LLNL-JRNL-2006716

Lawrence Livermore National Laboratory

Rocket Launch Detection with Smartphone Audio and Transfer Learning

S Popenhagen, M Garces, S Takazawa

August 2025

Sensors

Disclaimer

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

signals

MDPI

Article

Rocket Launch Detection with Smartphone Audio and Transfer Learning

Sarah K. Popenhagen[1,*], Samuel K. Takazawa[2] and Milton A. Garcés[1]

1 Infrasound Laboratory, University of Hawai'i at Mānoa, Honolulu, HI 96822, USA; milton@isla.hawaii.edu
2 Lawrence Livermore National Laboratories, Livermore, CA 94550, USA;
* Correspondence: spopen@hawaii.edu

Abstract: Rocket launches generate infrasound signatures that have been detected at great distances. Due to the sparsity of the networks that have made these detections, however, most signals are detected tens of minutes to hours after the rocket launch. In this work, a method of near-real-time detection of rocket launches using data from a network of smartphones located 10-70 km from launch sites is presented. A machine learning model is trained and tested on the open-access Aggregated Smartphone Timeseries of Rocket-generated Acoustics (ASTRA), Smartphone High-explosive Audio Recordings Dataset (SHAReD), and ESC-50 datasets, resulting in a final accuracy of 97% and false positive rate of <1%. The performance and behavior of the model are summarized, and its suitability for persistent monitoring applications is discussed.

Keywords: rocket launches; smartphones; machine learning; detection; data; infrasound; acoustics.

1. Introduction

Since the launch of Sputnik on October 4, 1957, humanity has successfully launched over six thousand rockets into orbit. As science and technology have advanced in the decades since, the specifications and capabilities of rocket engines have changed and grown while the underlying principles have remained constant. To escape Earth's atmosphere, rockets must be accelerated to great velocities. To achieve this acceleration, propellants are burned in a combustion chamber. The resulting exhaust is the working fluid; it is accelerated through a propelling nozzle and expelled at hypersonic velocities to produce thrust. The nature and state of the propellants used can vary between makes and models of rockets, but these basic principles remain constant.

During the launch sequence of a rocket, acoustic waves are generated by a number of source mechanisms, resulting in a complex acoustic signature. For convenience, we will divide these acoustic waves into three categories: waves generated by the engine (engine noise), waves generated by exhaust (exhaust noise), and waves generated by turbulent flow excitation (jet noise). Engine noise, while the least intense of the three, is well-studied due to combustion instabilities leading to catastrophic engine failure, material fatigue, etc. generating distinct acoustic signatures. Through ignition and liftoff, exhaust noise is the primary source of acoustic energy. As exhaust is expelled from the rocket, it collides with ambient air, creating shock waves. For many rockets, the intensity of these waves is high enough to damage the rocket and/or nearby structures, and thus water-based suppression

Academic Editor: Firstname Lastname

Received: date
Revised: date
Accepted: date
Published: date

Citation: To be added by editorial staff during production.

of exhaust noise has been frequently employed on launch pads since the Space Shuttle program. As the rocket accelerates into supersonic velocities, jet noise generated by the increasing shear flow and resulting turbulent eddies overtakes exhaust noise and becomes the dominant source of acoustic energy.

Like other large-scale events such as earthquakes and explosions, rocket launches generate low-frequency (<300 Hz) sound as well as infrasound (<20 Hz) [1]. These components of acoustic rocket signatures can remain detectable over great distances due to the frequency dependence of atmospheric attenuation. Infrasound sensors are employed by the International Monitoring System (IMS) of the Comprehensive Nuclear-Test-Ban Treaty Organization to detect and monitor large-scale events, including rocket launches [2]. IMS infrasound stations have successfully collected signatures from a variety of events, including numerous rocket launches [3], but the sparsity of the network necessitated by its global coverage and the expense of traditional infrasound sensors results in these signatures being collected only after propagating vast distances. The result is that detections can only be made long after a rocket has launched, at which point much of the information the signature once carried about its source has been lost to attenuation.

Rapid detection of rocket launches may be possible, however, through the use of non-traditional sensors such as smartphones. Smartphone microphones have been used to successfully collect infrasonic and low-frequency signatures [4,5], and smartphone audio data has been used to train machine learning models to accurately detect explosion signatures [6–8] in previous studies. In this work, we present a machine learning model designed and trained to detect acoustic rocket signatures in smartphone audio data from three open-access datasets using the machine learning method transfer learning. In the following sections, we will detail the methods used, evaluate performance under different conditions, and discuss the suitability of the solution for persistent monitoring.

1.1. Transfer Learning with YAMNet

Transfer learning is a machine learning method first developed by Stevo Bozinovski and Ante Fulgosi at the University of Zagreb in 1972 [9,10]. In theory, the technique improves learning efficiency by reusing knowledge gained from one task for a second, related task. The popularity of transfer learning has risen in recent years [6,11–14] due in part to the existence of publicly available models pre-trained on very large datasets for general tasks which can be used to boost performance on a related, specific task if the volume of available training data for the specific task is limited [15].

One such publicly available model is Google's Yet Another Mobile Network (YAMNet) [16], a deep neural network using Mobilenet_v1, a depthwise-separable convolution architecture [17]. YAMNet is pre-trained on AudioSet, a human-annotated collection of more than 2 million 10-second audio clips pulled from YouTube videos [18], to predict 521 classes of audio events. The intended input of the model is audio data sampled at 16 kHz. A stabilized log Mel spectrogram is computed from the input with a frequency range of 125 Hz to 7500 Hz, after which the results are divided into 0.96 s segments with 50% overlap. These segments are fed into the Mobilenet_v1 architecture. The output of Mobilenet_v1 is then averaged-pooled, resulting in 1024 embeddings. Up to this point, the procedures we have described are those performed by YAMNet for all applications. For the general case of using YAMNet for audio classification, a final output layer would then take in the embeddings and output the class scores, from which predictions would be made. For our chosen method of transfer learning, however, the final layer must be removed, and the embeddings used as the input of our smaller, more specific rocket detection model.

