Acoustic Signal-Based Binary Classification for Brick Wall Inspection by Hammering Test

***Abstract*—Bricks are the most popular building material in Nepal. However, aging, earthquakes, and poor construction can damage bricks and weaken walls. Traditional brick wall inspections in Nepal are typically conducted visually or through destructive testing. Visual inspection is not accurate. Destructive testing is not practical. So, we need a more effective and practical method. We propose a hammering test for brick walls as a reliable, non-destructive option. This technique uses only sound from hammer taps. It is simple and low-cost. However, it is difficult for inspectors to distinguish whether the sound is good or bad. To address this, we applied machine learning techniques. We compared Classical ML algorithms like SVM and KNN, and also Neural Net models such as 1D CNN and 2D CNN. We extracted time and frequency domain features, including MFCCs, ZCR, and spectral centroid. These were converted into both tabular features and spectrograms. The sound caused by the hammer consists of impact and vibration. We set up four types of data (original sound, impact only, vibration only, and combination of impact and vibration) for training and testing. As a result, the combination of impact and vibration yielded consistently better classification performance. The accuracy of classical models is over 95.5%. 1D CNN and 2D CNN achieved 96.5% and 97% respectively. To validate this model, we collected sound data from a different wall and had different individuals strike the wall. As a result, we achieved an accuracy of 95.3%. These results suggest that our study could support the use of real hammering test devices for brick wall inspection.**

***Keywords*—Brick Quality Assessment, Non-Destructive Testing, Machine Learning, CNN, MFCC, Spectrogram, Hammering Test**

I. Introduction

Background

Bricks are valuable building materials highly valued for their durability, strength, and cultural significance. Their use dates back to the eighth millennium BC when they used sun-dried clay bricks, advancing to burnt bricks around 4000 BC in Mesopotamia and in famous structures such as Egyptian houses, the Roman Colosseum, and the Great Wall of China [1]. In Nepal, brick masonry is the foundation for native shelters and cultural heritage, especially in the Kathmandu Valley, where mud-moist ash-fired bricks are used to build thousands of homes and temples [2].

The 2015 Gorkha earthquake, as illustrated in Figure 1, revealed the vulnerability of numerous buildings, affecting nearly half a million homes and 250,000 other structures—over 90% of which were unreinforced masonry [3]. Additionally, heritage temples made of brick and timber suffered significant damage due to substandard-quality bricks and non-standard construction practices [4].

Fig. 1. a) Kathmandu Durbar Square, b) Dharahara, and c) Gorkha Durbar after the 2015 earthquake. *Photos by Bibek Shrestha, Parazlaure, and Bidhan Rajkarnikar, respectively; licensed under CC BY-SA 4.0 via Wikimedia Commons.*

Applications of historic bricks in traditional brickwork with complex timber frames are highly susceptible to seismic stress and need reconstruction techniques that are safe and yield heritage conservation [5]. Poor-quality bricks are largely to blame. Unlike large schemes where strict controls have been applied, traditional buildings are more likely to use handmade or irregularly fired bricks. Builders also blend poor-quality bricks and paint over these to conceal flaws like cracks or hollow spans, making visual inspection unreliable, as illustrated in Figure 2. This dictates the use of independent, non-destructive testing.

Fig. 2. Brick wall at Bishnu Devi Temple, Kirtipur, showing both good and Bad bricks masked by fresh paint.

Current inspection methods rely on the tap-sound test or visual inspection, which are subjective and unsuitable for large-scale or heritage applications [6], [7]. Defects would not be detected due to the nature of human hearing and any existing ambient noise [8]. Standard tests like compressive strength are accurate but sample-destroying [9], whereas alternative non-destructive ones like SPR and IRT are costly to do and require specialists [10]. Scalable, autonomous, and accurate non-destructive methods are required.

Acoustic testing via hammer taps offers a potential solution. Traditionally, the quality was evaluated subjectively by sound-based judgment. Recent approaches utilize machine learning to map acoustic signals into frequency features or spectrograms, enabling AI-based classification [11]. This paper proposes an acoustic classification system with engineered features and spectrograms for the rapid, scalable, and non-destructive assessment of brick quality, aiding in manufacturing control as well as in-situ evaluation to safeguard Nepal's brick houses and heritage sites.

Research Questions

- Can acoustic signals from non-destructive hammering tests be effectively used to classify bricks as good or defective?
- Does comparing engineered audio features and spectrograms provide better insight into classification accuracy and generalization to real-world, in-situ brick conditions?

