Adaptive Noise Masking of Construction in Indoor Environments

Ian Straits
University of Virginia
Charlottesville, VA
zkp3zg@virginia.edu

Bhavya Boddu
University of Virginia
Charlottesville, VA
kvz9mm@virginia.edu

Nyla Gordon-Crocker
University of Virginia
Charlottesville, VA
nyd6bx@virginia.edu

Greg Zeckman
University of Virginia
Charlottesville, VA
kmt9gm@virginia.edu

1. INTRODUCTION

Due to the increasing student population at universities worldwide and the desire to provide more resources for the student body, construction is a constant presence on university campuses. The University of Virginia (UVA) alone is working on eight construction projects simultaneously that cover diverse usages, such as the Karsh Institute of Democracy at the Ivy Corridor, the fossil fuel-free energy plant in Fontaine Research Park, and the VERVE Apartments at the intersection of Jefferson Park Avenue, Emmet Street, and Stadium Road [1]. Despite the obvious need for new construction to support the growing community, the long duration of construction is often met with criticism. One of the primary concerns of construction is the highly disruptive noise of machinery, heavy equipment, backup alarms, and other on-site sources. This study in particular looks at Olsson Hall on Engineer's Way. A widely accessed hub for all engineering students at UVA, the Link Lab, located on the second floor of Olsson Hall, is often subject to excessive construction noise that has been acknowledged by occupants to disrupt comfort and productivity.

The noise level of a typical large office is about 50 dBA, and construction noises can often increase this to more than 80 dBA [2]. Additionally, the frequency of sound has been found to be a major determinant in how disruptive that sound was perceived; sounds of higher frequency, which are often generated by construction sites, are typically perceived as more annoying and aggressive than lower frequency sounds of the same loudness [3]. Thus, it is no surprise that construction noise is the primary source of acoustic-related complaints in large cities, and the sounds generated from construction can be difficult to predict because they encompass a wide range of sound energy levels [4]. Thus, it presents a unique opportunity for study.

Previous studies examining the effects of various noise levels on cognitive performance have shown that exposure to loud noises at 95 dBA causes significant changes in brain waves in the lobes responsible for attention and managing mental workload [5]. These reductions in brain function can greatly reduce productivity in the workplace. In a similar study, productivity was measured for workers in the automobile industry under various lighting and noise conditions. While lighting conditions resulted in no significant change in human productivity, an inverse relationship was found between noise level and productivity. The study concluded that a noise level below 85 dBA is needed to increase employee productivity [6]. A third source highlighting the effect of airport noise level on airport workers showed a similar result. Excessive noise was shown to increase work stress and decrease work productivity [7].

Disruptive noise exposure not only affects employee comfort and productivity but also may have significant long-term impacts on their physical health. For instance, a 4-year study of male night-shift factory workers found that those exposed to noise levels above 80 dB(A) during 8-hour shifts had a 3.36 times greater risk of developing insomnia compared to those with lower exposure [8]. Similarly, another cross-sectional study of industrial workers exposed to occupational noise found that those with higher exposure were 2.5 times more likely to experience both insomnia and hearing impairment, showing that constant exposure to noise can negatively impact sleep and hearing health [9]. Other studies have found that occupants exposed to persistent occupational noise experience significant increases in systolic blood pressure, heart rate, and white blood cell count, amongst other health effects, which were attributed to noise pressure alone [10].

The pronounced effect of noise on employee comfort and productivity at work calls for a solution. While passive noise control methods, such as soundproof walls, silencers, and enclosures, help block out noise, they are often insufficient to block all frequencies. Construction noises vary widely based on the type of equipment and activity, but passive noise control methods are most effective for high-frequency noises. Thus, several studies have looked at other noise control methods, such as active noise control methods and sound masking, to better control the negative effects of construction noise. Active noise control refers to canceling sound waves by producing a sound that can negate the sound being produced by the noise source, and has been shown to be useful for canceling low to mid-frequency sounds [11]. Another method, sound masking, which involves the application of sounds to alter the perception of the noise source, has also been examined in the context of construction noise [4].

Auditory masking refers to the perception of a target sound (in our case, construction noise) being made harder to hear, or "masked", by another sound, and this could refer to three types: simultaneous masking, forward masking, or backward masking [4]. In this study, we will study the sound recognition framework of such system that analyzes a stream of incoming data to detect changes in the acoustic environment indicative of external construction noise.

