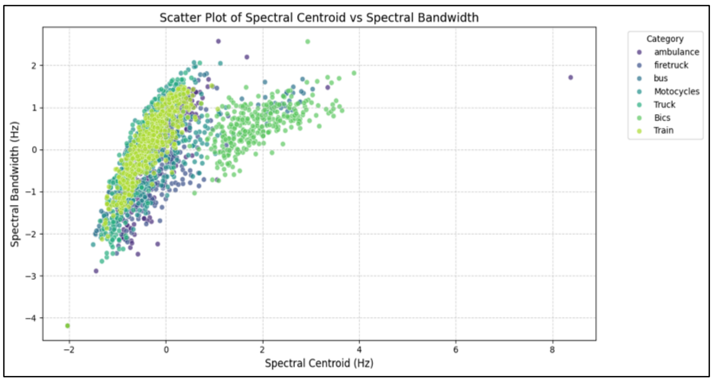


Figure 3: Scatter Plot of Spectral Centroid vs Spectral Bandwidth.
Source: Authors, (2025).

Figure 3 illustrates the relationship between Spectral Centroid and Spectral Bandwidth for each vehicle category in the dataset. Spectral Centroid represents the spectral energy center that is closely related to the dominant frequency in the vehicle's sound, while Spectral Bandwidth measures how wide the spectral energy distribution is around the centroid. The combination of these two features is often used to understand the frequency characteristics and complexity of sound. From Figure 3, it can be seen that most of the vehicle categories have overlapping distributions in the middle area of the graph. For example, categories such as Ambulance, Bus, and Firetruck exhibit fairly similar distribution patterns, suggesting that their acoustic characteristics may have similarities, especially in dominant frequencies and variations in sound bandwidth.

However, some important patterns can be observed, namely the Train and Bics categories show a more dispersed distribution, especially at higher spectral bandwidth values. This reflects that the voices of these categories have greater complexity than other categories. Train sounds, for example, often cover a wide range of frequencies, resulting in a wider bandwidth. The Truck category, although it mostly overlaps with other categories, has some unique data points with higher Spectral Centroid and Spectral Bandwidth values. These dots indicate that there is a certain sound pattern of the truck that is different from other vehicles. Outliers are seen in some categories, such as Ambulances, with very high Spectral Centroid values. Outliers like this can be challenging in the classification process because they can affect model learning and create bias towards certain categories.

Figure 3 also highlights a major challenge in voice-based classification, namely overlap between categories. This overlap suggests that categories such as Ambulance, Bus, and Firetruck have similar acoustic characteristics, making it difficult to separate them using only Spectral Centroid and Spectral Bandwidth. However, this feature remains relevant, especially for categories like Train and Bics, which have more unique acoustic patterns.

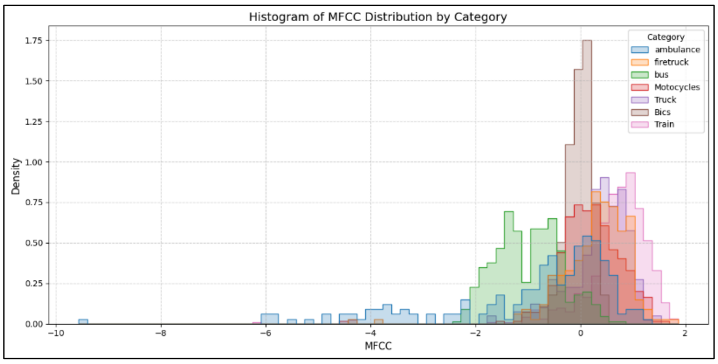


Figure 4: Histogram of MFCC Distribution by Category.
Source: Authors, (2025).

Figure 4 presents a distribution histogram of Mel-Frequency Cepstral Coefficients (MFCC) for each category of vehicles in the dataset. From Figure 4, several important patterns can be observed, namely the MFCC distribution for the Train category shows a high concentration around a small positive value, with a significant peak density compared to other categories. This shows that the sound of the train has a fairly distinctive dominant frequency pattern, which can be an important indicator in the classification process.

The MFCC of the Bics category also has a prominent distribution in the region of small positive values, but with a narrower spread than Train. This pattern reflects that bicycle sounds have a relatively simpler frequency pattern. The Ambulance, Firetruck, and Bus categories have overlapping MFCC distribution patterns, with concentrations around negative to zero values. This overlap suggests that the sounds from emergency vehicles and buses have fairly similar acoustic characteristics, which can be a challenge in differentiating these categories. The Truck and Motorcycles category has a more spread distribution of MFCC, with some negative values that are quite extreme. This suggests that the sound of these two types of vehicles has a greater variation in frequency patterns, perhaps due to differences in engine conditions or recording environments.