II. Related Works

Several studies have explored audio-based machine learning systems for detecting defects in construction materials. Koike et al. [12] introduced a tile defect detection system using tapping sounds analyzed by traditional ML models, achieving over 95% accuracy. Fukumura et al. [13] proposed a smartphone-based real-time sound classification approach for concrete structures, leveraging 5G connectivity for cloud-based inference. Ito et al. [14] demonstrated the use of deep neural networks and transfer learning to identify flaking in concrete walls, using frequency-domain features and rolling hammer impacts to improve data representation.

Ma et al. [15] applied machine learning to acoustic signal data from composite materials and found that features such as Mel frequency spectral coefficients (MFCC), zero-crossing rate (ZCR), and spectral centroid were effective in identifying internal defects. Wu et al. [16] applied impact-echo analysis combined with ML algorithms for defect detection in brick masonry, further validating the feasibility of sound-based inspection. Again, Fukumura et al. [17] leveraged transfer learning to enhance sound-based defect classification in hammering tests, demonstrating improved accuracy and training efficiency.

III. Experimental Design and Workflow

The proposed system, visualized in Figure 3, uses a specialized pipeline to classify brick quality from hammering sounds. Audio recordings are segmented into impact and vibration phases, from which time-frequency features are extracted for traditional models and a dual-input 1D CNN. Simultaneously, Mel-spectrograms are generated for a dual-input 2D CNN. This hybrid approach leverages both numerical features and spectrogram visuals for accurate Good and Bad classifications.

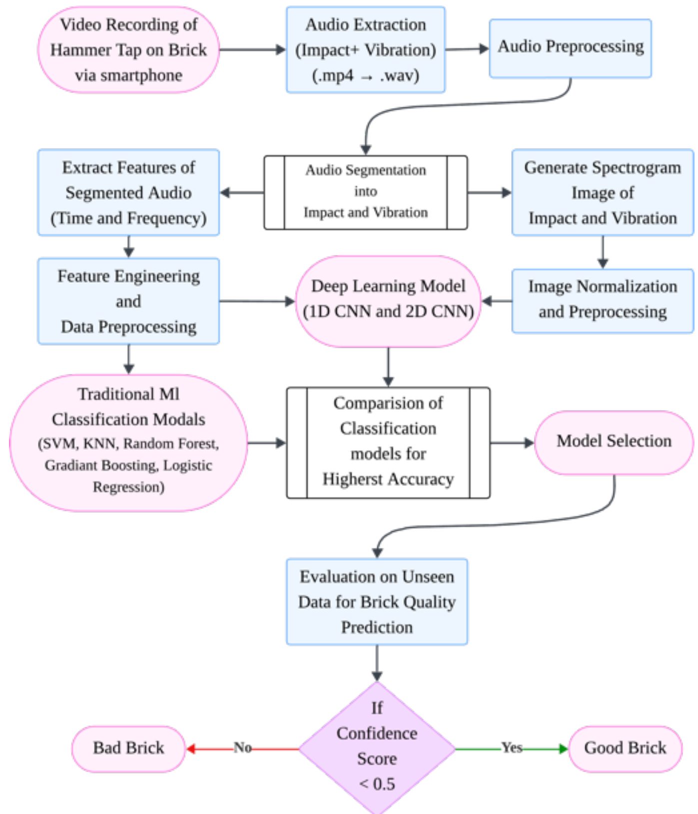

Fig. 3. Block diagram of the brick quality classification pipeline using audio processing, feature extraction, and machine/deep learning models.

IV. Methodology

A. Data Collection

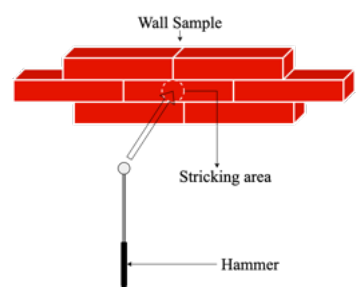

Fig. 4. Schematic of the hammering test setup showing the striking area on a wall-mounted brick sample

Acoustic data were collected from 1007 bricks (examples are given in Figure 6) using controlled non-destructive hammering tests, as demonstrated in Figure 4. A handheld metallic hammer was used for impact, and audio was recorded via a high-fidelity smartphone at 44.1 kHz to capture the full frequency response.