The Link Lab is a research space at the University of Virginia where students and faculty collaborate to advance innovation in the field of cyber-physical systems. One of the objectives of this lab is to collect and analyze real-time data and develop resources that can be used in smart cities and promote the construction of smart and healthy buildings. The lab is an excellent location for data collection as it serves as a space that is relatively quiet, with the greatest auditory interferences originating from human speech from meetings within each office space and, during the data collection period, construction noise from the VERVE construction site across the street.

1.1. Problem Statement

How can the intersection between reactive noise classification and white noise activation be implemented in a real space to minimize the impact of construction noises and improve acoustic comfort for room occupants?

1.2. Motivation

Studies have shown that disruptive noise can negatively impact work productivity [5]. In urban areas, construction and excessive street noise continue to disrupt common work environments, making it difficult to focus and communicate effectively.

More specifically, in the Living Link Lab, a collaborative space at the University of Virginia, the construction of a large apartment complex adjacent to Olsson Hall has contributed to unwanted construction noise. Traditionally, external sound is mitigated with soundproofing panels and windows; however, these solutions are often permanent and require a significant capital cost. Additionally, they promote a "one-size-fits-all" mentality by simply dampening the sound without limiting noise variability, which can be a distraction in itself.

This system has the potential to enhance occupant comfort and productivity by mitigating perceived noise bursts from external sources with an actuated white noise generator, thereby providing a consistent background noise that is ideal for maintaining occupant productivity and comfort.

1.3. Project Objectives

1. Record and manually label construction sounds in the link lab using ReSpeaker Mic Array v2.0 by Seeed Studio
2. Train a machine learning model to classify audio based on the presence of construction noise using Librosa Python software
3. Apply the model to a large audio dataset to identify construction and determine the theoretical actuation of a white noise generator

2. METHODOLOGY

The original approach to tackling the problem statement was to utilize real-time time-series sensor data retrieved from the University of Virginia's InfluxDB database, with the idea that we could leverage this existing sound pressure level (spl_a) data to train a model that could predict when acoustic anomalies may occur. However, due to the constraint of spl_a data being solely amplitude and not full waveform audio data, it could not be used to determine the actual source of the sound, as decibels alone are not indicative of construction, and the noise bursts often exist at too high a frequency to be captured every 3 seconds. Figure 1 shows the data stream for a spl_a sensor. Originally, it was assumed that thresholds could be set based on amplitude anomaly detection from spl_a data; however, it was quickly determined that loudness is not the best determinant for sounds that affect acoustic health and comfort in academic settings.

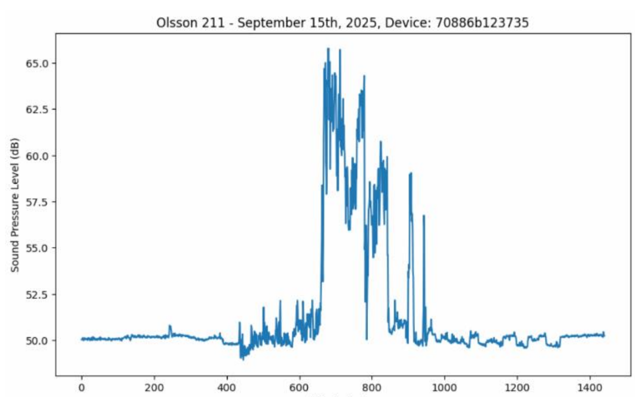

Figure 1. Mean Sound Pressure Level (dBA) for Olsson 211 Id:70886b123735 on September 15, 2025, grouped in one-minute intervals.

The maximum dBA level in Olsson 211 only reached 65 dBA, consistent with normal conversation. The baseline noise level, during periods of inoccupancy, was found to be approximately 52 dBA. All of these values are considered safe, compared to the 75 dBA limit enforced by the World Health Organization [12]. The lack of high frequency data, and variable sound pressure data evident in Figure 1, clearly illustrate why spl_a could not be used as a proxy for construction noise. This further emphasizes the need for sound identification to enhance occupant comfort, rather than sound pressure alone.