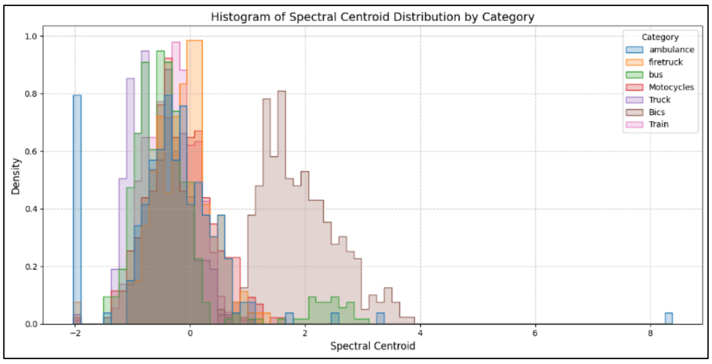


Figure 5: Histogram of Spectral Centroid Distribution by Category.
Source: Authors, (2025).

Figure 5 illustrates the histogram of the Spectral Centroid distribution for each category of vehicles in the dataset. Some important distribution patterns that can be observed from Figure 5 are that the Spectral Centroid distribution for the Train category has a wider range than other categories, with peak densities at values around zero to two. This reflects the complexity of train sounds, which often include a wide range of dominant frequencies due to large engines and wheel systems moving on rails.

The Bics and Truck categories show a fairly unique distribution of Spectral Centroids, with peak densities more spread in small to moderate positive value regions. This spread indicates that the sounds of these two categories have complex frequency patterns, which may reflect variations in engine noise or the influence of recording environmental conditions. The Ambulance, Firetruck, and Bus categories have significant overlap in the Spectral Centroid distribution, with density peaks concentrated around zero values. This overlap suggests that the sounds of emergency vehicles and buses have fairly similar dominant frequency characteristics, which can pose challenges in the classification process. In addition, there are some outliers in certain categories, such as Ambulance, that have Spectral Centroid values far beyond the main distribution. These outliers can be an indicator of the presence of recordings with very different sound characteristics or environmental conditions that affect frequency patterns.

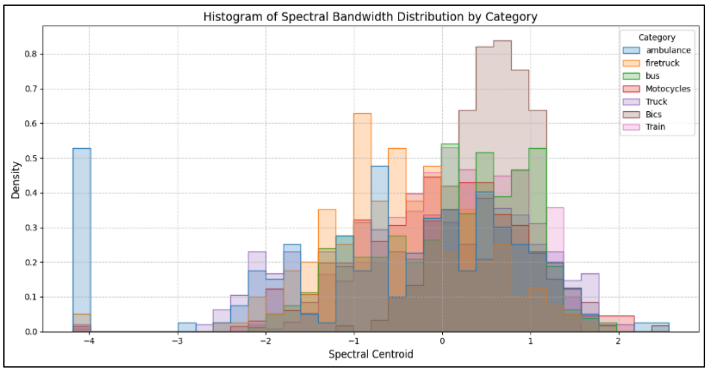


Figure 6: Histogram of Spectral Bandwidth Distribution by Category.
Source: Authors, (2025).

Figure 6 illustrates the distribution of Spectral Bandwidth for each category of vehicles in the dataset. Some important patterns that can be observed from Figure 6, namely the Train category, show a Spectral Bandwidth distribution that tends to be wider, with peak densities around values between 1 and 2. This pattern indicates that the sound of the train has a high frequency complexity, which is one of the hallmarks of this vehicle. The Bics category has a narrower distribution than Train, but still shows a fairly unique pattern in small positive regions. Trucks, on the other hand, have a distribution spread over almost the entire area of the histogram, reflecting the presence of large variations in sound patterns that may be caused by different types of trucks or different recording conditions. The Ambulance, Firetruck, and Bus categories show significant overlap in the distribution of Spectral Bandwidth. A similar distribution among these categories indicates that the complexity of the sounds of emergency vehicles and buses has characteristics that are difficult to distinguish.

Table 1: Descriptive Statistics of Dataset.