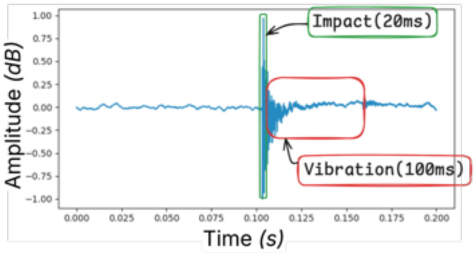

Fig. 5. Time-domain waveform showing the hammer impact followed by vibration, capturing key acoustic features for brick quality classification

To enable deeper analysis of the acoustic characteristics, the recorded signals were segmented into two distinct phases: the *impact phase* (initial strike 20 ms) and the *vibration phase* (subsequent resonant decay next 100 ms), as illustrated in Figure 5. This separation was performed to investigate whether modeling these components independently could provide more discriminative features for brick quality classification. Signals were normalized, denoised if needed, and manually labeled as Good (0) or Bad (1) based on expert assessment and material standards. Table I shows the statistics of the data collected.

Fig. 6. Variety of Brick Sample Collected to Train Classification Model to Predict Good or Bad Brick

TABLE I. Brick Audio Sample Distribution

Category	Count	Impact Samples	Vibration Samples	Total Samples
Good Bricks	507	3111	3111	6222
Bad Bricks	500	3190	3190	6380
Total	**1007**	**6301**	**6301**	**12,602**

B. Data Preprocessing

Following the data collection phase, two distinct approaches were prepared for classification:

- Approach 1: Feature extraction through signal analysis.
- Approach 2: Spectrogram image generation from audio signals.

1) Approach 1(Feature Extraction): All brick impact recordings underwent systematic preprocessing. Stereo signals were converted to mono by averaging the channels and resampled to 48 kHz to retain high-frequency content crucial for distinguishing Good and Defective bricks. Each audio sample was segmented into two phases: *impact* (first 20 ms) and *vibration* (next 100 ms), as shown in Figure 7.

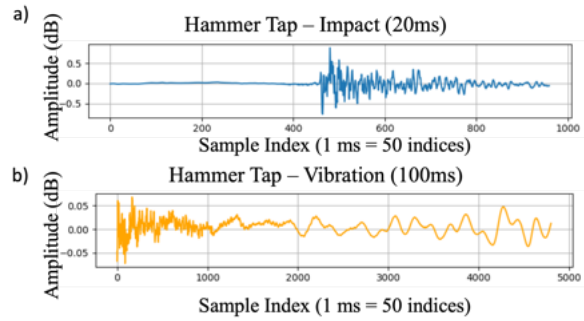

Fig. 7. Hammer tap waveforms showing a) impact (20 ms) and b) vibration (100 ms)

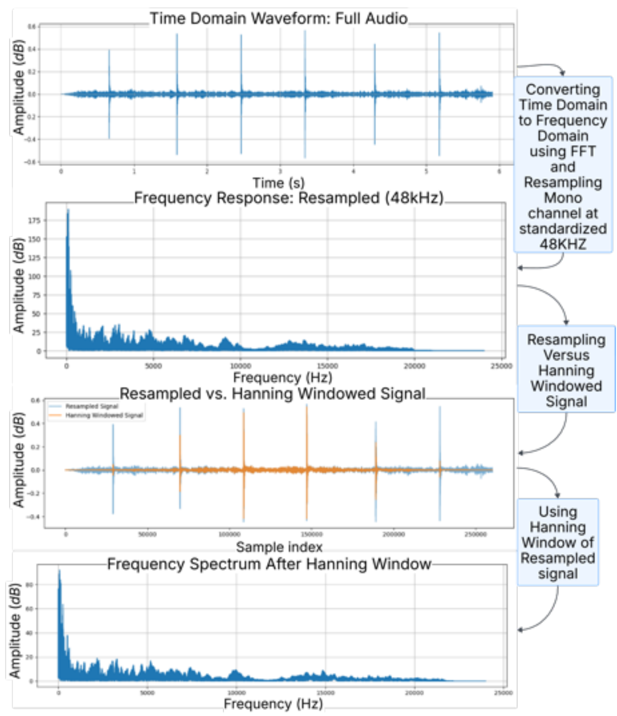

Fig. 8. Signal preprocessing pipeline: (top to bottom) original waveform, frequency response after resampling, windowed signal comparison, and FFT after Hanning window.

FFT-Based Feature Extraction: To minimize spectral leakage and improve frequency resolution, a Hanning window was applied to the signal before transformation. The Fast Fourier Transform (FFT) was then used to convert the time-domain signal into its frequency-domain representation, as shown in Figure 8.

From the resulting spectra, both time- and frequency-domain features were extracted, including spectral centroid, spectral entropy, spectral energy, dominant frequency, peak FFT values, selected FFT coefficients, and Mel-frequency cepstral coefficients (MFCCs).