Weighing these findings, our updated method leverages the use of waveform audio data (.wav) collected from a ReSpeaker Mic Array v2.0 by Seeed Studio, operating through Raspberry Pi OS. Waveform audio enables us to separate sound data into individual frequencies, allowing us to isolate sounds that may be considered disturbing or unwanted. We developed and trained our model on .wav data collected over a 24-hour period from November 12th to 13th, 2025, beginning at 11:30 am. The sensor was deployed in Olsson 241 inside a personal office, which is less than a tenth of a mile away from the VERVE construction site as shown in Figure 2. Full permission from occupants was obtained prior to recording.

Figure 2. Location of Personal Office (Olsson 241) in relation to the VERVE construction site.

To pre-process the data, the 24-hour waveform recording was split into 10-second chunks and stored in a directory using the makedirs function in Python. 10-second chunks were selected to reduce the manual load for labeling as an initial framework. While shorter windows could add temporal resolution, it would have required extensive manual labeling. The subsequent 10 second audio files were split into two subsections. A manual subsection was created, comprising 20% of the entire recording, to test and train the model. The other 80% of the data was further classified once the model had been trained.

To provide usable data for the model and Librosa to process, the manual dataset was labeled by listening to each audio chunk for acoustic signatures characteristic of construction. For each 10-second chunk, we recorded a 1 if no construction was detected and a 0 if construction was detected. If construction noise was identified at any point during the 10-second chunk, the chunk would be labeled with a 0, regardless of the type of noise. This includes impact tools, power tools, emergency sirens and backup alarms, heavy equipment, and material handling. It is important to note that for periods of talking within the collection room, a 1 was given to teach the model not to classify ordinary conversation as construction, which was otherwise the dominant acoustic source. The observations were recorded in a CSV file, totalling 1619 10-second chunks, or 4.5 hours.

Once the subsection was entirely labeled, a random forest classifier was trained on the 1619 manually labeled data points. A standard 80-20 train-test split was performed on this data with stratification to ensure that the train-test split consisted of the same proportion of each of the classification labels as the original data. A standard scaler was then used to normalize the features for the model, avoiding disproportionate weightage to any one feature. The model was then trained on this scaled training data and predictions were made on the scaled testing data. The outputted classification report and confusion matrix are shown below in Tables 1 and 2.

The model parameters and scaler settings were then saved so that they could be used to predict the existence of construction noise for the remaining 10-second clips (80% of the total data). For each data point, the sound features were first extracted using Librosa and then scaled with the Standard Scaler before prediction was made. The following information was captured and exported to a .csv file for each data point: filename, label, source, and model probability. The completed count of all 10-second audio chunks containing construction noise over the course of the collection period is illustrated in Figure 3 of the results section.

Additional processing was done on these model results to determine when the white noise actuator should be turned on or off. Because the classifier didn't output prediction labels in the order of the audio clips, data was first sorted in chronological order, and a "TOD" column was added to represent the time of day that the audio clip corresponded to. Deciding the ON/OFF settings for the white noise actuator system presents several unique challenges that require a precise balance. For instance, due to potential time latencies in the switch from an OFF → ON or ON → OFF state, switches were minimized for disruptions that occurred in small time windows. In addition, construction noise happens in unpredictable patterns and discrete discontinuous time blocks, so recent probability predictions should be given a higher weightage. Exponential moving average (EMA) and simple hysteresis were utilized to enforce these constraints and come up with state predictions for the actuator.

The following EMA formula was used to calculate the smoothed probabilities:

