SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V11I9P114

Volume 11 Issue 9, 149-162, September 2024

Original Article

Automated Detection of Anomalous Holes in Fuel Injector Nozzles Using High-Definition Microphones (DH-FINM) and Machine Learning Algorithms

T. A. Mohanaprakash[1], D. Siva[2], J. Jegan[3], S. Janagiraman[4], M. Therasa[5]

[1]*Department of CSE, School of Engineering and Technology, C M R University, Bangalore, Karnataka, India.*
[2]*Department of Computer Science and Applications, SRM Institute of Science and Technology, Ramapuram, Chennai, India.*
[3]*Department of CSE, School of Technology, The Apollo University, Chittoor, Andhra Pradesh, India.*
[4]*Department of CSE,St. Joseph's Institute of Technology, OMR, Chennai, Tamil Nadu, India.*
[5]*Department of CSE, Panimalar Engineering College, Poonamallee, Chennai, Tamil Nadu, India.*

[1]*Corresponding Author : tamohanaprakash@gmail.com*

Received: 04 July 2024 Revised: 14 August 2024 Accepted: 05 September 2024 Published: 30 September 2024

Abstract - *This study addresses a critical issue faced by fuel injector nozzle manufacturers: the identification of defects in holes caused by drill bit overheating. Manual inspection methods are often inadequate, leading to production delays and the creation of substandard products. To overcome this challenge, we propose a novel approach that combines high-resolution microphones with advanced machine learning algorithms. The system is designed to effectively isolate drilling sounds from background noise, enabling the use of a deep learning model to accurately differentiate between the sound signatures of a properly functioning drill bit and one that is damaged. This approach offers real-time monitoring and prompt alerts, ensuring that defective products are detected early in the production process, significantly reducing the likelihood of errors in the final output. By automating the detection process, the system not only enhances productivity but also improves overall product quality. This solution outperforms traditional methods by eliminating human error, minimizing downtime, and maintaining consistent production standards. The integration of high-resolution microphones allows for precise acoustic analysis, which is critical in identifying subtle differences in sound that may indicate defects. The machine learning model is trained on a comprehensive dataset of sound patterns, ensuring robust performance even in varied manufacturing environments. The proposed technology (DH-FINM –Detection of Holes in Fuel Injector Nozzles using Microphones) offers a cost-effective, scalable, and reliable solution that aligns with the industry's need for efficient and accurate quality control mechanisms. This innovative approach provides a substantial improvement over conventional inspection techniques, contributing to higher efficiency, reduced waste, and improved product reliability. The adoption of this system could set a new standard for quality control in fuel injector nozzle manufacturing, paving the way for broader applications of similar technologies across other sectors where precision and quality are paramount.*

Keywords - *Machine learning algorithms, Nozzle, Injector, Microphones, Noise*.

1. Introduction

Ensuring the production of high-quality goods is essential in the industrial sector, particularly for fuel injector nozzle manufacturers, where even minor deviations in the production process can result in costly defects and potential safety hazards. A critical challenge for these manufacturers is detecting errors in holes drilled at the bottom of the nozzles, which often result from internal drill bit breakage and overheating. These defects are difficult to identify through visual inspection alone. The current method involves a labor-intensive process where each product must be meticulously inspected by a specialized team after drilling. To address this issue and enhance the efficiency and accuracy of the production process, a solution utilizing high-resolution microphones and machine learning models is proposed. This system combines acoustic analysis with deep learning algorithms to enable real-time monitoring and detection of drilling faults, ensuring the production of defect-free final products. High-resolution microphones strategically placed in the drilling area form the core of this system, capturing the acoustic signals of the drilling process, including both the desired drilling sounds and unwanted background noises such as coolant sounds, metal rumbling, and ambient noise. The challenge lies in isolating the drilling sound from other auditory signals to focus on the critical data related to the drill bit's performance. To achieve this, a machine learning model is developed and trained using a diverse dataset of audio inputs, encompassing both drilling sounds and various unwanted noises typically present in drilling environments. Through exposure to this range of audio inputs, the model learns to filter out the irrelevant sounds and accurately identify the drilling sound within the audio spectrum. Beyond isolating the drilling sound, the proposed technology evaluates the drill bit's performance by employing a second machine learning model. This model distinguishes between the sound patterns produced by a functional drill bit and a damaged one. The analysis involves multiple preprocessing steps, including converting mono audio files to stereo, normalizing sample rates, standardizing audio lengths, and transforming audio files into Mel spectrograms, which visually represent the frequency content of the sound signals.

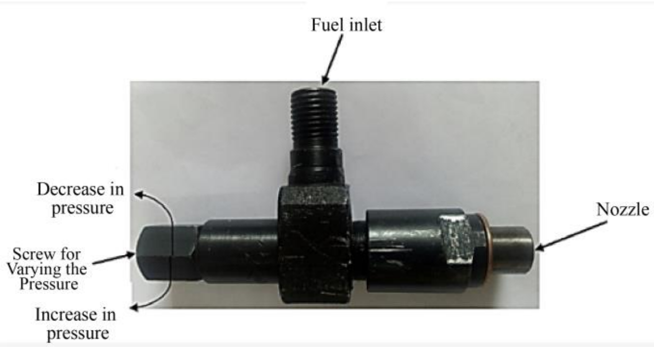

Fig. 1 Fuel injector parts

A deep learning model, pre-trained on a comprehensive dataset of labeled examples of both healthy and damaged drill bit sounds, utilizes these Mel spectrogram images. By leveraging the learned features within this model, the system efficiently differentiates between the acoustic patterns of a functional and a damaged drill bit. When a damaged drill bit sound pattern is detected, the technology promptly alerts the operator to stop the drilling process, preventing further production of defective products and mitigating potential risks associated with continued use of damaged drill bits. By adopting this approach (Figure 1), fuel injector nozzle manufacturers can significantly enhance quality control while reducing the reliance on manual inspections and the associated time and labor costs. Real-time monitoring facilitates the immediate detection of drilling faults, allowing for swift corrective action and preventing the production of substandard goods. Ultimately, this method enables manufacturers to streamline operations, improve product quality, and increase customer satisfaction.

2. Background Study and Related Work

Intricate drilling procedures are used throughout the production of the fuel injector nozzle, as depicted in (Figure 2), and these processes are essential to the product's entire performance. For manufacturers, a major issue is presented by the occurrence of erroneous holes being drilled at the bottom of the nozzle as a result of drill bit overheating. The conventional method of identifying these faults exclusively by human examination has shown to be time-consuming, prone to human error, and ineffective in detecting minute flaws. As a result, producing faulty items exposes producers to higher costs, lower productivity, and other hazards. Emerging technologies like machine learning and high-definition microphones have been used to create creative methods to get around these restrictions. These innovations have the potential to speed up the detection process, increase accuracy, and boost output throughout the industrial sector. Speech recognition, sound categorization, and anomaly detection are just a few of the audio analysis tasks for which machine learning models have recently shown exceptional proficiency.

Researchers have investigated the use of machine learning in monitoring and quality control procedures across a variety of sectors by using these improvements. Convolutional Neural Networks (CNNs), a kind of deep learning model, have also been shown to be quite successful at interpreting audio information. As a result of their excellent feature extraction skills, CNNs are highly suited for the identification of abnormal drill bit performance since they can recognize intricate patterns and fluctuations within sound data.

Previous research has effectively used machine learning and deep learning approaches in a variety of industrial areas, including product quality control, production line defect identification, and anomaly detection. However, the special difficulty of spotting drilling mistakes brought on by broken drill bits during the manufacture of fuel injector nozzles is still completely unexplored.