- MFCCs – Cepstral features derived from the logarithm of Mel-filtered energies.
- FFT Features – Selected frequency bin magnitudes and peak spectral values from the FFT spectrum.

- Dominant Frequency – The frequency corresponding to the maximum spectral amplitude in the FFT spectrum.

These dominant features act as primary indicators for determining whether a brick is Good or Bad.

2) Approach 2: Mel Spectrogram Generation: Each `.wav` file was converted into a Mel-scaled spectrogram, capturing both spectral and temporal characteristics of the brick acoustics, which can be seen in Figure 9.

a) STFT: The signal was segmented into frames and transformed using the Short-Time Fourier Transform (STFT) with a Hanning window to preserve time-localized frequency content.

b) Mel-Scale Mapping: The STFT magnitudes were mapped to the Mel scale using 128 triangular filters, which more closely represent frequency perception in human hearing. This process produced a Mel spectrogram matrix with 128 Mel bands over time.

c) Logarithmic Scaling: To compress the dynamic range and enhance low-energy components, the Mel spectrogram was converted to the decibel scale.

d) Image Generation: The matrix S_{dB} was visualized with the viridis colormap and saved as a 128×128 RGB image using matplotlib, without axes or padding.

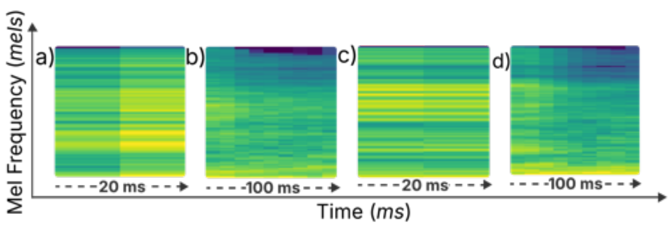

Fig. 9. Sample spectrograms of Good brick [a) impact and b) vibration], and Bad brick [c) impact and d) vibration], and colors show amplitude (*dB*)

C. Classification Frameworks

For classification, two complementary frameworks were utilized based on the two data modalities:

- Traditional Machine Learning: Feature vectors obtained from FFT and MFCCs were fed into classic classifiers, including K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Random Forest (RF), Gradient Boosting (GB), and Logistic Regression (LR).
- Deep Learning:
  – A 1D Convolutional Neural Network (1D-CNN) was employed to classify bricks based on extracted features from the time-series signal.
  – A 2D Convolutional Neural Network (2D-CNN) was trained on Mel spectrogram images to capture spatial and frequency-temporal patterns.

V. Model Development

Before applying learning models, the original brick samples were tested for their physical responses through impact and vibration. The classification task in this study then leveraged both traditional machine learning and deep learning approaches to classify bricks as Good or Bad

A. Comparison of Impact and Vibration Signal Components

To find which type of input sound signal is best to feed as input to models and to evaluate the impact of separating acoustic segments in brick classification, we tested four input strategies across five classical models: (i) *Original(whole) sound* — 507 good and 500 bad samples (5–6s audio clips); (ii) *impact-only*; (iii) *vibration-only* — each with 3111 good and 3190 bad samples; and (iv) *combined impact and vibration* as distinct inputs — comprising 6222 good and 6380 bad samples. As shown in Table II and Figure 10, combining impact and vibration consistently improved accuracy. For example, SVM achieved 0.9556 (vs. 0.95736 for impact-only and 0.92147 for vibration-only), and KNN reached 0.9564 (vs. 0.95400 and 0.92449). These results confirm the complementary nature of impact and vibration signals, supporting their joint use for enhanced classification performance. Thus, for developing both ML and DL models, we used a combined (Hit + Vibration) as the input type and checked performance results.

TABLE II. Model Accuracy Comparison Across Different Acoustic Input Types to Set Baseline for DL Models from ML Models

Model	Original Sound	Hits Only	Vibrations Only	Hit + Vibration
Logistic Regression	0.8812	0.9009	0.8210	0.9310
Random Forest	0.8861	0.9429	0.8175	0.9500
Gradient Boosting	0.8911	0.9358	0.9128	0.9445
Support Vector Machine	0.9059	0.9574	0.9215	0.9556
K-Nearest Neighbour	0.9059	0.9540	0.9245	0.9564