$$EMA_t = (Prob_t \times k) + (EMA_{t-1} \times (1 - k))$$

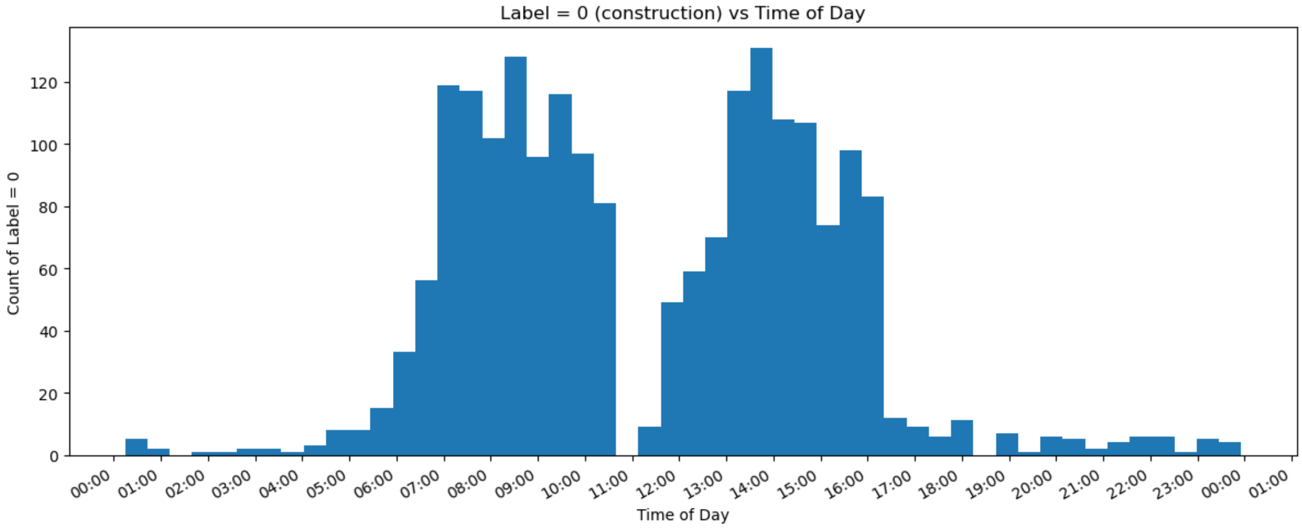

Figure 3. Distribution of construction noise events across the 24-hour data collection period (Nov 12th, 2025, 11:30 am - Nov 13, 2025) grouped into half-hour intervals. Indications of high construction activity are evident throughout the morning and afternoon consistent with standard work hours.

After manually testing several values for the EMA hyperparameter, k, a value of 0.3 was chosen. The small magnitude is associated with an increased emphasis on the more recent data points. Once the smoothed probability is generated for a data point, it is compared with a low threshold and a high threshold value to determine the actuator setting. After several rounds of manual testing, the low threshold value was chosen to be 0.15, while the high threshold value was chosen to be 0.4. If the smooth probability is less than the low threshold, the actuator is turned off. If the smoothed probability is higher than the high threshold, the actuator is turned on. However, if the smoothed probability lies in between the two thresholds, the actuator setting is the same as what it was for the previous data point. The low threshold values reflect the prioritization of true positives over true negatives in the choosing of the actuator setting.

3. RESULTS

The outcomes of model training are represented in two subsections. The manual results pertain to the ground-truth determination made by listening, and the model results are the remaining audio chunks labeled based on the random forest classifier model.

3.1. Manual Results

For the 1619 10-second chunks listened to, we recorded 735 instances where construction was audible (0) and 883 instances where it was not detected (1). Most data were gathered from near the beginning and end of the recording, as these were most in line with hours of the standard workday, when construction would be most present. This was an attempt to capture a similar quantity of 0s and 1s to provide the model.

3.2. Model Results

Using a random forest classification model, the manual data set is then split 80/20 into training and testing sets, respectively. The testing set was evaluated on its performance, outlined in the classification report in Table 1 below.

Table 1. Classification Report of Training Set

	precision	recall	f1-score	support
0	0.7908	0.8231	0.8067	147
1	0.8480	0.8192	0.8333	177
accuracy			0.8210	324
macro avg	0.8194	0.8212	0.8200	324
weighted avg	0.8220	0.8210	0.8212	324

Table 1 indicates that the model correctly predicted construction 79% of the time and non-construction 84% of the time, as indicated by the precision. The recall, on the other hand, examines how many of the detected events it accurately identified within n support values. The f1-score is the harmonic mean of both the precision and recall for each identifier. All of this together led to a set accuracy of 82% for this model. The relative breakdown of these decisions is displayed in Table 2.

Table 2. Confusion Matrix of Training Set

	positive	negative
true	121	145
false	32	26

For the 324 testing points, the model correctly identified 121 construction events and 145 non-construction results. There were also 32 false construction events and 26 false non-construction events.

The trained model then labeled all data over the 24-hour collection period, amassing 8,646 audio chunks. The model predicted 6,662 instances of no construction and 1,983 instances of construction over the 24-hour period. The total occurrences of construction (Label = 0) at a specific time of day (TOD) are shown in Figure 3 above. Construction was most prevalent during the hours of 7:00 am to 11:00 am and 12:00 pm to 5:00 pm.

4. DISCUSSION

The results of this study indicate that a reactive construction noise classification system can be implemented using waveform audio and machine learning based detection. Assuming that construction noise is acoustically distinct from other building/city/occupant-generated sounds, and that 10-second audio segments are an appropriate interval for proper classification, this application could be used in similar settings where construction noise is a major complaint for building occupants.