Category	MFCC		Spectral Centroid		Spectral Bandwidth	
	Mean	Std	Mean	Std	Mean	Std
Bics	-0.227	1.449	1.90	0.655	0.56	0.484
Motocycles	-0.217	0.897	-0.189	0.558	-0.025	0.923
Train	0.319	0.634	-0.316	0.425	0.146	0.777
Truck	0.135	0.965	-0.55	0.459	-0.238	1.163
Ambulance	-0.263	1.212	-0.298	1.018	-0.611	1.596
Bus	0.066	0.84	-0.297	0.867	0.093	0.856
Firetruck	-0.213	0.963	-0.115	0.499	-0.489	0.899

Source: Authors, (2025).

In addition to being explained using Figures 2 to 6. Table 1, descriptive statistical analysis provides a deeper insight into the feature characteristics of MFCC, Spectral Centroid, and Spectral Bandwidth extracted from the dataset. The mean and standard deviation (std) values for each feature in each category provide an overview of the data distribution pattern, as well as show the potential of the feature to differentiate vehicle categories. In the MFCC feature, the category with the highest average score is Train (0.319). This reflects that the sound pattern of the train has better stability than other categories. In contrast, the lowest average MFCC value was found in Ambulance (-0.263), indicating that the sound of these emergency vehicles tends to have a lower dominant frequency. In addition, the greatest variation in MFCC was seen in the Bics category, with a standard deviation of 1,449. This suggests that the sound of the bicycle has a more diverse frequency pattern, which is most likely due to factors such as engine variations or recording environmental conditions. In the Spectral Centroid feature, the highest average value was found in the Bics category (1.90), which indicates that the bike has a sound with a higher dominant frequency than other categories.

In contrast, the lowest average value for this feature is found in the Truck category (-0.55), which reflects that truck sounds tend to have lower dominant frequencies. In addition, the highest standard deviation for this feature is found in the Ambulance category (1,018), indicating a large variation in the sound patterns of these emergency vehicles. This large variation can be caused by differences in intensity or the type of siren used. Meanwhile, on the Spectral Bandwidth feature, the category with the highest average value is Bics (0.56), which indicates that bicycle sounds have a wider spectrum of energy, reflecting a higher complexity of sound. In contrast, the lowest average bandwidth value was found in the Ambulance category (-0.611), which indicates a lower sound complexity for these emergency vehicles. The largest variation in bandwidth was also found in the Ambulance category, with a standard deviation of 1,596. This suggests that the sound from these emergency vehicles can vary greatly depending on the recording situation or the intensity of the siren used. The vehicle dataset based on sound was trained and tested using random forest classification. The division of training and test data was used, namely 90:10, 80:20, 70:30, and 60:40 percent. Table 2 is the division of training data and test data.

Table 2: number of training data and test data based on ratio.

Kendaraan	Ratio	Jumlah Data Latih	Jumlah Data Uji
Bics	90:10	270	30
	80:20	240	60
	70:30	210	90
	60:40	180	120
Motocycles	90:10	295	33
	80:20	262	66
	70:30	229	99
	60:40	196	132
Train	90:10	495	55
	80:20	440	110
	70:30	385	165
	60:40	330	220
Truck	90:10	216	24
	80:20	192	48
	70:30	168	72
	60:40	144	96
Ambulance	90:10	180	20
	80:20	160	40
	70:30	140	60
	60:40	120	80
Bus	90:10	360	40
	80:20	320	80
	70:30	280	120
	60:40	240	160
Firetruck	90:10	180	20
	80:20	160	40
	70:30	140	60
	60:40	120	80

Source: Authors, (2025).

Based on Table 3, this study shows that the effectiveness of the dimension reduction method is highly dependent on the ratio of training data and test data used, with each method showing superiority in certain scenarios. At a training data:test data ratio of 90:10, Incremental PCA recorded the best performance with an accuracy of 0.982 using 18 components. Methods such as PCA, ICA, and SVD also showed excellent results with an accuracy of 0.977. Meanwhile, the PCA and LDA kernels lag slightly behind with an accuracy of 0.973. When the proportion of test data increases to 20% (80:20 ratio), the performance of most methods remains competitive. Incremental PCA, PCA, Kernel PCA, SVD, and LDA all recorded the same accuracy of 0.973. However, the ICA showed an advantage with an accuracy of 0.975 using 13 components, which reflects its ability to handle a larger proportion of test data without losing significant information.