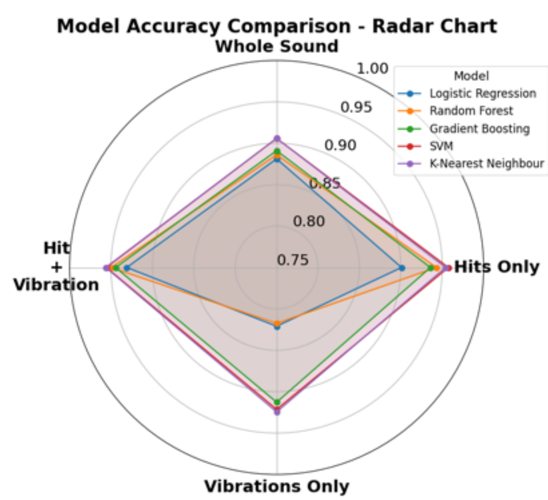

Fig. 10. Radar chart showing model accuracy across four input types. The combined Hit + Vibration input consistently outperforms others, highlighting the complementary value of these inputs.

B. Traditional Machine Learning Models

Traditional models were trained using the numerical features extracted from the FFT and MFCC-based signal analysis.

Supervised learning was adopted using labeled data for accurate classification. Five baseline classifiers were trained on extracted features: Logistic Regression, Random Forest, Gradient Boosting, SVM, and KNN.

The dataset was split 80/20 with stratified sampling to maintain class balance. The results are shown in Figure 11, with KNN achieving the highest accuracy score (0.9564). This indicates its effectiveness among the traditional classifiers.

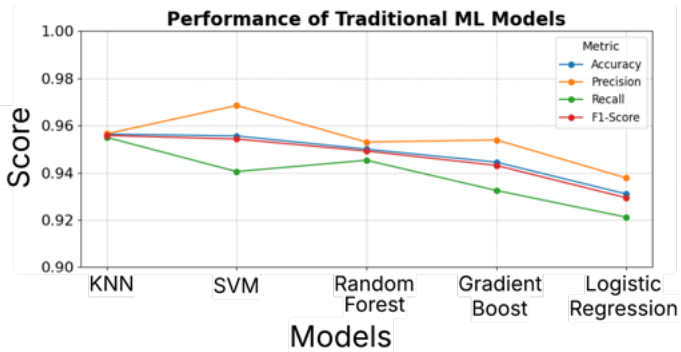

Fig. 11. Performance comparison of traditional ML models.

C. Deep Learning Models

1) Dual-Input 1D Convolutional Neural Network (CNN): A dual-input 1D CNN was developed to process impact and vibration feature vectors ($\mathbf{x}_{\text{imp}}, \mathbf{x}_{\text{vib}} \in \mathbb{R}^{N\times 1}$) simultaneously for improved acoustic analysis.

Architecture: Each branch uses Conv1D–ReLU layers with MaxPooling and Dropout. Outputs are flattened and concatenated:

$$\mathbf{h} = [\text{Flatten}(\mathbf{z}_{\text{imp}}) \,\|\, \text{Flatten}(\mathbf{z}_{\text{vib}})]$$

The combined vector passes through a dense layer with sigmoid activation for binary prediction:

$$\hat{y} = \sigma(\mathbf{w}^\top \mathbf{h} + b)$$

Training: The model used Binary Cross-Entropy loss:

$$\mathcal{L}(y, \hat{y}) = -[y \log(\hat{y}) + (1-y)\log(1-\hat{y})]$$

Training employed Adam optimizer, batch size 32, and early stopping over 50 epochs. It achieved 96.53% accuracy, surpassing all traditional classifiers.

2) Dual-Input 2D Convolutional Neural Network (CNN): To capture both temporal and spectral features, paired spectrograms from hit and vibration data were used in a dual-input 2D CNN architecture, as illustrated in Figure 12.

Spectrogram Preprocessing: Spectrograms were resized to 128×128 and normalized to $[0, 1]$. Paired inputs were matched by filename for consistency.

Model Architecture: Each input is passed through separate CNN branches (Conv2D + MaxPooling), followed by flattening and concatenation for joint feature learning.