2. Data and Methods

2.1. Aggregated Smartphone Timeseries of Rocket-generated Acoustics (ASTRA)

The acoustic rocket signatures used to train and test the models in this work are from Aggregated Smartphone Timeseries of Rocket-generated Acoustics (ASTRA), a publicly available dataset containing 1089 smartphone audio recordings of 243 rocket launches [19]. The data in ASTRA were recorded at 800 Hz sampling rate by Android smartphones stationed 10-70 km from launch pads at Cape Canaveral Space Force Station and the Kennedy Space Center. Along with the recordings themselves, the estimated signal start and peak times at each station, as well as the unique launch identification strings were used. The process used to calculate these estimates is detailed in Popenhagen & Garcés (2025) [20].

In preparation for machine learning, a subset of ASTRA was compiled by removing any recordings for which confidence in the alignment was reduced due to missing data or high environmental noise. The remaining 789 recordings from 233 launches were upsampled to 16 kHz, and 4.8 s windows centered on the estimated launch signal peak were selected. The selected 4.8 s windows were then divided into 0.96 s segments, resulting in 5 samples. The unique launch identification strings associated with each recording in ASTRA were used to ensure that recordings from the same launch were not split between the training, validation, and test sets. If a recording contained data taken from >120 s before the associated estimate of the arrival time of the beginning of the signal, that segment was separated and divided into up to 50 samples for additional noise data. These samples were processed similarly to the rocket signature samples, being divided into 0.96 s segments and split by associated event.

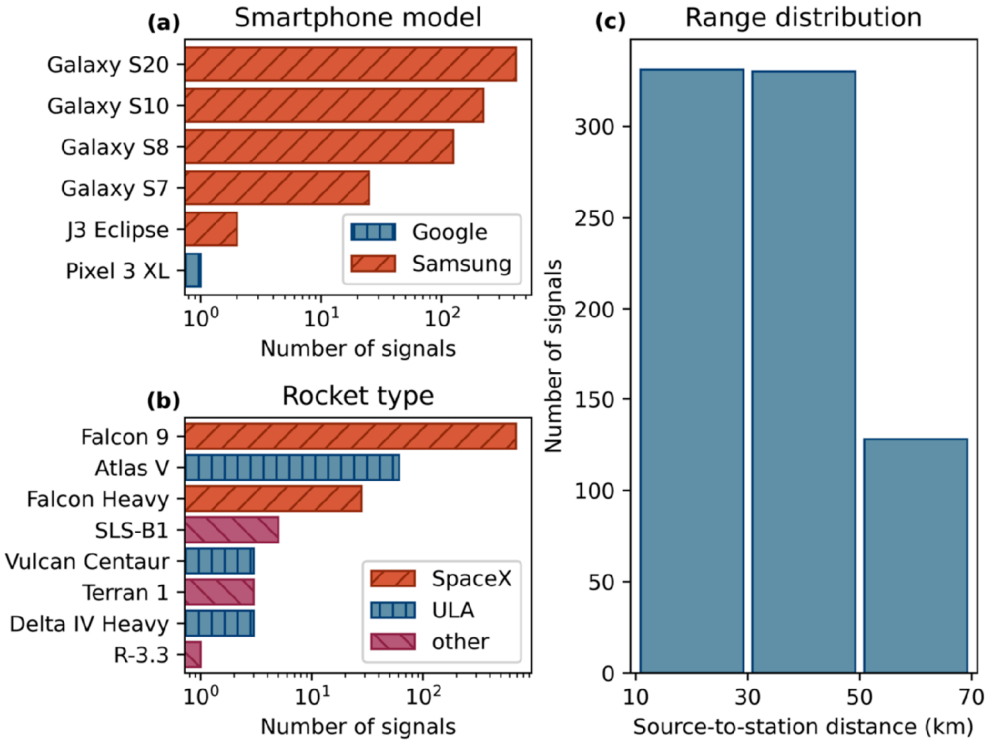

Figure 1. Bar plots of the distribution of signals in the high-confidence subset of ASTRA (a) recorded on different makes and models of smartphones, (b) originating from different types of rockets, and (c) collected in different range categories.

2.2. Smartphone High-explosive Audio Recordings Dataset (SHAReD)

The Smartphone High-Explosive Audio Recordings Dataset (SHAReD) [21] is an open-access dataset containing 326 multi-modal smartphone data collected during 70 surface high explosive events. In this work, only the smartphone microphone data were used. The explosion signals in SHAReD all have signal-to-noise ratios of >3, effective yields from 1 to 100 kg TNT, and ranges between 10g and 4 tons. SHAReD audio were collected at sampling rates of either 800 Hz or 8000 Hz. For an in-depth overview of SHAReD, we direct the reader to its accompanying papers [4,6]. For each station during each event, SHAReD includes one 0.96 s long 'explosion' waveform and one 0.96 s long 'silence' waveform (sampled from before the explosion signal's arrival). Like the ASTRA data, the SHAReD data were split by event so that the model was never trained and tested on data from the same event.