An overall accuracy of 82% between precision and recall statistics for the classification model is promising, considering the relatively small training and testing sets and that the manually labeled data was labeled by humans with only audio data, without the use of supplementary video validation data. This is especially impressive considering that 10 seconds is a relatively long interval for construction noise identification, since many construction noises often occur in short bursts of less than a second; model precision would likely improve with a larger data set of shorter interval audio chunks. Additionally, the use of manually labeled data introduces subjectivity to the labeler which could have contributed to the 18% of inaccurate data.

Nonetheless, the model yields promising results for identifying construction noises. Using a random forest classifier is useful considering the variable nature of construction noise. Different sounds, including alarms, heavy machinery, and impacts, vary in their acoustic parameters. A forest classifier allows the model to consider multiple decision trees, all without placing undue weight on a specific sound or trend in the data. In our case, the model was able to classify a variety of different construction noises emitting from the VERVE construction site.

In part, this is also due to the audio classification provided by Librosa that pairs with the random forest model. Using Librosa, we were able to extract specific acoustic features from the waveform audio that are characteristic of the specific type of construction that is occurring. These features provide a qualitative representation of each audio chunk, allowing the model to learn the relationships with their classification label. Figure 4 below displays a mel spectrogram illustrating these acoustic patterns and their classification label.

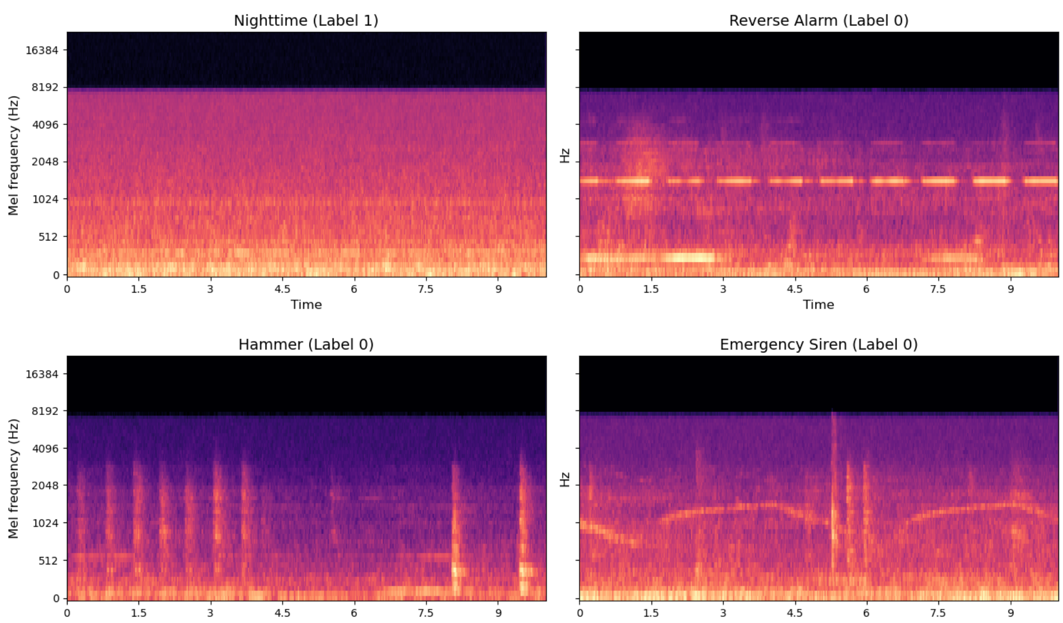

Figure 4. Mel-spectrogram representations of sample audio samples, comparing nighttime ambience (1) to several construction noise types (0), including a vehicle reverse alarm, hammer impacts, and emergency siren beeps.f