At a 70:30 ratio, ICA recorded the highest accuracy of 0.979 using only 13 components, demonstrating its efficiency in retaining critical information even though the test data was more dominant. Incremental PCA remains competitive with an accuracy of 0.970 using 20 components. The PCA and SVD methods showed the same results with an accuracy of 0.968, while the PCA and LDA kernels experienced a decrease in accuracy to 0.965 and 0.967, respectively. When the ratio of training data:test data was 60:40, SVD showed the highest performance with an accuracy of 0.976 using the best 18 components, followed by ICA with an accuracy of 0.972 using 19 components. Incremental PCA remains relevant with an accuracy of 0.967 even though it uses only 13 components. LDA also gave good results with an accuracy of 0.974 using 17 components. On the other hand, PCA and PCA Kernel recorded an accuracy of 0.967 and 0.970.

The results of this study show how each dimension reduction method provides optimal performance in various ratios of training data and test data, confirming the relevance of this approach in different scenarios. Incremental PCA is the most superior method in situations with a high training data ratio, such as 90:10, where a larger amount of training data (e.g., 495 training data for the Train category) allows the method to utilize more information to maintain high accuracy (0.982). This highlights the advantages of Incremental PCA in managing large and information-rich datasets. In contrast, at larger test data ratios, such as 60:40, ICAs stand out with their best performance, achieving an accuracy of up to 0.979 at a 70:30 ratio with only 13 components.

ICA's ability to retain critical information even though the number of components is less indicates its efficiency in situations where test data is more dominant. This is particularly relevant for applications where dimension reduction must be achieved without sacrificing accuracy on larger test data. SVD, with its stability in all ratios, reflects its reliability as a reliable method for various data proportions. Regardless of the data ratio, SVD consistently maintains high accuracy, making it a safe choice for scenarios that require versatile dimension reduction. In addition, LDA exhibits competitive performance, especially in applications where component interpretability is a priority. Although the accuracy of LDA is slightly lower than that of Incremental PCA or ICA, its stability makes it a viable option for certain cases. PCA and Kernel PCA, although recording a slight decrease in performance at smaller training data ratios, still demonstrated high adaptability with an accuracy close to 0.97 at most ratios. These results show that these two methods remain relevant in scenarios where the dataset has a more varied

Table 3: Best evaluation results.

Method	Ratio	Best n_components	Accuracy	Precision	Recall	F1-Score
PCA	90:10	17	0,977	0,978	0,977	0,977
	80:20	19	0,973	0,973	0,973	0,973
	70:30	17	0,968	0,969	0,968	0,968
	60:40	20	0,967	0,967	0,967	0,967
Kernel PCA	90:10	16	0,973	0,973	0,973	0,973
	80:20	19	0,973	0,973	0,973	0,973
	70:30	14	0,965	0,965	0,965	0,965
	60:40	17	0,970	0,970	0,970	0,970
Incremental PCA	90:10	18	0,982	0,982	0,982	0,982
	80:20	19	0,973	0,973	0,973	0,973
	70:30	20	0,970	0,970	0,970	0,970
	60:40	13	0,967	0,967	0,967	0,967
ICA	90:10	17	0,977	0,978	0,977	0,977
	80:20	13	0,975	0,976	0,975	0,975
	70:30	13	0,979	0,979	0,979	0,979
	60:40	19	0,972	0,972	0,972	0,972
SVD	90:10	18	0,977	0,978	0,977	0,977
	80:20	12	0,973	0,973	0,973	0,973
	70:30	19	0,968	0,969	0,968	0,968
	60:40	18	0,976	0,976	0,976	0,976
LDA (Latent Dirichlet)	90:10	19	0,973	0,973	0,973	0,973
	80:20	17	0,973	0,973	0,973	0,973
	70:30	17	0,967	0,968	0,967	0,967
	60:40	17	0,974	0,974	0,974	0,974

Source: Authors, (2025).

distribution. This study emphasizes that the selection of dimension reduction methods must consider the distribution of training data and test data. These findings make a significant contribution to the understanding of the effectiveness of various dimension reduction approaches, both for datasets with a high proportion of training data and for datasets with more dominant test data. With in-depth analysis and measurable results, the study offers clear guidance for other researchers in choosing the method that best suits their needs.

V. CONCLUSIONS

This study shows the effectiveness of dimension reduction techniques in overcoming the challenge of high-dimensional acoustic data for vehicle recognition. By comparing six methods on different ratios of training and test data, Incremental PCA proved to be the most effective approach for scenarios with larger training datasets, achieving the highest accuracy of 0.982. Meanwhile, ICA excels with a smaller number of components under the condition of a higher proportion of test data, reflecting its efficiency in maintaining important features. SVD's stability across all ratios highlights its reliability for a wide range of applications, while LDA remains relevant for tasks that prioritize interpretability. These findings emphasize the importance of selecting appropriate dimension reduction methods based on data characteristics and application needs, and provide valuable insights for the development of efficient and accurate intelligent transportation systems.