2.3. Environmental sound data from ESC-50

The ESC-50 dataset is a collection of 2000 environmental sound recordings originally recorded as part of the Freesound project [22]. Each of the 2000 clips is 5 s in duration, has a sampling rate of 44.1 kHz, and is labeled as belonging to 1 of 50 semantical classes. There are exactly 20 clips from each class (ex. 'dog', 'snoring', 'chainsaw', etc.), with some classes ('thunderstorm', 'fireworks', 'airplane', etc.) including low-frequency and/or infrasonic sources. ESC-50 is widely used as a benchmarking dataset for environmental sound classification models, and its inclusion has been shown to improve robustness and performance of transfer learning audio classification models [6]. 5 0.96 s long samples were taken from each ESC-50 clip. These samples were kept together when splitting the dataset into training, validation, and testing sets.

2.4. Preprossessing and splitting

All audio samples recorded at >800 Hz sampling rate were downsampled to 800 Hz to match the ASTRA audio samples. For each recording in ASTRA, the mean signal-to-noise ratio (SNR) of the 'rocket' and 'noise' samples was calculated, and those with ratios of <2.0 were removed to prevent training on misaligned or exceptionally noisy samples. All data except the 'rocket' data from ASTRA were labeled as 'noise'. For the proposed use-case of persistent monitoring, it is beneficial to prioritize a low false positive rate over other metrics, and thus the training data was left intentionally unbalanced, including approximately 10 'noise' samples for every 'rocket' sample in the training set. This imbalance was intended to increase robustness by exposing the model to a greater amount and variety of 'noise' data, as well as encouraging the model to favor false negatives over false positives.

The data from ASTRA, SHAReD, and ESC-50 were each split randomly into training, validation, and testing sets, with target percentages of 80%, 10%, and 10%, respectively. This random splitting was performed 15 times, in preparation for training and testing the model 15 separate times on each of the 15 randomly split datasets. The target distributions were not always met exactly, due to the constraint of isolating the data from each individual launch, explosion, or Freesound clip to one of the three sets.

2.5. Model design

Before building, training, and testing the transfer learning model, the full dataset was run through YAMNet alone, and the resulting predictions were analyzed. The distributions of the classes predicted by YAMNet for data from ASTRA, SHAReD, and ESC-50 are shown in Figure 4.2. Due to the data having been upsampled from 800 Hz and thus being devoid of any content above 400 Hz, the highly inaccurate predictions we observe are unsurprising, even for the ESC-50 data, which contains types of sounds similar to those YAMNet was trained to classify. What is interesting, however, is that the distributions of

predicted classes vary significantly between subsets of the dataset, indicating that YAMNet is able to see some differences between these subsets despite having no context for data collected at 800 Hz.

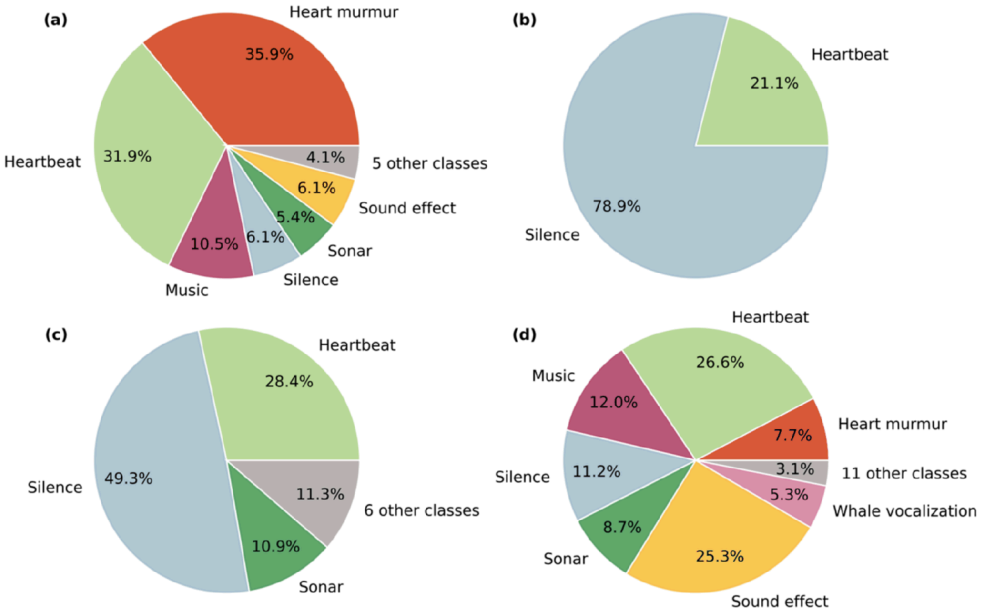

Figure 2. Bar plots of the distribution of signals in the high-confidence subset of ASTRA (a) recorded on different makes and models of smartphones, (b) originating from different types of rockets, and (c) collected in different range categories.

Transfer learning was chosen to compensate as much as possible for the limited quantity of rocket audio data. In theory, we can transfer some of the knowledge YAMNet has on the general problem of audio classification to our more specific problem of rocket detection. To construct the transfer learning model, the final output layer of YAMNet was removed and replaced with three new layers. The first of these was a fully connected layer with 32 nodes, utilizing leaky rectified linear unit (leaky ReLU) activation with an alpha value of 0.01. The second was a dropout layer added to minimize overfitting, and the final layer was an output layer with 2 nodes corresponding to the two classes ('rocket' and 'noise'). Sparse categorical cross-entropy was used for loss, and the Adamax optimizer was used. The number of nodes, activation function, alpha value, and optimizer were selected after analyzing the results of cross-validation, with lower numbers of nodes being given preference to avoid overfitting. The number of epochs was set to 300, with early stopping implemented to further reduce the likelihood of overfitting.