In order to specifically address the problem of drilling mistake detection, the suggested system expands upon these pillars by integrating high-definition microphones, machine learning models, and deep learning strategies. The method seeks to distinguish between the intended drilling sound and other acoustic disturbances by training the machine learning models on a heterogeneous dataset made up of ambient noises and drilling sounds. The ability to reliably categorize the audio signals is provided by the use of a deep learning model that has been trained on labeled data representing both good and damaged drill bit sounds.

This method makes it possible to monitor the drilling process in real-time and makes it easier to act quickly to stop the manufacturing of faulty goods. The suggested method is intended to considerably enhance the fuel injector nozzle manufacturing process by completing a thorough background study on relevant research and using the developments in audio analysis and machine learning. High-definition microphones and sophisticated machine learning models are used to identify drilling mistakes in real-time, resulting in improved quality control, shorter manufacturing times, and eventually error-free final goods.

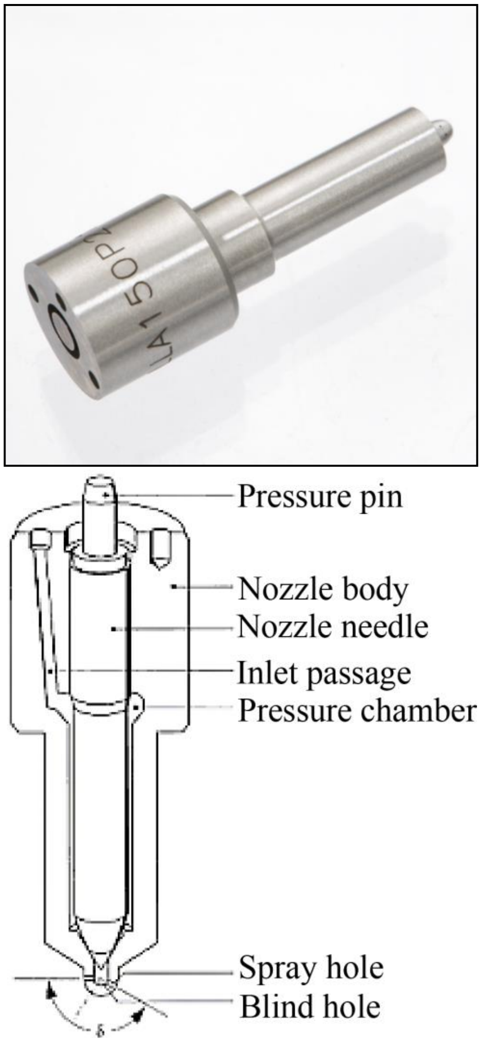

Fig. 2 Fuel injector nozzle

Kim, Kj., and Sun, JO. "Isolation of speaker-related vibrations in audio-visual electronic devices without deteriorating speaker cone vibration." The authors of this paper suggest a technique for isolating speaker-related vibrations in audio-visual electronic devices. [1] The objective is to enhance audio quality without adversely influencing speaker cone vibration. The researchers successfully lessen the transfer of vibrations from the speaker to the device's construction by using a passive isolation technique. The suggested technique successfully isolates vibrations, resulting in improved audio performance without affecting the speaker's use.

S. Xiao, H. Zhang, L. Jiang, and J. Zhang. The phrase "Isolated Word Recognition with Audio Derivation and CNN." This study employs convolutional neural networks (CNNs) and audio derivation to analyze solitary word recognition. The authors create a CNN-based model that identifies isolated words and extracts pertinent information from audio inputs. The proposed model achieves excellent accuracy in word recognition tasks by making use of CNNs [2]. The study allows the creation of voice-controlled devices with enhanced word recognition skills and advances speech recognition technology. The H. Phan et al. "Multi-label Multi-task Convolutional Recurrent Neural Networks for Isolated and Overlapping Audio Event Detection." The difficulty of auditory event recognition is addressed in this study, with a focus on isolated and [3] overlapping events. By combining convolutional and recurrent layers, the model captures both temporal and spectral information for enhanced event recognition. The authors present a multi-label, multi-task Convolutional Recurrent Neural Network (CRNN) model that can identify and categorize audio events. The research shows how the suggested method works well for correctly classifying both overlapping and isolated audio occurrences.

G. Fazekas, M. Sandler, K. Choi, and K. Cho's article titled "A Comparison of Audio Signal Preprocessing Methods for Deep Neural Networks on Music Tagging." [4] This study analyzes several audio signal preprocessing techniques to increase the accuracy of music tagging using deep neural networks. The authors look at a number of preprocessing methods, including the wavelet transform, constant-Q transform, and mel-spectrogram. The research offers insights into the efficacy of each technique for music tagging tasks by assessing the performance of deep neural networks with various preprocessing strategies.Fabien Ringeval, Zixing Zhang, Björn Schuller, and Jouni Pohjalainen. The study is titled "Spectral and Cepstral Audio Noise Reduction Techniques in Speech Emotion Recognition." In this study, noise reduction methods for voice emotion identification are investigated. [5] The authors examine several methods to minimize noise and improve the precision of emotion identification systems with an emphasis on spectral and cepstral audio components. The research aids in the creation of more reliable and accurate emotion identification models by examining the effects of noise reduction strategies on speech characteristics.

K. W. Cheuk, D. Herremans, H. Anderson, and K. Agres. A toolbox for on-the-fly GPU audio to spectrogram conversion using 1D convolutional neural networks is called "nnAudio." The authors provide nnAudio, a toolset that allows instantaneous GPU-accelerated conversion of audio to spectrograms. [6] The toolkit offers real-time conversion of audio inputs into spectrograms, which are often employed in different audio processing applications. This is accomplished by using 1D CNNs. Faster training and inference for applications using audio are made possible by the effective and seamless integration of audio preprocessing pipelines into deep learning frameworks made possible by nnAudio. R. Bammer, T. Grill, and M. Dörfler. "Convolutional Neural Networks in Audio Processing: Inside the Spectrogram." The use of Convolutional Neural Networks (CNNs) for audio processing tasks is examined in this research. The authors [7] investigate the use of CNNs for audio event detection and music genre classification. The researchers show how well CNNs capture complicated audio patterns and enhance the efficiency of audio processing tasks by training them using spectrogram representations of audio data. M. S. A. Hashmi and others. The article is titled "Railway Track Inspection Using Deep Learning Based on Audio to Spectrogram Conversion: An on-the-Fly Approach." This paper suggests a deep learning-based strategy for inspecting railroad tracks. The spectrogram representations are analyzed by the authors using deep learning models and audio to spectrogram conversion methods. [8] The suggested method permits on-the-fly examination of railway lines based on auditory inputs by using deep learning algorithms. The study offers a practical and useful approach for inspecting railroad tracks improving maintenance and safety standards.

Using deep learning methods and audio to spectrogram conversion, this research study suggests a revolutionary method for inspecting railroad tracks. The authors automatically evaluate acoustic signals recorded from railroad tracks using the strength of deep learning algorithms. [9] They identify significant characteristics from the audio data and train a deep learning model for track inspection by transforming the audio signals into spectrograms, which are visual representations of sound frequency content across time.

The suggested method enables real-time and on-the-fly processing, allowing precise and effective tracking of anomaly or defect diagnosis. The research illustrates the potential for improving railway maintenance and safety while demonstrating the usefulness of the suggested strategy in identifying different track issues. Using spectrogram analysis and Convolutional Neural Networks (CNNs), the authors of this study describe a unique method for Radio Frequency (RF) fingerprint detection in LoRa (Long Range) communication systems.