VI. AUTHOR'S CONTRIBUTION

Conceptualization: Yudhi Darmawan and Muladi.
Investigation: Yudhi Darmawan.
Discussion of results: Yudhi darmawan, Muladi, Heru Wahyu Herwanto, and Nur Rachman Supadmana Muda.
Methodology: Yudhi Darmawan, Muladi.
Writing – Original Draft: Yudhi Darmawan, and Muladi. **Writing – Review and Editing:** Yudhi Darmawan, Muladi.Heru Wahyu Herwanto and Nur Rachman Supadmana Muda.
Supervision: Muladi, Heru Wahyu Herwanto, and Nur Rachman Supadmana Muda.
Approval of the final text: Yudhi darmawan, Muladi, Heru Wahyu Herwanto, and Nur Rachman Supadmana Muda.

VII. ACKNOWLEDGMENTS

We thank the promoters for their support in realizing this article. This research is part of the doctoral program in the electrical engineering and informatics study program at the State University of Malang. All authors contributed to the research and writing of this article with the authority and approval to submit the manuscript.

VIII. REFERENCES

[1] T. M. Nithya, P. Dhivya, S. N. Sangeethaa, and P. Rajesh Kanna, "TB-MFCC multifuse feature for emergency vehicle sound classification using multistacked CNN – Attention BiLSTM," Biomed. Signal Process. Control, vol. 88, no. PA, p. 105688, Feb. 2024, doi: 10.1016/j.bspc.2023.105688.

[2] S. Damiano, L. Bondi, S. Ghaffarzadegan, A. Guntoro, and T. van Waterschoot, "Can Synthetic Data Boost the Training of Deep Acoustic Vehicle Counting Networks?," in ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, Apr. 2024, pp. 631–635. doi: 10.1109/ICASSP48485.2024.10447868.

[3] C.-Y. Chiang, R. Zhong, M. Jaber, and P. Hayward, “A Distributed Acoustic Sensor System for Intelligent Transportation using Deep Learning,” IEEE Internet Things Mag., vol. 7, no. 5, pp. 88–96, Sep. 2024, doi: 10.1109/IOTM.001.2300287.

[4] K. K. M. Shariff, M. A. Haron, M. S. A. M. Ali, I. M. Yassin, M. H. Azami, and M. A. Kechik, “Comparison of Spectrograms for Classification of Vehicles from Traffic Audio,” in 2024 IEEE Symposium on Wireless Technology & Applications (ISWTA), IEEE, Jul. 2024, pp. 12–17. doi: 10.1109/ISWTA62130.2024.10651848.

[5] R. Evans, S. Mullins, and G. Noel, “Challenges of Vehicle Classification Using Acoustics,” Int. FLAIRS Conf. Proc., vol. 35, May 2022, doi: 10.32473/flairs.v35i.130658.

[6] A. I. Yassin, K. K. Mohd Shariff, M. A. Kechik, A. M. Ali, and M. S. Megat Amin, “Acoustic Vehicle Classification Using Mel-Frequency Features with Long Short-Term Memory Neural Networks,” TEM J., vol. 12, no. 3, pp. 1490–1496, Aug. 2023, doi: 10.18421/TEM123-29.

[7] M. Ashhad, U. Goenka, A. Jagetia, P. Akhtari, S. K. Ambat, and M. Samuel, “Improved Vehicle Sub-type Classification for Acoustic Traffic Monitoring,” in 2023 National Conference on Communications (NCC), IEEE, Feb. 2023, pp. 1–6. doi: 10.1109/NCC56989.2023.10067994.

[8] J. Murovec, J. Prezelj, D. Ćirić, and M. Milivojčević, “Zero Crossing Signature: A Time-Domain Method Applied to Diesel and Gasoline Vehicle Classification,” IEEE Sens. J., vol. PP, no. Xx, pp. 1–1, 2024, doi:
10.1109/JSEN.2024.3516876.

[9] A. Chandra and G. Singh, “Hybrid Classification Model For Acoustic-Based Emergency Vehicle Detection,” in 2023 6th International Conference on Contemporary Computing and Informatics (IC3I), IEEE, Sep. 2023, pp. 1199–1205. doi: 10.1109/IC3I59117.2023.10397746.