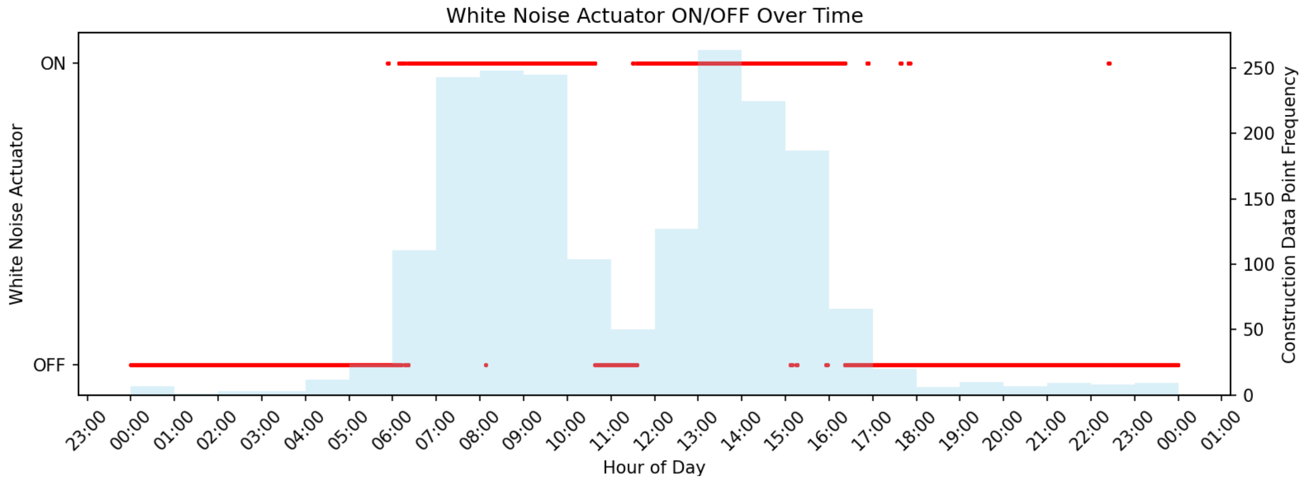

Figure 5. Same histogram as shown in Figure 3, grouped by the hour, overlaid with a plot showing the tuned suggestions for turning off/on the white noise actuator system. The suggestions demonstrate the use of the EMA hyperparameter and hysteresis logic to limit off/on cycles of the actuator, essential for occupant comfort.

The nighttime spectrogram, labeled 1 by the model, shows no variability over the 10-second chunk with a low general amplitude compared to other spectrograms. The other three spectrograms were all labeled as construction by the model, representing a reverse alarm, a hammer, and an emergency siren upon listening to the specific audio chunks. Visual inspection of these spectrograms shows unique patterns related to the construction noise they are attributed to. All three exhibit low-amplitude, high-frequency repetitive noises at around 2000 Hz; conditions that seem to be evident in most construction events, as they originate from outside the building. Using the audio features extracted by Librosa, the model assigned each chunk a probability to designate its classification. The individual model-based probability and smoothed context-dependent probability are presented in Table 3 below.

Table 3. Model-based probability to contain construction

Sound	Model prob.	Smoothed prob.
Nighttime	0.005	0.0615
Reverse Alarm	0.94	0.799
Hammer	0.88	0.797
Emergency Siren	0.86	0.818

Despite all three construction sounds having a generally high model-based probability, the smoothed probability shows a lower value, taking into account instances of no-construction in time proximity to the selected audio chunk.

Utilizing the smoothed probabilities processed by the EMA and hysteresis methodology, Figure 5 displays the suggestions for the actuator settings as either "ON" or "OFF". It is evident that the suggested actuator settings at the various time points directly correlate with the increase in construction noise datapoints in any particular hour. In addition, though the raw labels for the data points, as predicted by the model, were shifting continuously between 1s and 0s, the EMA and hysteresis method greatly reduced the number of transitions from ON → OFF or OFF → ON, which was one of the goals of the actuator system.

This work was completed under the assumption that a consistent white-noise is a preferable alternative to infrequent construction noises by the occupant. The output of the model indicated that the white noise generator would run essentially non-stop during the workday and only turn off during lunchtime or after the construction workers leave for the day. Balancing between responsiveness and stability for the actuator and the psychological effect that differing actuator sensitivities could have on the occupant is an important consideration that has not been accounted for in this study but could be an avenue for future work.

Overall, this study demonstrated the feasibility of using machine learning audio classification to support a white noise generator as a means to mitigate an acoustic environment polluted by construction noise There were clear limitations to dataset size and labeling assumptions, which can be easily expanded on for future work. Nonetheless, the findings suggest that construction noise can be autonomously identified using machine learning, which serves as motivation for the future development of a fully automated dynamic noise masking system.