[10-13] The authors create a CNN-based model to recognize distinct RF fingerprints by recording and examining the spectrogram of LoRa RF signals, which offers a visual representation of the frequency and time-domain features. Based on their characteristic spectrogram patterns, the proposed solution allows the safe identification and authentication of LoRa devices. The paper emphasizes the approach's potential for boosting the security and dependability of LoRa-based communication networks and shows how successful it is via experimental assessments. [14-16]. In [17], S. Sreethar et al. implemented a cross-layer design in Wireless Sensor Networks (WSNs) using deep learning to overcome the limitations of traditional layered architectures. By sharing information across layers, they improved localization accuracy, energy efficiency, and network performance. In [18], V. E. Jesi et al. focused on enhancing road safety and anomaly detection through sensor technology and machine learning. Their system integrates data from multimodal sensors (cameras, radar, LIDAR) and uses machine learning to detect road anomalies, vehicle issues, and unsafe driving behaviors in real-time.

3. Materials and Methods

The suggested method is a cutting-edge approach to quality control in the production of drill bits. High-definition microphones are used to record audio samples from the drilling site, including both desired and undesired noises. The technology can isolate the sound of the drilling by filtering out other sounds using a number of sophisticated algorithms. The frequency content of the sound signals is then represented by Mel spectrograms using the filtered audio. The spectrograms are analyzed using a pre-trained deep learning model, which divides the drill bit sounds into two groups: good drill bits and damaged drill bits.

The algorithm can precisely detect any indications of drill bit wear by comparing the input spectrograms to the learnt patterns. The system continually observes the drilling operation and conducts the categorization of drill bit sounds, as demonstrated in (Figure 3) in real-time. An instant alarm is created whenever a damaged drill bit sound pattern is found, informing the concerned worker or operator to stop the operation.

This stops the creation of faulty goods and enables prompt action to fix the problem with the broken drill bit. The suggested solution provides a dependable and effective way to improve manufacturing quality by combining high-definition microphones, noise filtering techniques, Mel spectrogram analysis, and deep learning models. It assures the manufacture of error-free final goods, improves accuracy, and decreases the need for human inspection.

3.1. Data Collection

The gathering of a broad and extensive dataset of audio samples from the drilling region is a key initial step in the proposed system. The system's machine learning models are trained and evaluated using this dataset as the basis. High-definition microphones are placed strategically close to the drilling region to record the acoustic signals produced throughout the drilling operation in order to collect the required data. These microphones were chosen with care to provide high-fidelity audio recording, enabling accurate analysis of the drilling noises.

Drilling noises are recorded during the data collecting phase in order to identify the distinctive audio patterns connected to the drilling activity. The typical rhythmic and repeating noises made when the drill enters the material are included in these drilling sounds. The collection also includes a wide range of undesirable sounds that are often found in drilling environments. These sounds might be coolant noises from the flow of cooling fluids used during drilling, as well as rumbling noises from the movement of metal pieces in the equipment.

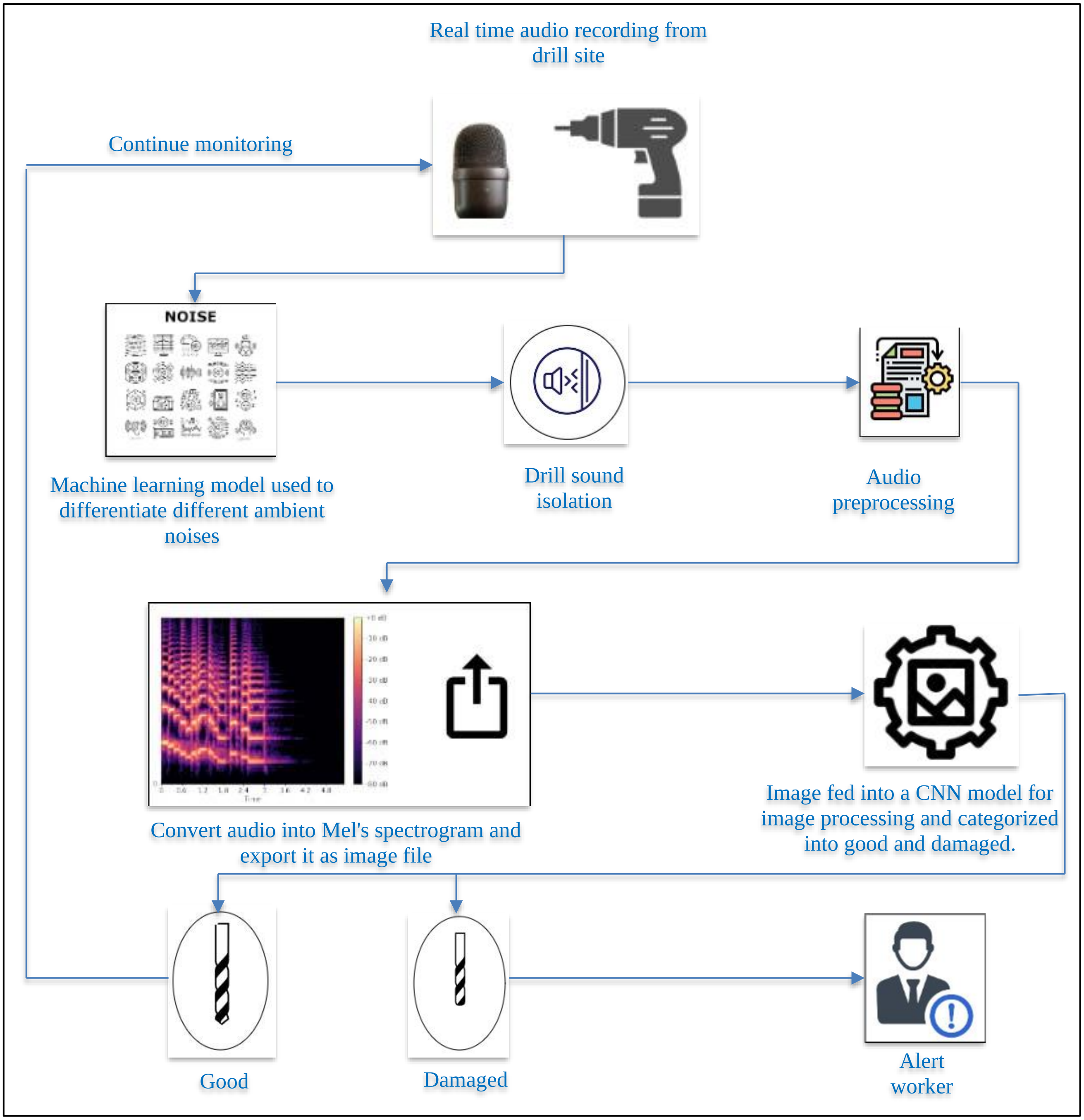

Fig. 3 Architectural representation of the proposed system

To mimic actual settings, ambient noises are also recorded, such as background noise from the surroundings. For the machine learning models to be trained to distinguish between the intended drilling sound and the ambient disturbances, these undesirable sounds must be included in the dataset. The models learn to extract and separate the important drilling sound from the background noise by being exposed to a variety of acoustic changes and interferences. In order to provide a wide and representative sample of drilling and undesired noise recordings, the dataset gathering procedure is carried out thoroughly. In order to account for any changes in the sound patterns, care is made to record a variety of drilling situations, drilling bit types, and environmental variables. The generated dataset serves as the foundation for the next steps of model training, assessment, and validation. The system may successfully learn to distinguish between the intended drilling sound and other acoustic interferences by assembling a vast and varied dataset that includes both desirable noises and undesired noises. This guarantees reliability and precision when spotting drilling faults brought on by broken drill bits, allowing the system to make choices in real-time and take preventative action to stop the manufacture of defective items.