[10] Y. Luo, L. Chen, Q. Wu, and X. Zhang, “Sound-Convolutional Recurrent Neural Networks for Vehicle Classification Based on Vehicle Acoustic Signals,” in 2021 International Conference on Smart City and Green Energy (ICSCGE), IEEE, Nov. 2021, pp. 98–102. doi: 10.1109/ICSCGE53744.2021.9654357.

[11] Y. Pan, Y. Dong, D. Wang, J. Di, and A. Chen, “Vehicle Detection and Classification by Voiceprint Recognition Based on Single Acoustic Sensor Under Bridge Expansion Joint,” IEEE Trans. Instrum. Meas., vol. 73, pp. 1–15, 2024, doi: 10.1109/TIM.2024.3428610.

[12] R. Zhong, C.-Y. Chiang, and M. Jabor, “Distributed Acoustic Sensor System for Vehicle Occupancy Detection,” in 2024 IEEE World Forum on Public Safety Technology (WFPST), IEEE, May 2024, pp. 132–137. doi: 10.1109/WFPST58552.2024.00022.

[13] C.-Y. Chiang, M. Jaber, K. K. Chai, and J. Loo, “Distributed Acoustic Sensor Systems for Vehicle Detection and Classification,” IEEE Access, vol. 11, no. March, pp. 31293–31303, 2023, doi: 10.1109/ACCESS.2023.3260780.

[14] I. Corera, E. Piñeiro, J. Navallas, M. Sagues, and A. Loayssa, “Long-Range Traffic Monitoring Based on Pulse-Compression Distributed Acoustic Sensing and Advanced Vehicle Tracking and Classification Algorithm,” Sensors, vol. 23, no. 6, p. 3127, Mar. 2023, doi: 10.3390/s23063127.

[15] J. Abeber, S. Gourishetti, A. Katai, T. Claub, P. Sharma, and J. Liebetrau, “IDMT-Traffic: An Open Benchmark Dataset for Acoustic Traffic Monitoring Research,” in 2021 29th European Signal Processing Conference (EUSIPCO), IEEE, Aug. 2021, pp. 551–555. doi: 10.23919/EUSIPCO54536.2021.9616080.

[16] S. Damiano, L. Bondi, A. Guntoro, and T. van Waterschoot, “A framework for the acoustic simulation of passing vehicles using variable length delay lines,” EURASIP J. Audio, Speech, Music Process., vol. 2024, no. 1, p. 49, Oct. 2024, doi: 10.1186/s13636-024-00372-4.

[17] Y. Houkpati et al., “Siamese Networks for Autonomous Classification of Battlefield Ground Vehicles Using Acoustic Data,” in 2023 International Conference on Computational Science and Computational Intelligence (CSCI), IEEE, Dec. 2023, pp. 1–6. doi: 10.1109/CSCI62032.2023.00008.

[18] V.-T. Tran and W.-H. Tsai, “Acoustic-Based Emergency Vehicle Detection Using Convolutional Neural Networks,” IEEE Access, vol. 8, pp. 75702–75713, 2020, doi: 10.1109/ACCESS.2020.2988986.

[19] A. Y. Yazli, K. K. Mohd Shariff, S. A. Mutalib Al Junid Syed Abdul Rahman, and M. A. Md Ali, “Multimodal Vehicle Classification Based on Radar and Acoustic Sensors Using Hybrid Shallow CNN,” in 2024 IEEE Symposium on Wireless Technology & Applications (ISWTA), IEEE, Jul. 2024, pp. 264–268. doi: 10.1109/ISWTA62130.2024.10652023.

[20] A. Barros et al., “Beyond noise levels: vehicle classification using psychoacoustic indicators from pass-by road traffic noise and their correlations with speed and temperature,” Appl. Acoust., vol. 214, no. June, p. 109716, Nov. 2023, doi: 10.1016/j.apacoust.2023.109716.

[21] Kaggle, “Vehicle Sounds Dataset.” Accessed: Jan. 30, 2025. [Online]. Available: https://www.kaggle.com/datasets/janboubiabderrahim/vehicle-sounds-dataset

[22] Kaggle, “Emergency Vehicle Siren Sounds.” Accessed: Jan. 30, 2025. [Online]. Available: https://www.kaggle.com/datasets/vishnu0399/emergency-vehicle-siren-sounds/