4.1. Future Work

More work is needed to apply this system to real-world applications. Expanding the dataset to include a more diverse set of audio recordings from different rooms in

Olsson Hall would allow the model to capture a broader spectrum of acoustic conditions. The current dataset only covers one space from a single 24-hour period in Olsson 241, but this does not fully capture the range of intensities and frequencies that are heard throughout the building. Future work could also prioritize the collection of audio from outside the building concurrently with indoor recordings, enabling more accurate labeling by reducing acoustic indoor interference that may lead to misclassifications. Additionally, future refinement of the model could include sliding frame windows instead of 10-second audio chunks for more accuracy, or different actuation parameters optimized to prevent frequent cycles, while maintaining acoustic comfort.

The ultimate goal of this system would be to process audio in real-time. Currently, all processing is reactive, analyzing pre-recorded audio. In the future, the model could process real-time audio recordings so that it can predict construction bursts as they occur and activate the white noise generator autonomously. Evaluating the effectiveness of a white noise generator was beyond the scope of this project but should be immediately tested with the implementation of real time processing, and tested for occupant perception.

5. CONCLUSION

This study demonstrates that machine learning audio classification, when paired with white noise actuation, may serve as a viable approach for mitigating the effects of acoustic disturbances generated from a construction site. Through waveform audio data collection and manual labeling, a random forest classifier was trained on features extracted from Librosa, achieving an accuracy of 82% in construction noise identification. The actuation model correctly signaled the white noise generator to turn on and stay on during times of peak construction activity, and correctly signaled the generator to turn off and stay off during periods of site inactivity. These findings suggest that a reactive system that distinguishes between construction noises and non-construction noises and acts based on this classification is feasible, with room for expansion.

5.1. Acknowledgements

We would like to thank Saja Alnusair for her expertise and guidance throughout this study, which included providing recording materials and contributing to model development. Additionally, we would like to thank Brad Campbell for providing access to his office for data collection.

REFERENCES

[1] Bailey, L. (2025, October 6). *University construction wraps up summer projects, what now?*. University construction wraps up summer projects, what now? - The Cavalier Daily - University of Virginia's Student Newspaper.

[2] Cochary, J. (2021, May 15). *Common noise levels*. Noise Awareness Day.

[3] Rimskaya-Korsakova, L. K., Pyatakov, P. A., & Shulyapov, S. A. (2022). Evaluations of the Annoyance Effects of Noise. *Acoustical Physics, 68*(5), 502–512.

[4] Wu, Z.F. and Zhao, X.Q. (2024) *Reducing construction noise: Sound masking effect on soundscape dominated by construction noise - International Journal of Environmental Science and Technology*, *SpringerLink*, 22, 797-832.

[5] Jafari, M. J., Khosrowabadi, R., Khodakarim, S., & Mohammadian, F. (2019). *The Effect of Noise Exposure on Cognitive Performance and Brain Activity Patterns. Open access Macedonian journal of medical sciences, 7(17), 2924–2931.*

[6] Akbari J, Dehghan H, Azmoon H, Forouharmajd F. (2013). *Relationship between lighting and noise levels and productivity of the occupants in automotive assembly industry. Journal of Environmental and Public Health.*

[7] Ichwan, N.A.A., Saleh, L.M., Thamrin, Y., Wahyu, A., Syafar, M., Mallongi, A. (2024). *The Effect of Noise and Work Fatigue on Work Productivity: Work Stress as a Mediator Among Airport Workers.*

[8] Cho, S., Lim, D. Y., Kim, S., Kim, H., Kang, W., & Park, W. J. (2023). *Association between Occupational Noise Exposure and Insomnia among Night-Shift Production Workers: A 4-Year Follow-up Study.* Noise & health, 25(118), 135–142. https://doi.org/10.4103/nah.nah_15_23

[9] Lim, H. M., Kang, W., Park, W. J., Jang, K. H., Ann, J. S., & Moon, J. D. (2017). *Insomnia and hearing impairment among occupational noise exposed male workers. Annals of occupational and environmental medicine*, 29, 36.

[10] Huang, C.-Y. *et al.* (2024) *Effect of occupational noise on employee health: A longitudinal study*, *Noise & health*, *26(123), 514-522.*

[11] Kwon, N., Park, M., Lee, H.S., Ahn, J., Shin, M. (2016, Jan 27). *Construction Noise Management Using Active Noise Control Techniques.*

[12 World Health Organization. (1999). *Guidelines for community noise. Geneva, Switzerland: World Health Organization.*