3.2. Training the Noise Filtering Model

The suggested method then develops and trains a machine learning model specially designed to filter out undesirable noises and isolate the drilling sound once the dataset of drilling sounds and unwanted noises has been gathered. This noise filtering model is essential for improving the precision of future classifications and analysis. The gathered dataset is split into two subsets: a training set and a validation set—in order to train the noise filtering model. The validation set is used to evaluate the model's performance and adjust its parameters, while the training set is used to teach the model how to distinguish between the drilling sound and other sounds. The machine learning model used for noise filtering may be based on deep learning or conventional signal processing methods, among other methodologies. Recurrent Neural Networks (RNNs) and Convolutional Neural Networks (CNNs), in particular, have shown outstanding aptitude at processing sequential input, such as audio signals. The model is exposed to a variety of auditory inputs throughout the training phase, including both drilling sounds and undesirable noises. The algorithm learns to recognize and extract features relevant to the drilling sound while ignoring or lessening the influence of undesirable sounds by studying the patterns and characteristics within the dataset. The dataset is labeled to make training more efficient, specifying whether each audio sample contains drilling noise or other undesired sounds. The model is able to create correlations between the audio inputs and their associated categories thanks to this labeled data. The noise filtering model gains the ability to distinguish between the intended drilling sound and the undesirable sounds via an iterative procedure. It modifies its internal settings to enhance its capacity to remove unwanted noise from audio streams.

The validation set is used to evaluate the noise filtering model's performance. To assess the model's capability to accurately detect and isolate the drilling sound while limiting false positives and false negatives, metrics including accuracy, precision, recall, and F1-score are computed. The suggested technique significantly improves the precision and dependability of future studies by training the noise filtering model. The system can concentrate on the crucial acoustic signals throughout the drilling process because the model successfully separates the drilling sound from undesired noise. This improves the system's accuracy and efficiency in spotting drilling mistakes brought on by broken drill bits and lays the groundwork for later categorization and real-time monitoring.

3.3. Pre-processing the Audio Files

In the proposed system for detecting anomalous holes in fuel injector nozzles, preprocessing steps such as converting mono audio to stereo and normalizing sample rates play a crucial role in enhancing model performance. The conversion from mono to stereo is essential as it provides additional spatial information, allowing the system to better distinguish between drilling noises and background interference. This spatial data enriches feature extraction and improves noise reduction, leading to more accurate and robust anomaly detection. Additionally, normalizing sample rates ensures consistency across audio recordings, allowing the model to process data uniformly and effectively. By standardizing the sample rate, the system reduces variability that could affect model training and evaluation, contributing to more stable and reliable performance. Together, these preprocessing steps refine the quality of the input data, significantly enhancing the system's ability to accurately identify defects and anomalies in the production process.

The audio recordings go through a number of preparation procedures before being fed into the system in order to guarantee consistent and uniform input for the deep learning model. These procedures are designed to improve the audio data's compatibility and quality, enabling precise analysis and categorization. Making the mono audio files into stereo is the initial preprocessing step. Stereo audio uses two channels to record a larger audio spectrum than mono audio, which uses a single channel to record sound. The audio data is actually changed into a format suitable for further processing by duplicating the mono audio channel and allocating it to both channels in the stereo file.

The deep learning algorithm can extract more complex information from the drilling sound and undesired sounds by converting the audio recordings into stereo format. The model performs better overall and is better able to discern between various sound patterns thanks to this enlarged audio representation. The audio files' sample rates are standardized in the next stage of preprocessing. The number of samples obtained from the analog audio stream per second during the digitization process is referred to as the sampling rate. The dataset is homogeneous thanks to standardizing the sample rate, which also makes it easier to do reliable analysis. The proposed system's audio files typically have sampling rates that vary from 41,000Hz to 53,000Hz.

This range was selected to preserve a manageable file size and processing efficiency while capturing a broad frequency spectrum. Any possible discrepancies in the audio data resulting from changes in the recording setup or equipment are eliminated by standardizing the sample rate. The system makes sure the audio files are in a consistent and compatible format for subsequent analysis by completing these preprocessing processes. While standardizing the sampling rate provides consistency and compatibility throughout the dataset, the conversion to stereo format widens the scope of the audio representation and captures more granular information. These preprocessing processes help the deep learning model perform more accurately and effectively overall throughout the succeeding analysis and classification phases. The model can recognize and discriminate between drilling sounds and undesired noises with better accuracy and resilience because of the dependable inputs provided by the converted audio files.

3.4. Standardizing Audio Length

The duration of the audio files within the dataset must be standardized in order to preserve consistency and enable effective processing. In this phase, all audio recordings are resized to have the same duration in order to maintain consistency and compatibility for the analysis and classification stages that will follow. Due to variables like drilling speed, material characteristics, or unique drilling conditions, the length of audio recordings may change during the drilling operation. The audio recordings are chopped or padded to a preset length to handle this variance and provide a constant input size for the deep learning model. The audio recordings are modified as part of the standardization process to fit a predetermined time that is chosen based on the specifications of the deep learning model and the features of the drilling sound. Typically, this preset length of time is used to adequately collect the pertinent auditory data and to give sufficient context for precise interpretation. Padding is used to lengthen the duration of audio files whose length is less than the predefined amount. Padding extends the duration of the audio clip without changing the substance of the drilling sound by inserting quiet or low-intensity parts. The padding segments keep the dataset consistent while not contributing to the analysis. On the other hand, if an audio file is longer than the set length, it is cut to fit. Trimming is the process of cutting out the extraneous audio signal while keeping the necessary drilling sound for the allotted time. This makes sure that only the instructive and relevant bits are kept for examination.

The suggested approach provides a fair comparison and consistent processing of the audio data by standardizing the audio duration. It makes sure that the deep learning model is fed inputs of the same length, which speeds up computation and lessens any bias that different audio durations can cause. The system's precision and dependability in identifying drilling mistakes brought on by damaged drill bits are improved by this standardization phase. The deep learning model can efficiently learn and identify the sound patterns associated with good and damaged drill bits by receiving consistent audio inputs, enabling more precise and reliable diagnosis of defective items.

3.5. Generating Mel Spectrograms

The suggested approach then transforms the audio data into Mel spectrograms after standardizing the duration of the audio files. A Mel spectrogram is a graphic depiction of a sound signal's frequency content that shows how the energy is dispersed over time across various frequency bands. The audio recordings go through a procedure known as spectrum analysis to produce Mel spectrograms. In order to extract the frequency components, a Fourier transform is used to split the audio stream into manageable chunks called frames. The resultant spectrogram shows how each frequency bin's intensity or magnitude has changed over time. Traditional spectrograms, on the other hand, fall short of portraying the perceptual qualities of human hearing. As a result, the audio data is converted into Mel spectrograms, which are created expressly to simulate the perception of sound by the human auditory system. Applying a transformation known as the Mel scale transforms a spectrogram into a Mel spectrogram. The Mel scale is a perceptual scale that converts the linear frequency scale to a non-linear scale, highlighting frequency ranges more important to human perception. The suggested approach takes use of the benefits of visual representation by turning the audio data into Mel spectrograms. Mel spectrograms may be regarded as photos, allowing deep learning models based on images to be used for analysis. These models may make use of their talents in sound analysis by utilizing Mel spectrograms as visual inputs. They have shown exceptional effectiveness in a variety of computer vision applications. The capacity of the system to record significant frequency characteristics and patterns that are essential for identifying between the drilling sound generated by a good drill bit and a damaged one is improved when Mel spectrograms are used as input data. The Mel spectrograms' visual aspect enables the deep learning model to assess and interpret the audio data more logically and thoroughly.