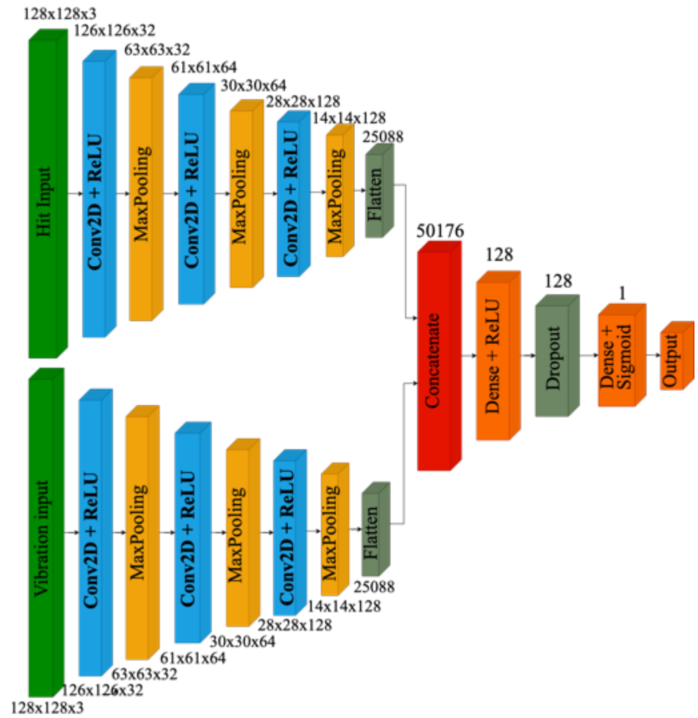

Fig. 12. Proposed Dual-Input 2D CNN architecture processing hit and vibration spectrograms in parallel branches, followed by feature fusion and classification.

Feature Fusion and Classification: Concatenated features from both branches were fed into a dense layer (128 ReLU units), Dropout (0.4), and a sigmoid output for binary classification.

Training Configuration:

- Input: $128 \times 128 \times 3$ RGB spectrograms
- Data: 12,604 samples (6,302 Good / 6,302 Bad), 80% train, 20% val (stratified)
- Training: Adam optimizer, Binary Crossentropy loss, batch size 64, learning rate 1×10^{-4}
- Platform: Google Colab Pro (NVIDIA T4 GPU)

Regularization and Optimization: To prevent overfitting and enhance generalization:

- Dropout (rate = 0.4) was used after the dense layer.
- Early stopping (patience = 10) monitored validation loss.
- Learning rate and batch size were tuned iteratively.

Activations: ReLU for hidden layers: $f(x) = \max(0, x)$ Sigmoid for output: $\sigma(x) = \frac{1}{1+e^{-x}}$

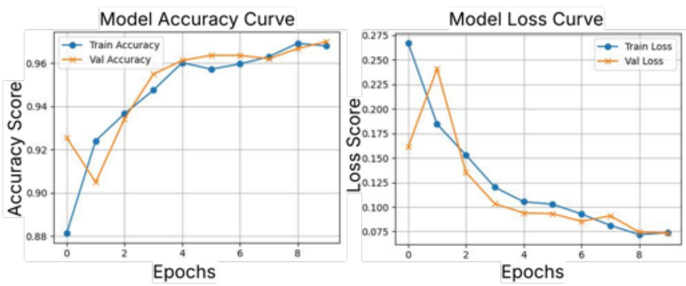

Fig. 13. Training and validation accuracy (left) and loss (right) for the 2D CNN model over 10 epochs.

The dual-input CNN architecture, utilizing Hit + Vibration data, enhances brick quality assessment by processing each signal type in separate branches, allowing filters to specialize in their distinct acoustic patterns before feature fusion. Concatenating these complementary temporal and spectral features provides a richer and more discriminative representation than using either input alone, which is proven by the superior performance of the Hit + Vibration combination over single-input strategies. This design also improves generalization and robustness, as noise in one stream has less impact, while dropout, early stopping, and normalization further prevent overfitting. Moreover, the branch-wise structure supports clearer interpretation of the model's decisions, making it both highly accurate (97% test accuracy) and explainable.

The Dual-Input 2D CNN achieved 97% test accuracy, surpassing all traditional classifiers and the 1D CNN, demonstrating the strength of integrating impact and vibration spectrograms for reliable brick quality assessment.

VI. Model Evaluation

A. Performance Summary

To assess model performance, a composite weighted score was computed to capture overall effectiveness across multiple evaluation metrics. Table III provides a performance summary of the models on the test dataset.

Composite Weight Score: To ensure fair evaluation, a Composite Score was calculated by equally averaging Accuracy, Precision, Recall, and F1-Score (each weighted 0.25):

$$\text{Composite Score} = \frac{\text{Accuracy} + \text{Precision} + \text{Recall} + \text{F1}}{4}$$

This equal weighting avoids bias toward any single metric, providing a holistic view of model performance. It is particularly suitable in this exploratory phase, where both false positives and false negatives are treated with equal importance.

B. Performance on best on 2D CNN

Performance results highlight the 2D CNN as the top model, which we validated using unseen real-world data.