3. Results

The model was evaluated using the samples assigned to the training set for each run, consisting of roughly 10% of the data from each of the three open-access datasets. All testing samples were taken from events (rocket launches, explosions, or Freesound clips) that were not included in the training or validation sets. The overall performance of the model is presented, along with the performance of the best-performing split. In Figures 4.3 and 4.4, results are shown using normalized confusion matrices, which show the true positive rate in the upper left corner and the true negative rate in the lower right corner. Since the categories are unbalanced, the confusion matrices are all normalized for clarity, but the original counts were preserved and are included in parentheses on the figures.

3.1. Overall performance

The accuracy of the model ranged between 93.53% and 97.00%, with the mean over all 15 splits lying at 95.40% accuracy. The mean performance of the model is shown as a confusion matrix in Figure 4.3. In Figure 4.4, a confusion matrix of the best-performing split is shown. When trained on this split, the model performed slightly better, with no indications of overfitting. The best split accuracy was 97.00%, with a false positive rate of 0.98%.

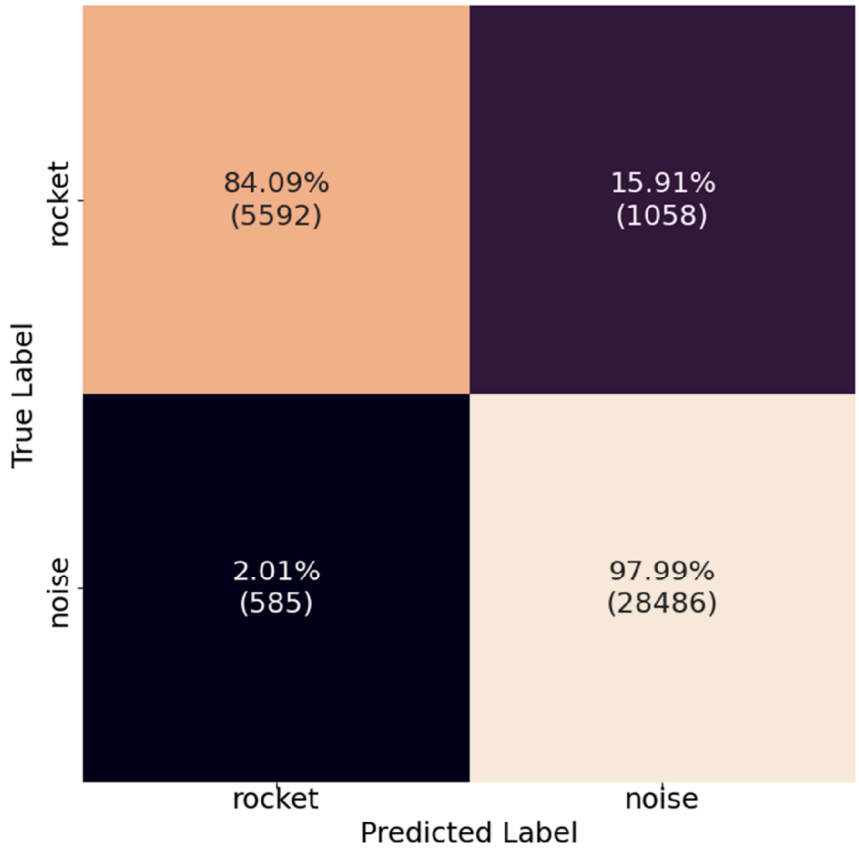

Figure 3. The confusion matrix of the model on all 15 test sets. Each quadrant shows the relevant rate as a mean percentage over all iterations, as well as the sum of the total number of samples in the category over all iterations in parentheses.

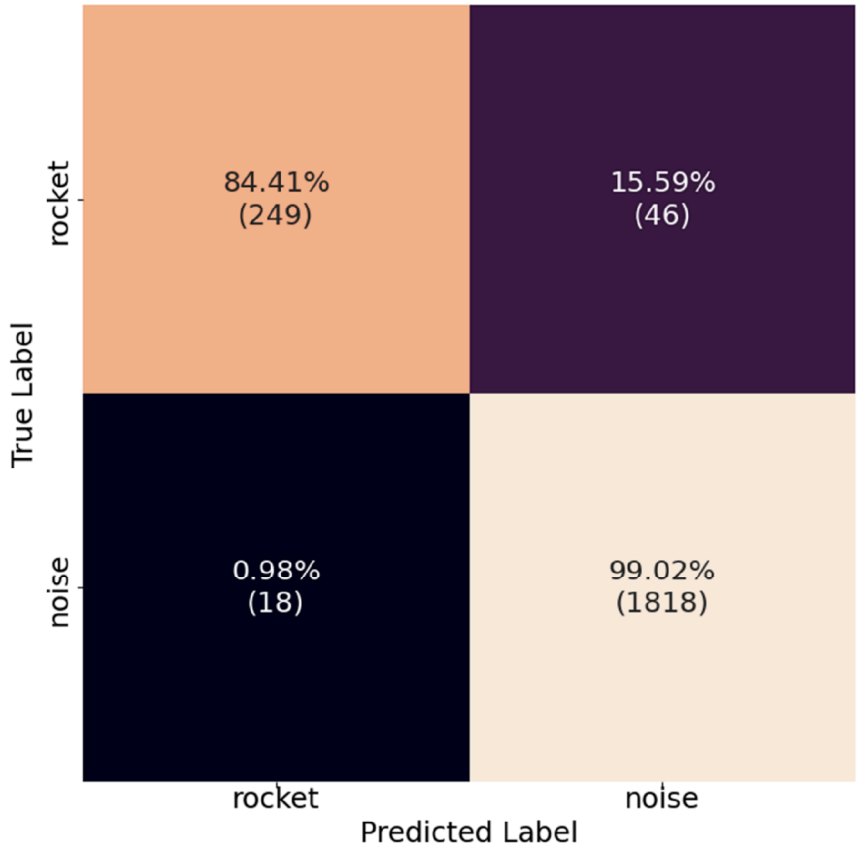

Figure 4. The confusion matrix of the best performing split. Each quadrant shows the relevant rate as a percentage, as well as the total number of samples in the category in parentheses.