The suggested method achieves improved accuracy and reliability in categorizing the drilling sound by using the strength of image-based deep learning models. Mel spectrograms' visual representation helps the model to catch tiny frequency content fluctuations and fine-grained information, leading to a more accurate diagnosis of drilling mistakes brought on by broken drill bits.

Algorithm: Spectrogram Augmentation

Step 1: Define a function, spectro_augment, that takes the following parameters:
- spec: The input spectrogram.
- max_mask_pct: The maximum percentage of the spectrogram to mask.
- n_freq_masks: The number of frequency masks to apply.
- n_time_masks: The number of time masks to apply.

Step 2: Get the shape of the input spectrogram, spec, as (_, n_mels, n_steps).

Step 3: Set the mask_value as the mean value of the input spectrogram.

Step 4: Initialize aug_spec as the input spectrogram, spec.

Step 5: Calculate the freq_mask_param as max_mask_pct * n_mels.

Step 6: Apply frequency masking n_freq_masks times: a. Use a transformation function FrequencyMasking that takes freq_mask_param and mask_value as arguments to mask the frequency components of aug_spec. b. Update aug_spec with the masked spectrogram.

Step 7: Calculate the time_mask_param as max_mask_pct * n_steps.
Step 8: Apply time masking n_time_masks times: a. Use a transformation function TimeMasking that takes time_mask_param and mask_value as arguments to mask the time components of aug_spec. b. Update aug_spec with the masked spectrogram.
Step 9: Return the augmented spectrogram, aug_spec.

3.6. Training the Drill Bit Classification Model

A pre-trained deep learning model is used to properly categorize the audio signals into two groups: good drill bit sounds and damaged drill bit sounds. A Convolutional Neural Network (CNN), which is well-known for its efficiency in image classification tasks, is a preferred option for this purpose. The Mel spectrogram pictures are sent to the CNN as inputs during the training phase. These Mel spectrogram pictures were created by using the method previously mentioned for the audio samples. The CNN is made to examine and extract important details from these pictures, allowing it to understand the differences between healthy drill bit sounds and damaged drill bit sounds. A large collection of labeled instances of both good and damaged drill bit sounds is needed prior to training the algorithm.

These labelled samples act as the model's starting point for learning. A broad variety of sound changes reflecting various drilling situations and circumstances should be included in the dataset. A large dataset that most likely included a variety of sound patterns from multiple sources was used to train the pre-trained CNN. Transfer learning is used to modify it, particularly for classifying drill bits. The pre-training phase's learnt characteristics and information may be used in the present job of classifying drill bits by the model thanks to transfer learning. The pre-trained CNN is adjusted on the dataset of Mel spectrogram pictures throughout the training phase. To fine-tune the CNN's performance for the particular classification job, it updates the layers' weights and biases. Through this approach, the model is able to learn the distinctive sound patterns connected to both healthy and damaged drill bit noises. By examining the characteristics collected from the Mel spectrogram pictures, the CNN eventually learns to distinguish between the two sound patterns as training goes on. In order to increase its capacity to effectively categorize the audio signals, the model modifies its internal parameters. Several measures, including accuracy, precision, recall, and F1 score, are used to assess the drill bit classification model's performance. Usually, a separate validation dataset that is different from the training dataset is used for the assessment. This guarantees a fair evaluation of the model's effectiveness. The suggested approach achieves a high degree of accuracy in recognizing the sound patterns associated with good and damaged drill bits by training the deep learning model exclusively for drill bit classification. Making accurate and trustworthy predictions, thanks to CNN's capacity to assess the characteristics retrieved from the Mel spectrogram pictures, ensures that the production process may be stopped right away upon discovering a defective drill bit.

3.7. Real-Time Monitoring and Classification

The proposed system makes use of the capabilities of high-definition microphones and sophisticated machine learning models to enable real-time monitoring and categorization of the drilling process. The following is how the system works: High-definition microphones are positioned strategically close to the drilling region during the actual drilling procedure. These microphones record the acoustic signals produced during drilling, including the drilling sound, as well as different undesirable noises, such as coolant sounds, the rattling of metal components, and ambient noise.

The noise filtering model, which has been trained to distinguish between the drilling sound and the undesired sounds, processes the collected audio signals after that. This model successfully isolates the drilling sound by filtering out the undesirable noise components. The device concentrates just on the sound associated with the drilling operation by cutting out the disturbance brought on by other sounds. The filtered audio is then converted into a Mel spectrogram, which only includes the drilling sound. The audio data is transformed into visual pictures that correspond to the frequency content of the sound stream.

The Mel spectrogram accurately depicts the frequency components of the drilling sound and their temporal distribution. The pre-trained deep learning model is given the Mel spectrogram picture, and it is taught particularly to categorize the drilling sound into two groups: good drill bit sounds and damaged drill bit sounds. To provide precise predictions, the deep learning model uses the learnt characteristics from its training on a large collection of labeled samples. The Mel spectrogram picture is analyzed by the deep learning model to identify whether the sound is associated with a healthy drill bit or a damaged one. It uses advanced pattern recognition algorithms to detect minute variations and irregularities in the drilling sound's frequency patterns.

The model produces an output reflecting the state of the drill bit based on its categorization. Throughout the drilling operation, the real-time monitoring and categorization process is ongoing. The system continuously records acoustic data, uses a noise-filtering model to separate the sound of drilling, transforms the audio into Mel spectrograms, and then feeds the Mel spectrograms into the deep learning model for categorization. This enables the quick detection and attribution of any drilling mistakes brought on by broken drill bits.

The suggested system provides a number of advantages by integrating real-time monitoring and categorization into the production process:

3.7.1. Less Workforce Needed

The technology does not need manual inspection or human interaction to detect drilling problems. This saves time and costs by reducing the need for a team of people to manually verify each product for flaws.

3.7.2. Reliable Product Quality

The system for in-process monitoring and categorization guarantees that only goods created with high-quality drill bits move on to the next step in the production process. The technology stops the production of faulty or subpar goods by quickly spotting and stopping them in the event of a broken drill bit.

3.7.3. Improved Reputation and Business Opportunities

The business using the suggested approach may boost its standing within the sector by showcasing a dedication to quality control and dependable production procedures. Increased consumer and possible partner trust may result from this, making it easier to close large agreements and forge lasting commercial ties.

3.8. Alert Generation and Process Control

In the proposed system, preserving product quality and avoiding the creation of faulty goods are mostly dependent on alert generation and process management. To do this, the system includes the following steps:

3.8.1. Real-time Sound Analysis

The technology constantly analyzes the drilling sound patterns as the operation moves forward. It distinguishes between excellent and damaged drill bit sounds by using high-definition microphones and sophisticated machine learning algorithms.

3.8.2. Damaged Drill Bit Sound Pattern Detected

If a damaged drill bit sound pattern is discovered during real-time analysis, the system immediately sends out an alarm. This notice is produced based on the deep learning model's classification output, which has been taught to distinguish between the sounds of a good drill bit and one that has been damaged.

3.8.3. Alert generation

The warning is especially addressed to the employee or operator in charge of monitoring the drilling operation. Visual indications, auditory alerts, or messages sent to the operator's workstation or mobile device are just a few of the several ways it may be created. The alarm is intended to catch the operator's eye and demand quick action. Stop drilling: Upon getting the warning, the concerned employee or operator is told to immediately stop drilling. This guarantees that the manufacture of flawed goods is halted right once, preventing any further processing using a broken drill bit. Preventing the production of defective goods: By stopping the drilling operation when it hears the sound of a damaged drill bit, the system successfully stops the production of defective goods. By taking a proactive approach, you may reduce the amount of damaged products, waste, and possible consumer unhappiness.