TABLE III. Performance Comparison of Models Using Hit + Vibration Input on Test Set with Standard Evaluation Metrics

Model	Accuracy	Precision	Recall	F1-Score	Comp-Score
2D CNN	0.9700	0.9700	0.9700	0.9700	0.9700
1D CNN	0.9653	0.9604	0.9712	0.9658	0.9657
KNN	0.9564	0.9565	0.9550	0.9558	0.9559
SVM	0.9556	0.9685	0.9405	0.9543	0.9547
Random Forest	0.9500	0.9530	0.9453	0.9492	0.9494
Gradient Boost	0.9445	0.9539	0.9325	0.9431	0.9435
Logistic Reg.	0.9310	0.9378	0.9212	0.9294	0.9299

1) Validation Data Collection (Unseen Data for Model Test): To validate the performance of our best model, the 2D CNN, we used a separate test dataset not involved in training, given in Table IV. This unseen data was collected to assess the model's generalization ability. Good samples came from intact wall assemblies with confirmed sound bricks and from loose, unused bricks struck on a sand bed to capture clean impact vibrations. Defective samples were taken from in-wall bricks, as loose defective ones were unavailable or hard to identify. A licensed civil engineer verified all defective bricks to ensure accurate labeling.

TABLE IV. Validation Data Distribution for 2D CNN Model

Group	Category	Count	Impact Samples	Vibration Samples	Total Samples
Good	Loose Brick	37	208	208	416
	Wall Brick	15	101	101	202
Bad	Wall Brick	50	305	305	610
Total		**102**	**614**	**614**	**1,234**

2) Model Validation: Dual-Input 2D CNN: The generalization of the 2D CNN was tested on a held-out 10% validation set (614 samples), achieving 95.3% accuracy. Precision, recall, and F1-scores ranged from 0.95 to 0.96 for both Good and Bad classes, indicating balanced and reliable performance. Table V shows the model validation results.

TABLE V. Classification Report from 10% Unseen Test Data of Total Dataset

Class	Precision	Recall	F1-Score	Support
Good	0.95	0.96	0.95	309
Bad	0.96	0.95	0.95	305
Accuracy	–	–	**0.95**	614
Macro Avg	0.95	0.95	0.95	614
Weighted Avg	0.95	0.95	0.95	614

a) Confusion Matrix Analysis: The Dual-Input 2D CNN achieved 95.3% accuracy on unseen data, with 16 false positives and 13 false negatives, as illustrated in Figure 14. These low errors suggest strong generalization and robust performance despite real-world acoustic variability and labeling uncertainties.

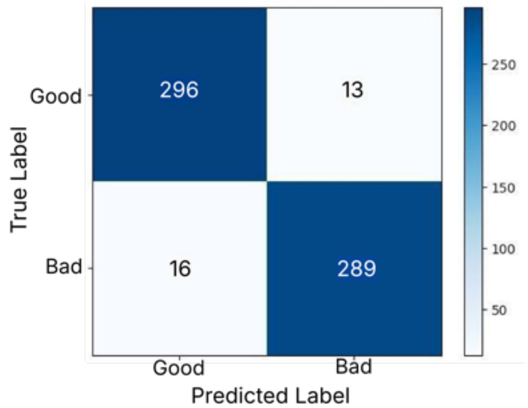

Fig. 14. Confusion matrix of 2D CNN model performance on unseen brick data for validation.

VII. Discussions

This study compared traditional ML models with deep learning approaches for classifying bricks based on acoustic signals. While conventional models, such as KNN, SVM, and Random Forest, performed reasonably well on handcrafted features, they struggled to capture complex patterns in the data. In contrast, deep learning models—especially CNNs using spectrograms—significantly outperformed traditional methods by automatically learning rich, time-frequency features. Their superior accuracy and generalization make them more suitable for real-world, noise-prone environments. Although traditional ML offers simplicity and lower computational cost, deep learning proves to be the better choice for robust, scalable, and automated non-destructive testing applications.

VIII. Conclusion and Future Works

The study demonstrates that acoustic signals from hammering tests can be utilized to classify bricks as either good or defective. Our best model, a 2D CNN, achieved 97% accuracy during training and 95.3% accuracy on real-world validation data, performing better than traditional machine learning models. Additionally, comparing audio features and spectrograms showed that spectrograms gave better results and helped the model work well in real conditions. Notably, combining impact and vibration as inputs improved classical model performance, so this approach was used for developing CNN models, confirming the benefit of joint signal utilization. This proves that the method is effective and practical for testing brick quality without causing damage.