3.2. Misclassification analysis

To better understand the model's behavior, it is important to remember that 5 positive samples were taken from each rocket recording, and that there are multiple recordings for most launches. In Table 4.1, the ASTRA events with false negatives in the best split model are listed, along with the other classifications from the event, which show that for every event with false negatives, there are always more true positives. For every recording in the best split model's test set, there is at least one 'rocket' classification within 10 seconds of the estimated peak arrival time included in ASTRA. From an application perspective, then, there are no false negatives, only delayed true positives, and we can focus on reducing the false positive rate.

Table 1. All events with 'rocket' samples misclassified as 'noise' by the best split model.

Event	No. false negatives	No. positive samples	No. true positives
ASTRA_218	7	30	21
ASTRA_219	3	15	12
ASTRA_220	1	25	24
ASTRA_221	1	10	9
ASTRA_222	10	25	15
ASTRA_224	5	20	15
ASTRA_225	4	25	21
ASTRA_226	2	20	18
ASTRA_227	3	10	7

ASTRA_229	5	15	10
ASTRA_230	1	20	19
ASTRA_232	4	20	16

The simplest method of reducing the false positive rate is to increase the threshold for positive 'rocket' classifications slightly, from 0.5 to 0.6. With the increased threshold, the best split model's false positive rate decreases from 0.98% to 0.76%. This improvement in the false positive rate is accompanied by an increase of the false negative rate from 15.59% to 19.66%. However, there is still at least one 'rocket' classification within 10 seconds of the estimated peak arrival time of each recording, thus the practical effect of the increased false negative rate is minimal. The best split model's remaining false positives after increasing the threshold are shown in Table 4.2. Upon examination, it is clear that false positives tend to appear chronologically isolated or in small clusters of 5 or less samples.

Table 2. All events with 'noise' samples misclassified as 'rocket' by the best split model.

Event	No. false positives	No. negative samples	No. true negatives
ASTRA_225	5	50	45
ESC50_1660	3	5	2
ESC50_1667	1	5	4
ESC50_1675	1	5	4
ESC50_1683	1	5	4
ESC50_1731	2	5	3
ESC50_1761	1	5	4
ESC50_1779	2	5	3
ESC50_1808	2	5	3

To visualize this behavior, we can plot the 'rocket' classifications as vertical green lines with a transparency inversely proportional to the model's confidence in its classification. In Figure 4.5, all recordings from one of the launches in the testing set are plotted this way, showing a number of false positives before the arrival of the signal at two of the three stations. Since the rocket launch signal persists over at least 100 s, we can eliminate chronologically isolated positives by setting all scores less than the 0.6 threshold to 0 and applying a 4.8 s long rolling median filter. In Figure 4.6, the remaining 'rocket' classifications after applying the median filter. After applying this method to the scores of pre-launch segments of the ASTRA recordings included in the test set, the best split model's false positive rate on pre-launch ASTRA data decreased from 0.76% to 0.11%. This improvement in the false positive rate has minor tradeoffs in two areas. First, the detection time is delayed from the first 'rocket' classification over 0.5 by a mean of 4.92 s. Second, despite the increase in the false negative rate appearing negligible on the sample level, 1 of the 100 recordings in the test set now has no 'rocket' classifications within 10 seconds of its estimated peak arrival time.

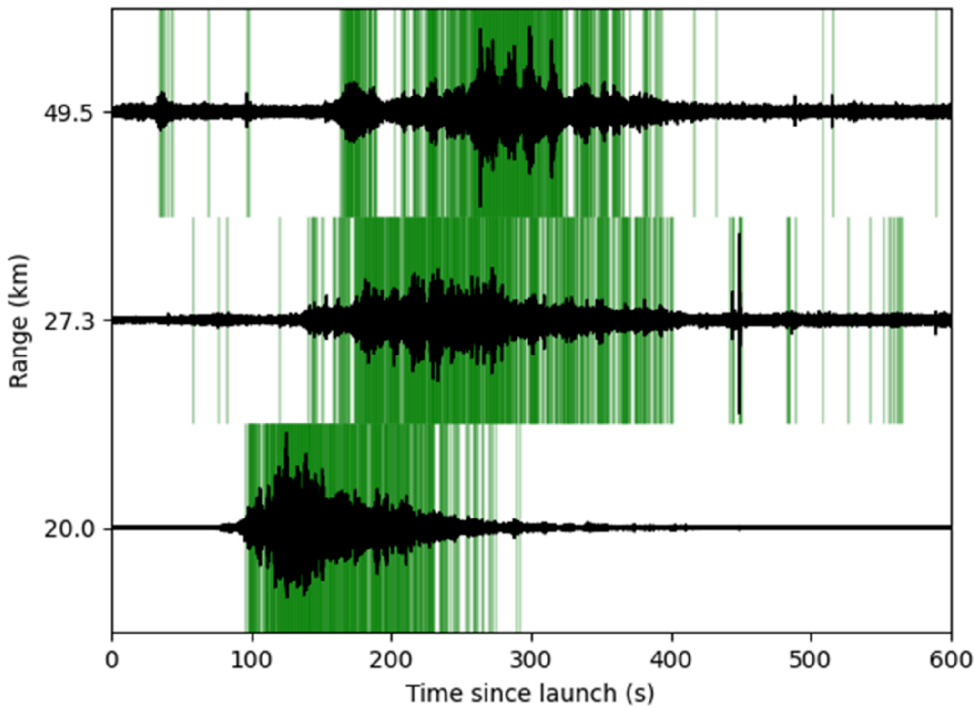

Figure 5. Plot showing the best split model's unfiltered performance on all signals from the SpaceX Falcon 9 launch, Transporter-4. Green lines indicate 'rocket' classification scores >0.6.