3.8.4. Process Management and Resolving

The system starts a process control mechanism as soon as the drilling operation is stopped. Maintenance workers or supervisors who are in charge of fixing the problem with the damaged drill bit may be involved in this. Before starting up the production again, the broken drill bit may be replaced and any required maintenance or alterations can be done. Continuous Feedback and Monitoring Throughout the drilling operation, the real-time monitoring and alarm-generating system is still in use, giving constant input and allowing prompt action anytime a broken drill bit is discovered. This cycle of iterative feedback ensures that quality control procedures are followed consistently, which raises the proportion of error-free goods. The suggested method improves quality control and reduces the production of faulty goods by integrating alert generation and process management into the system. The rapid alert generation allows for quick operator response, limiting the production of defective finished goods and lowering the total cost of quality problems. The proactive approach to process management and real-time monitoring helps the manufacturing organization build a solid reputation while also increasing customer satisfaction and product dependability.

The proposed system offers a range of innovative features that make it a valuable tool for fuel injector nozzle manufacturers, setting it apart from existing practices. It automates error identification by using high-resolution microphones and machine learning models to detect drilling mistakes caused by broken drill bits, eliminating the need for manual inspection and significantly reducing labor costs. With real-time monitoring and prompt alerts, the system quickly identifies faults and immediately alerts operators to halt operations, preventing the production of defective goods and saving time and resources.

This approach improves product quality by accurately detecting drilling errors, thereby reducing the occurrence of defects that could compromise performance or safety, ultimately enhancing customer satisfaction. The system's automated nature streamlines the manufacturing process, reducing delays associated with manual inspections and increasing overall productivity and efficiency. By employing advanced technologies, such as high-resolution microphones and machine learning, the system demonstrates a commitment to innovation, enhancing the company's reputation as a leader in the industry. Furthermore, the system provides cost savings by reducing expenses related to labor, rework, and customer returns, offering a competitive advantage and making the company a more attractive partner for major contracts and collaborations.

4. Results and Discussion

Compared to the conventional manual inspection method that relies on human labor, the proposed drill bit quality control system, which integrates high-resolution microphones and machine learning models, offers several advantages. The proposed system demonstrates superior accuracy due to its use of advanced algorithms and deep learning models in analyzing and classifying drill bit sounds. Trained on a comprehensive dataset, it can detect subtle differences and patterns associated with damaged drill bits, whereas manual inspection is prone to human error and subjectivity, leading to inconsistencies and misinterpretations.

The proposed system also includes real-time monitoring capabilities that allow for the immediate detection of issues related to damaged drill bits, enabling rapid response and minimizing the production of defective goods. (Fig.4-5) This proactive approach helps reduce waste and unnecessary rework. In contrast, manual inspection relies on periodic checks, which may overlook defects and result in faulty products reaching the market. Additionally, the proposed system significantly reduces the need for manual labor by automating the inspection process, eliminating the requirement for a dedicated team to manually examine each drill bit.

This not only lowers labor costs but also frees up resources for other critical industrial tasks. Manual inspection, on the other hand, is time-consuming and labor-intensive, which can be both costly and inefficient. Overall, the proposed system surpasses the traditional manual inspection method in terms of reliability, efficiency, and cost-effectiveness. It offers enhanced precision, real-time monitoring, reduced labor dependence, and improved quality control. By implementing this system, manufacturers can ensure the production of reliable and high-quality final products while optimizing operational efficiency shown in Tables 1-3.

Table 1. Summarizing the results and discussion comparing the proposed method

Feature	Proposed System (Machine Learning + High-Resolution Microphones) - DH-FINM	Conventional Manual Inspection	Discussion/Advantages
Accuracy	High accuracy due to advanced algorithms and deep learning models.	Prone to human error and subjectivity, leading to inconsistencies.	Superior precision in detecting subtle differences and patterns in drill bit sounds.
Real-Time Monitoring	Provides immediate detection and response to damaged drill bits.	Relies on periodic checks, which may delay fault detection.	Minimizes waste, reduces rework, and prevents defective products.
Labor Requirement	Significantly reduced due to automation.	High labor demand for manual examination of each drill bit.	Lowers labor costs and reallocates resources to other essential tasks.
Cost-Effectiveness	More cost-effective by reducing labor costs and rework expenses.	Costly due to high labor costs and inefficiencies.	Reduces overall operational costs, improving profitability.
Operational Efficiency	High efficiency due to automation and continuous monitoring.	Low efficiency due to time-consuming manual processes.	Enhances production speed and reduces downtime.
Quality Control	Improved quality control with consistent monitoring and analysis.	Inconsistent due to human factors and periodic inspections.	Ensures higher standards in final product quality, enhancing customer satisfaction.
Reliability	High reliability through consistent and accurate detection.	Lower reliability due to variability in human judgment.	Strengthens trust in product quality and manufacturing process.
Error Reduction	Minimizes error rates by detecting faults early in the process.	Higher error rates due to delayed detection and human mistakes.	Decreases the likelihood of defective products entering the market.
Waste and Rework	Reduces waste and unnecessary rework by preventing defective output.	More waste and rework due to late identification of defects.	Enhances sustainability and resource management.

Table 2. Comparative analysis of parameter accuracy and speed

Aspect	Proposed System	Cheuk et al. (2020) [6]	Dorfler et al. (2017) [7]	Hashmi et al. (2022) [8]	Ouyang et al. (2019) [9]	Shen et al. (2021) [10]	Cheng et al. (2017) [11]
Accuracy	98.50%	93%	94%	96%	95%	93%	90%
Speed ms	1	1.5	2	2	1.8	2.5	2