Considering the limited 5G and high-speed internet in regions like Nepal, future work will focus on an offline, edge-computing IoT device for real-time, on-site classification without cloud reliance. The model can also be extended to identify defect types and adapted for broader non-destructive evaluation of concrete, bridges, roads, and pavements, providing a practical tool for infrastructure monitoring in remote, low-resource settings.

References

[1] Amjad Almusaed and Asaad Almssad. Introductory chapter: Bricks between the historical usage and sustainable building concept. In Amjad Almusaed and Asaad Almssad, editors, *Masonry for Sustainable Construction*, chapter 1. IntechOpen, Rijeka, 2023.

[2] Shyam Sundar Basukala and Prem Nath Maskey. Seismic vulnerability of traditional masonry building a case study of byasi, bhaktapur. *Journal of Science and Engineering*, 4:24–30, Apr. 2017.

[3] Rohit Kumar Adhikari and Dina D'Ayala. 2015 nepal earthquake: seismic performance and post-earthquake reconstruction of stone in mud mortar masonry buildings. *Bulletin of Earthquake Engineering*, 18(8):3863–3896, Jun 2020.

[4] Dipendra Gautam, Hugo Rodrigues, Krishna Kumar Bhetwal, Pramod Neupane, and Yashusi Sanada. Common structural and construction deficiencies of nepalese buildings. *Innovative Infrastructure Solutions*, 1(1):1, Mar 2016.

[5] Rizwan Azam, Muhammad Rizwan Riaz, Ehtasham Ul Haq, Ayman Shihata, and Mohamed Zawam. Development of quality assessment criteria for burnt clay bricks of different ages based on ultrasonic pulse velocity test. *Buildings*, 12(8), 2022.

[6] Hamidreza Nemati and Ehsan Dehghan-Niri. Bio-inspired robotic tap testing: An innovative approach for nondestructive testing of wooden structures. 02 2024.

[7] Masahito Ishioka, Kazuo Yamamura, Tomohiro Harada, and Tsuneyoshi Sugimoto. Visualization technology for internal defects in refractory walls of furnace plants using non-contact acoustic inspection method. *Mitsubishi Heavy Industries Technical Review*, 61(1), 3 2024.

[8] Henning Kaiser and Vistasp M. Karbhari. Quality and monitoring of structural rehabilitation measures, part 2: review and assessment of non-destructive testing (ndt) techniques. 2002.

[9] Jiaxing Ye, Takumi Kobayashi, Masaya Iwata, Hiroshi Tsuda, and Masahiro Murakawa. Computerized hammer sounding interpretation for concrete assessment with online machine learning. *Sensors*, 18(3), 2018.

[10] Aslam Hussain and Saleem Akhtar. Review of non-destructive tests for evaluation of historic masonry and concrete structures. *Arabian Journal for Science and Engineering*, 42(3):925–940, Mar 2017.

[11] Dipendra Gautam and Hugo Rodrigues. Seismic vulnerability of urban vernacular buildings in nepal: Case of newari construction. *Journal of Earthquake Engineering*, 25, 07 2018.

[12] Masafumi Koike, Masako Saito, Atsushi Ito, and Katsuhiko Hibino. Hammering tests of tiles fixed with organic adhesives with the aid of machine learning. In *2024 Twelfth International Symposium on Computing and Networking Workshops (CANDARW)*, pages 123–129. IEEE, 2024.

[13] Tsubasa Fukumura, Hayato Aratame, Atsushi Ito, Masafumi Koike, and Katsuhiko Hibino. Improvement of sound classification method on smartphone for hammering test using 5g network. *International Journal of Networking and Computing*, 12(2):359–371, 2022.

[14] Atsushi Ito, Masafumi Koike, and Katsuhiko Hibino. Smart non-destractive test of a concrete wall using a hammer. *Int. J. Electron. Telecommun.*, pages 45–50, March 2024.

[15] Sayf Majeed, HAFIZAH HUSAIN, Salina Samad, and Tarik Idbeaa. Mel frequency cepstral coefficients (mfcc) feature extraction enhancement in the application of speech recognition: A comparison study. *Journal of Theoretical and Applied Information Technology*, 79:38–56, 09 2015.

[16] Yanqi Wu, Jian Zhang, Chenhao Gao, and Junde Xu. Internal defect detection quantification and three-dimensional localization based on impact echo and classification learning model. *Measurement*, 218:113153, 2023.

[17] Tsubasa Fukumura, Hayato Aratame, Atsushi Ito, Masafumi Koike, Katsuhiko Hibino, and Yoshihisa Kawamura. An efficient learning method for sound classification using transfer learning for hammering test. In *2020 IEEE SENSORS*, pages 1–4, 2020.