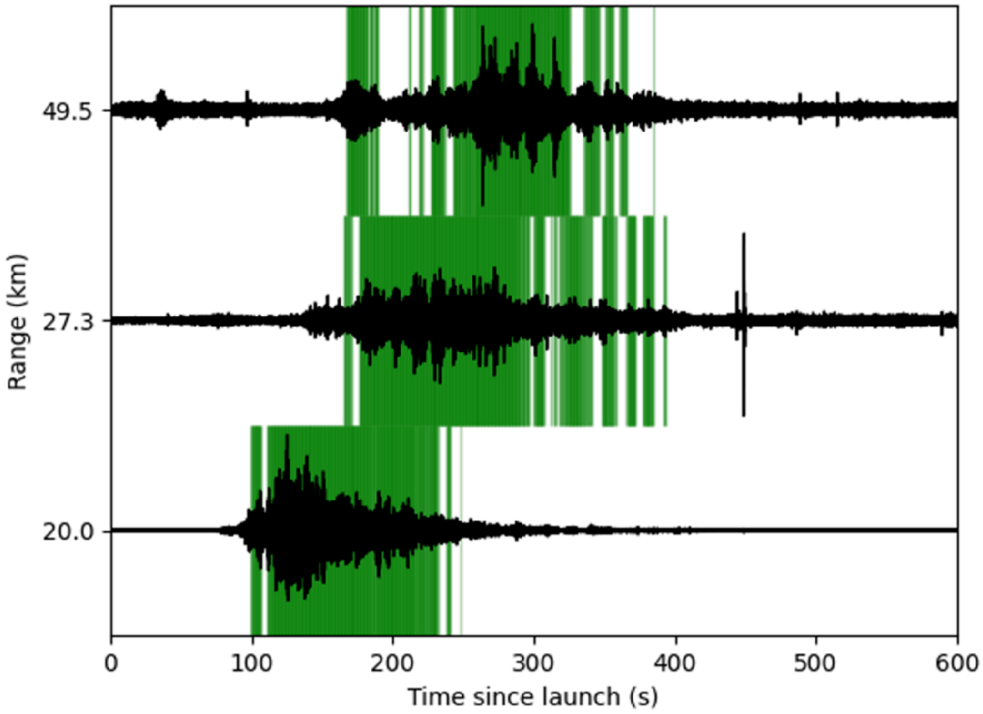

Figure 6. Plot showing the best split model's median-filtered performance on all signals from the SpaceX Falcon 9 launch, Transporter-4. Green lines indicate 'rocket' classification scores >0.6 after median filtering is applied.

When applied to the ESC-50 data, this method also decreases the false positive rate significantly. Figure 4.7 shows the false positive rates before and after thresholding and

median filtering for each ESC-50 class containing samples the best split model misclassified, as well as the mean false positive rates over all ESC-50 classes. As can be seen in Figure 4.7, the overall false positive rate on ESC-50 data dropped to 0.05% after applying thresholding and median filtering. As the recordings from SHAReD were each only 0.96 s long, we were unable to determine the effect of median filtering on those data. However, the model performed exceptionally well on SHAReD data even before adjusting the threshold, with no misclassifications by the best split model and false positive rates <1% for all 15 models.