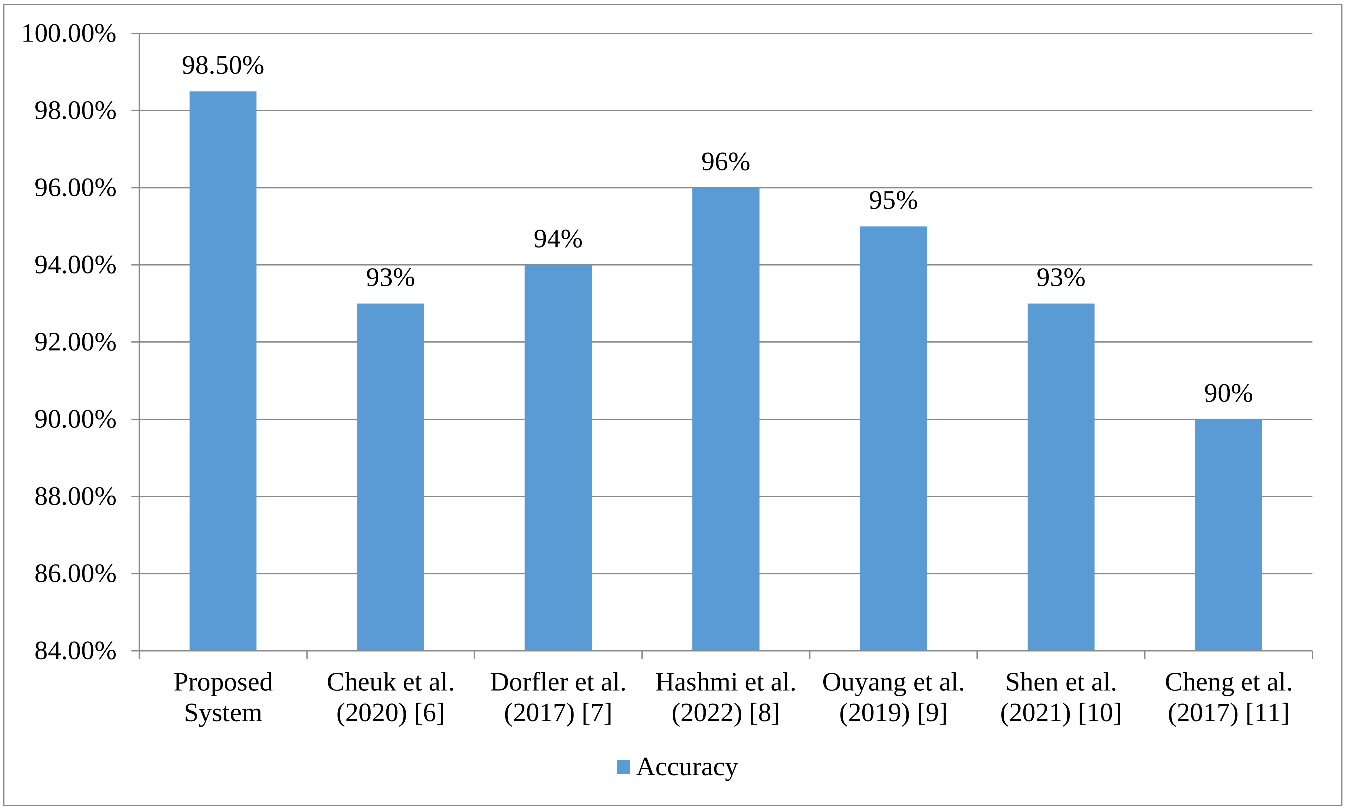

Fig. 4 Comparative analysis of parameter accuracy

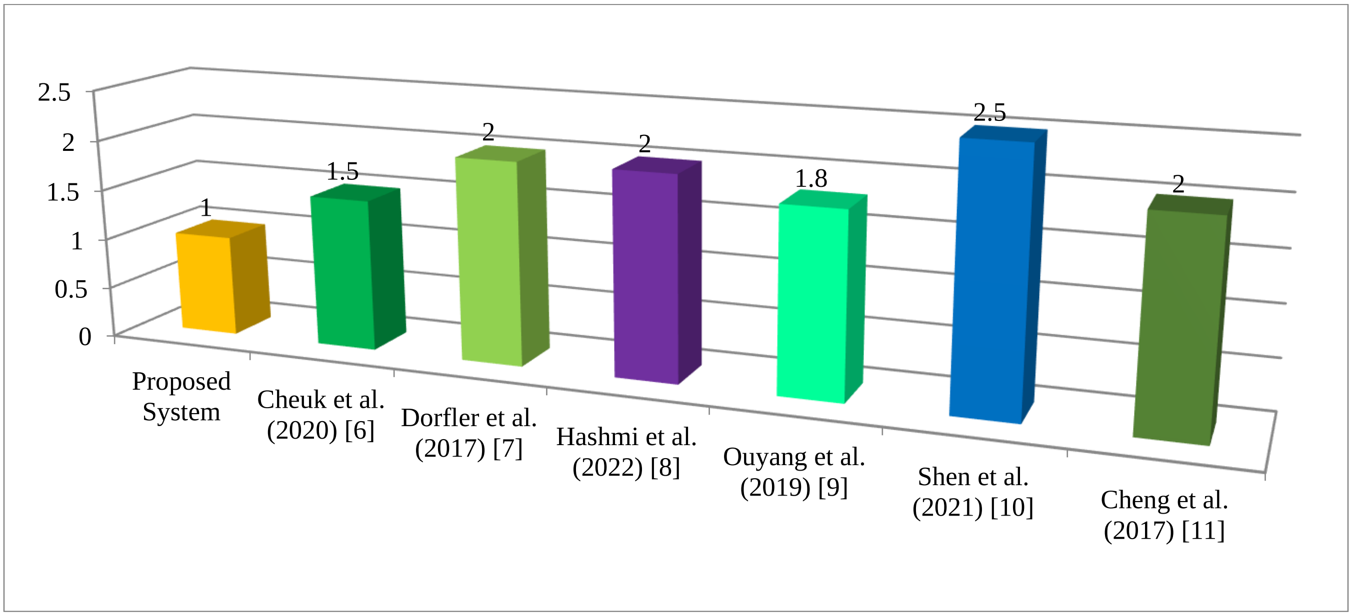

Fig. 5 Comparative analysis of time complexity (ms)

Table 3. Comparative analysis proposed system with existing methods

Aspect	Proposed System DH-FINM	Cheuk et al. (2020) [6]	Dörfler et al. (2017) [7]	Hashmi et al. (2022) [8]	Ouyang et al. (2019) [9]	Shen et al. (2021) [10]	Cheng et al. (2017) [11]	Maccagno et al. (2021) [12]	Cheng et al. (2019) [13]
Primary Objective	Real-time detection of anomalous holes in fuel injector nozzles using audio and machine learning.	Real-time GPU-accelerated conversion of audio to spectrograms.	Application of CNNs for processing audio spectrograms.	Real-time inspection of railway tracks using audio to spectrogram conversion.	Enhancement of speech using CNNs for spectrogram processing.	Identification of RF fingerprints using spectrograms and CNNs.	Analysis of construction equipment activity using audio and SVM.	Audio classification in construction sites using CNNs.	Analysis of construction operations using audio signals.
Technology Used	High-definition microphones, machine learning models, noise filtering, Mel spectrograms.	1D CNNs for audio to spectrogram conversion.	Convolutional neural networks (CNNs) for audio data.	Deep learning models for real-time audio to spectrogram conversion.	Fully convolutional neural networks (CNNs) for speech enhancement.	Spectrogram analysis with CNNs for RF fingerprint identification.	Audio signals and support vector machines (SVMs) for equipment activity analysis.	Convolutional neural networks (CNNs) for audio classification.	Audio signals for construction operations analysis.
Real-Time Capability	Yes, real-time monitoring and anomaly detection.	Yes, real-time spectrogram conversion.	Not specifically real-time; focuses on spectrogram processing.	Yes, real-time track inspection using spectrograms.	Not specifically real-time; focuses on speech enhancement.	Not specifically real-time; focuses on RF fingerprint identification.	Not real-time; focuses on activity analysis.	Not real-time; focuses on classification accuracy.	Not real-time; focuses on audio-based analysis accuracy.
Accuracy and Precision	High accuracy in detecting drill bit anomalies.	High accuracy in spectrogram conversion, not for anomaly detection.	High accuracy in processing audio spectrograms.	High accuracy in track inspection with spectrograms.	High accuracy in speech enhancement using CNNs.	High accuracy in RF fingerprint identification.	High accuracy in activity analysis with SVMs.	High accuracy in audio classification with CNNs.	High accuracy in audio-based construction operations analysis.
Labor Requirements	Significantly reduced, automated inspection process.	Reduced labor for spectrogram conversion.	Requires manual intervention for CNN training and application.	Reduced labor with automated inspection.	Reduced labor for spectrogram processing.	Reduced labor for RF fingerprint analysis.	Requires manual data analysis and SVM training.	Requires manual configuration and CNN training.	Requires manual setup and audio analysis.
Cost Efficiency	Cost-effective by reducing labor and minimizing defects.	Cost-effective in spectrogram conversion but not directly related to inspection.	Cost-effective for audio processing but not related to nozzle inspection.	Cost-effective for track inspection but not directly applicable to nozzle inspection.	Cost-effective in speech enhancement but not related to nozzle inspection.	Cost-effective in RF identification but not related to nozzle inspection.	Cost-effective for equipment activity analysis but not applicable to nozzle inspection.	Cost-effective for audio classification but not related to nozzle inspection.	Cost-effective for construction operation analysis but not applicable to nozzle inspection.
Operational Efficiency	High efficiency with automated and real-time processing.	High efficiency in spectrogram conversion.	High efficiency in spectrogram processing but not related to operational inspection.	High efficiency in real-time track inspection.	High efficiency in processing for speech enhancement.	High efficiency in RF fingerprint identification.	High efficiency in activity analysis with SVMs.	High efficiency in audio classification with CNNs.	High efficiency in audio analysis for construction operations.