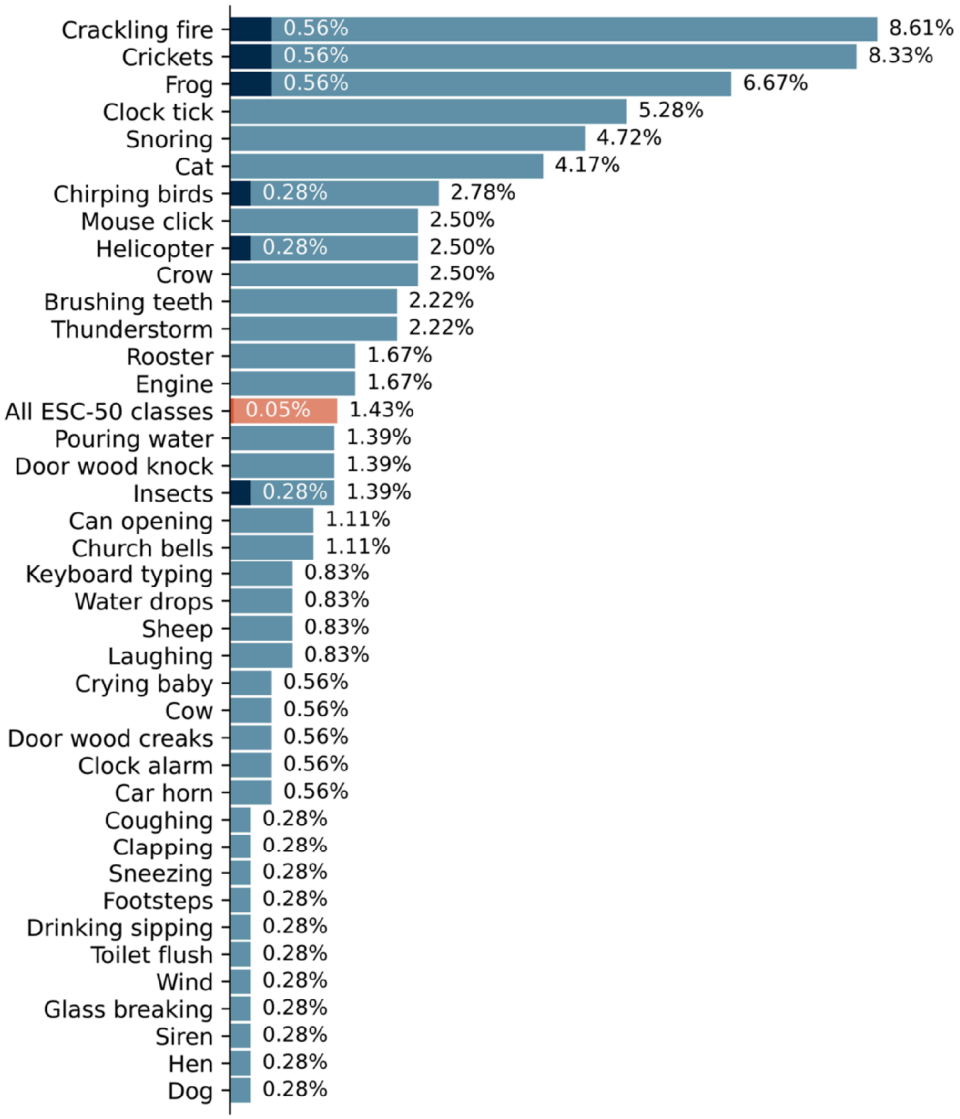

Figure 7. A horizontal bar plot representing the best split model's false positive classifications of ESC-50 data, plotted according to class labels included with ESC-50. Light colored bars represent the false positive rates before thresholding and median filtering, and dark colored bars represent the same rates after thresholding and median filtering are applied. Results for individual classes are shown in blue and labeled with black text, while the overall results are shown in orange and labeled with white text.

4. Discussion

After training and testing on 15 random train-test splits, we found that the transfer learning model's behavior is exceptionally stable, with only slight variations in the results between different splits. Analyzing the behavior of the best split model in detail, it was discovered that despite the nominal false negative rate of 15.59%, the model made at least one 'rocket' classification within 10 seconds of the estimated arrival times of 99/100 recordings in the training set. In addition, the already low false positive rate of 0.98% can be reduced even further by increasing the positive classification threshold from 0.5 to 0.6 and applying a 4.8 s long median filter. The application of these two measures resulted in the best split model's false positive rates on ESC-50 data decreasing from 1.43% to 0.05%, prelaunch 'noise' data from ASTRA decreasing from 0.76% to 0.11%, and the overall false positive rate decreasing from 0.98% to 0.072%. The increased threshold and addition of the median filter combined were found to increase the time between launch and detection by 4.92 s on average, but this delay is relatively short compared to the estimated propagation times (~60-200 s) and its negative effects would likely be outweighed by the benefits of the decreased false positive rate in most use cases. Applying the median filter also reduced the true positive rate slightly, resulting in the model failing to detect the launch within 10 seconds of the estimated peak arrival time of 1 of the 100 recordings in the dataset. It is worth noting, however, that the recording in question was 1 of 4 recordings of the same launch, and the model was able to detect the launch successfully in the other 3 recordings.

Although the initial success of this transfer learning model is encouraging, the dataset is too small to truly determine its usefulness for persistent monitoring. As more and more rockets continue to be launched every year, the amount of available data will likely increase rapidly in the next few years. Future work should include long-term deployment and testing of the model to fully evaluate its potential for monitoring. If further decreases in the false positive rate are required in the future, there are a number of different avenues to pursue, including training against additional sources of infrasound such as earthquakes or volcanic activity, incorporating ensemble learning [6], and combining input from multiple smartphones, which may also prove useful for trajectory modeling in the future.

Author Contributions: Conceptualization, S.K.P. and M.A.G.; methodology, S.K.P.; software, S.K.P.; validation, S.K.P.; formal analysis, S.K.P.; investigation, S.K.P.; resources, M.A.G.; data curation, S.K.P.; writing—original draft preparation, S.K.P.; writing—review and editing, S.K.P. and M.A.G.; visualization, S.K.P.; supervision, M.A.G.; project administration, M.A.G.; funding acquisition, M.A.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Department of Energy National Nuclear Security Administration under Awards Nos. DE-NA000390 (MTV) and DE-NA0003921 (ETI).

Data Availability Statement: ASTRA is available as a pandas DataFrame [23] and can be found in the Harvard Dataverse open access repository under the following Digital Object Identifier doi: 10.7910/DVN/ZKIS2K. The ERA5 temperature and wind data used in this study are available through the Copernicus Climate Change Service [24,25].

Acknowledgments: The authors are grateful for the support of the U.S. Department of Energy, National Nuclear Security Administration, Office of Defense Nuclear Nonproliferation, Research and Development. They would also like to thank Garrett Apuzen-Ito, Helen Janiszewski, Sloan Coats, and Brian Powell for their comments and suggestions on this manuscript and work, as well as all those who supplied feedback on this project at MTV and ETI conferences. This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

IMS	International Monitoring System
ASTRA	Aggregated Smartphone Timeseries of Rocket-generated Acoustics
SHAReD	Smartphone High-explosive Audio Recordings Dataset
YAMNet	Yet Another Mobile Network
ReLU	Reticulated Linear Unit
MTV	Consortium for Monitoring, Technology, and Verification
ETI	Consortium for Enabling Technologies and Innovation