5. Conclusion

The proposed system for real-time monitoring and classification of drilling noises utilizes high-definition microphones and machine learning algorithms to address challenges faced by fuel injector nozzle manufacturers. By leveraging advanced technology and sophisticated algorithms, the system offers significant improvements over traditional manual inspection methods. It excels in processing and classifying data by effectively distinguishing drilling sounds from background noise through a noise filtering model. Mel spectrograms, processed by a pre-trained deep learning model, enhance classification accuracy and efficiency, enabling precise detection of defective drill bits. This approach reduces labor requirements, streamlines the inspection process, and minimizes human error through real-time monitoring and categorization. The system quickly identifies faulty drill bit sound patterns, issuing alerts to halt operations and prevent defective products, thus ensuring error-free final outputs. Prompt notifications allow operators to take immediate action, conserving time and resources and enhancing overall quality control. This proactive method not only improves the manufacturer's reputation but also establishes the company as a reliable partner capable of delivering high-quality products.

Additionally, the system's real-time analysis of drilling noises boosts operational efficiency, reduces waste, and increases customer satisfaction by preventing the production of defective items. This capability results in cost savings related to rework, repairs, and customer complaints. By integrating high-definition microphones, noise filtering models, Mel spectrograms, and deep learning algorithms, the system represents a significant advancement in production operations. It reflects the company's commitment to quality and customer satisfaction, offering benefits such as reduced labor needs, enhanced product quality, improved reputation, and increased business opportunities.

Author Contribution

Mohanaprakash T A conceptualized the study, developed the methodology, and supervised the project. He led the design and implementation of the machine learning models and played a significant role in writing the manuscript. D. Siva was responsible for data analysis, the design and deployment of the high-definition microphone system, and the development of algorithms. He also assisted in manuscript preparation and editing. J. Jegan handled data collection, preprocessing, and validation processes, contributed to model training and evaluation, and participated in writing and reviewing the manuscript. S. Janagiraman supported the development of software tools for data analysis and model integration, conducted experiments to validate system accuracy, and contributed to the final revision of the manuscript. Therasa M contributed to the literature review and background research, coordinated research activities across different institutions, and supported manuscript drafting and proofreading.

References

[1] Jong-Oh Sun, and Kwang-Joon Kim, "Isolation of Vibrations due to Speakers in Audio-Visual Electronic Devices without Deteriorating Vibration of Speaker Cone," *Journal of Mechanical Science and Technology*, vol. 26, pp. 723-730, 2012. [CrossRef] [Google Scholar] [Publisher Link]

[2] Jingjing Zhang et al., "Isolated Word Recognition with Audio Derivation and CNN," *2017 IEEE 29th International Conference on Tools with Artificial Intelligence*, Boston, MA, USA, pp. 336-341, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[3] Huy Phan et al., "Unifying Isolated and Overlapping Audio Event Detection with Multi-Label Multi-Task Convolutional Recurrent Neural Networks," *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing*, Brighton, UK, pp. 51-55, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[4] Keunwoo Choi et al., "A Comparison of Audio Signal Preprocessing Methods for Deep Neural Networks on Music Tagging," *2018 26th European Signal Processing Conference (EUSIPCO)*, Rome, Italy, pp. 1870-1874, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[5] Jouni Pohjalainen et al., "Spectral and Cepstral Audio Noise Reduction Techniques in Speech Emotion Recognition," *Proceedings of the 24th ACM International Conference on Multimedia*, Amsterdam, Netherlands, pp. 670-674, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[6] Kin Wai Cheuk et al., "nnAudio: An on-the-Fly GPU Audio to Spectrogram Conversion Toolbox Using 1D Convolutional Neural Networks," *IEEE Access*, vol. 8, pp. 161981-162003, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[7] Monika Dörfler, Roswitha Bammer, and Thomas Grill, "Inside the Spectrogram: Convolutional Neural Networks in Audio Processing," *2017 International Conference on Sampling Theory and Applications (SampTA)*, Tallinn, Estonia, pp. 152-155, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[8] Muhammad Shadab Alam Hashmi et al., "Railway Track Inspection Using Deep Learning Based on Audio to Spectrogram Conversion: An On-the-Fly Approach," *Sensors*, vol. 22, no. 5, pp. 1-21, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[9] Zhiheng Ouyang et al., "A Fully Convolutional Neural Network for Complex Spectrogram Processing in Speech Enhancement," *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing*, Brighton, UK, pp. 5756-5760, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[10] Guanxiong Shen et al., "Radio Frequency Fingerprint Identification for LoRa Using Spectrogram and CNN," *IEEE INFOCOM 2021 - IEEE Conference on Computer Communications*, Vancouver, BC, Canada, pp. 1-10, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[11] Chieh-Feng Cheng et al., “Activity Analysis of Construction Equipment Using Audio Signals and Support Vector Machines,” *Automation in Construction*, vol. 81, pp. 240-253, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[12] Alessandro Maccagno et al., “A CNN Approach for Audio Classification in Construction Sites,” *Progresses in Artificial Intelligence and Neural Systems, Smart Innovation, Systems and Technologies*, vol. 184, pp. 371-381, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[13] Chieh-Feng Cheng et al., “Evaluation of Software and Hardware Settings for Audio-Based Analysis of Construction Operations,” *International Journal of Civil Engineering*, vol. 17, pp. 1469-1480, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[14] R. Huber, and B. Kollmeier, “PEMO-Q—A New Method for Objective Audio Quality Assessment Using a Model of Auditory Perception,” *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 14, no. 6, pp. 1902-1911, 2006. [CrossRef] [Google Scholar] [Publisher Link]

[15] V.A. Abramov et al., “On-Board Transmission Quality Assessment Using Short Audio Signal,” *2021 Systems of Signals Generating and Processing in the Field of on Board Communications*, Moscow, Russia, pp. 1-6, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[16] Nallamilli P.G. Bhavani et al., “Real-Time Inspection in Detection Magnetic Flux Leakage by Deep Learning Integrated with Concentrating Non-Destructive Principle and Electromagnetic Induction,” *IEEE Instrumentation & Measurement Magazine*, vol. 25, no. 7, pp. 48-54, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[17] S. Sreethar et al., “Implementation of Cross Layer Design with Localization Techniques in Wireless Sensor Networks Using Deep Learning,” *2024 International Conference on Expert Clouds and Applications*, Bengaluru, India, pp. 607-613, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[18] V. Elizabeth Jesi et al., “Advancements in Sensor Technology and Computer Vision for Enhanced Road Safety and Anomaly Detection using Machine Learning Approach,” *2023 Third International Conference on Ubiquitous Computing and Intelligent Information Systems*, Gobichettipalayam, India, pp. 457-462, 2023. [CrossRef] [Google Scholar] [Publisher Link]