References

1. Schwardt, M.; Pilger, C.; Gaebler, P.; Hupe, P.; Ceranna, L. Natural and Anthropogenic Sources of Seismic, Hydroacoustic, and Infrasonic Waves: Waveforms and Spectral Characteristics (and Their Applicability for Sensor Calibration). *Surv Geophys* 2022, *43*, 1265–1361.
2. Hupe, P.; Ceranna, L.; Le Pichon, A.; Matoza, R.S.; Mialle, P. International Monitoring System Infrasound Data Products for Atmospheric Studies and Civilian Applications. *Earth Syst Sci Data* **2022**, *14*, 4201–4230, doi:10.5194/essd-14-4201-2022.
3. Pilger, C.; Hupe, P.; Gaebler, P.; Ceranna, L. 1001 Rocket Launches for Space Missions and Their Infrasonic Signature. *Geophys Res Lett* **2021**, *48*, doi:10.1029/2020GL092262.
4. Takazawa, S.K.; Popenhagen, S.K.; Ocampo Giraldo, L.A.; Cárdenas, E.S.; Hix, J.D.; Thompson, S.J.; Chichester, D.L.; Garcés, M.A. A Comparison of Smartphone and Infrasound Microphone Data from a Fuel Air Explosive and a High Explosive. *J Acoust Soc Am* **2024**, *156*, 1509–1523, doi:10.1121/10.0028379.
5. Popenhagen, S.K.; Bowman, D.C.; Zeiler, C.; Garcés, M.A. Acoustic Waves From a Distant Explosion Recorded on a Continuously Ascending Balloon in the Middle Stratosphere. *Geophys Res Lett* **2023**, *50*, doi:10.1029/2023GL104031.
6. Takazawa, S.K.; Popenhagen, S.K.; Ocampo Giraldo, L.A.; Hix, J.D.; Thompson, S.J.; Chichester, D.L.; Zeiler, C.P.; Garcés, M.A. Explosion Detection Using Smartphones: Ensemble Learning with the Smartphone High-Explosive Audio Recordings Dataset and the ESC-50 Dataset. *Sensors* **2024**, *24*, doi:10.3390/s24206688.
7. Thandu, S.C.; Chellappan, S.; Yin, Z. Ranging Explosion Events Using Smartphones. In Proceedings of the 2015 IEEE 11th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob); IEEE, October 2015; pp. 492–499.
8. Thandu, S.C.; Bharti, P.; Chellappan, S.; Yin, Z. Leveraging Multi-Modal Smartphone Sensors for Ranging and Estimating the Intensity of Explosion Events. *Pervasive Mob Comput* **2017**, *40*, 185–204, doi:10.1016/j.pmcj.2017.06.012.
9. Bozinovski, S. Reminder of the First Paper on Transfer Learning in Neural Networks, 1976. *Informatica (Slovenia)* **2020**, *44*, 291–302, doi:10.31449/INF.V44I3.2828.
10. Bozinovski, S.; Fulgosi, A. The influence of pattern similarity and transfer of learning upon training of a base perceptron B2. In Proceedings of the Symposium Informatica 3-121-5; Bled, 1976.
11. Brusa, E.; Delprete, C.; Di Maggio, L.G. Deep Transfer Learning for Machine Diagnosis: From Sound and Music Recognition to Bearing Fault Detection. *Applied Sciences* **2021**, *11*, 11663, doi:10.3390/app112411663.
12. Tsalera, E.; Papadakis, A.; Samarakou, M. Comparison of Pre-Trained CNNs for Audio Classification Using Transfer Learning. *Journal of Sensor and Actuator Networks* **2021**, *10*, 72, doi:10.3390/jsan10040072.
13. Ashurov, A.; Zhou, Y.; Shi, L.; Zhao, Y.; Liu, H. Environmental Sound Classification Based on Transfer-Learning Techniques with Multiple Optimizers. *Electronics (Basel)* **2022**, *11*, 2279, doi:10.3390/electronics11152279.
14. Hyun, S.H. Sound-Event Detection of Water-Usage Activities Using Transfer Learning. *Sensors* **2023**, *24*, 22, doi:10.3390/s24010022.
15. Pan, S.J.; Yang, Q. A Survey on Transfer Learning. *IEEE Trans Knowl Data Eng* 2010, *22*, 1345–1359.
16. Plakal, M.; Ellis, D. YAMNet 2029.
17. Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications 2017.
18. Gemmeke, J.F.; Ellis, D.P.W.; Freedman, D.; Jansen, A.; Lawrence, W.; Moore, R.C.; Plakal, M.; Ritter, M. Audio Set: An Ontology and Human-Labeled Dataset for Audio Events. In Proceedings of the 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP); IEEE, March 2017; pp. 776–780.

19. Popenhagen, S.K. Aggregated Smartphone Timeseries of Rocket-Generated Acoustics (ASTRA) Available online: https://doi.org/10.7910/DVN/ZKIS2K (accessed on 21 November 2024).
20. Popenhagen, S.K.; Garcés, M.A. Acoustic Rocket Signatures Collected by Smartphones. *Signals* **2025**, *6*, 5, doi:10.3390/signals6010005.
21. Takazawa, S.K. Smartphone High-Explosive Audio Recordings Dataset (SHAReD) 2024.
22. Piczak, K.J. ESC: Dataset for Environmental Sound Classification. In Proceedings of the the 23rd ACM International Conference on Multimedia; Association for Computing Machinery: Brisbane, Australia, 2015; pp. 1015–1018.
23. The pandas development team Pandas-Dev/Pandas: Pandas Available online: https://doi.org/10.5281/zenodo.3509134 (accessed on 21 November 2024).
24. Hersbach, H.; Bell, B.; Berrisford, P.; Biavati, G.; Horányi, A.; Muñoz Sabater, J.; Nicolas, J.; Peubey, C.; Radu, R.; Rozum, I.; et al. ERA5 Hourly Data on Single Levels from 1940 to Present.
25. Copernicus Climate Change Service (2023): ERA5 Hourly Data on Single Levels from 1940 to Present Available online: https://cds.climate.copernicus.eu/datasets/reanalysis-era5-single-levels (accessed on 26 November 2024).