ShotSpotter: Real-Time Gunshot Detection in Urban Noise

Lin Xu, Vincent Lu, Benjamin Shan, Andreas G. Andreou

Whiting School of Engineering | The Johns Hopkins University | Baltimore, MD

Introduction

- Gun violence remains a critical public safety challenge worldwide, with delayed emergency responses often leading to injuries and deaths. The motivation behind developing a machine learning (ML)-based gunshot detection system lies in its potential to save lives by enabling faster, more accurate identification of firearm incidents.
- Traditional detection methods, such as audio sensors or human reporting, suffer from high false alarms, latency, or limited coverage. Leveraging ML algorithms to analyse acoustic signatures offers a scalable, real-time solution to distinguish gunshots from similar noises (e.g., fireworks, traffic) and pinpoint their location.
- This project aims to create a model that enhances public safety infrastructure, empowers law enforcement with actionable data, and ultimately contributes to safer communities. By integrating technology with societal needs, this work underscores the transformative role of ML in addressing urgent humanitarian challenges.

Methods

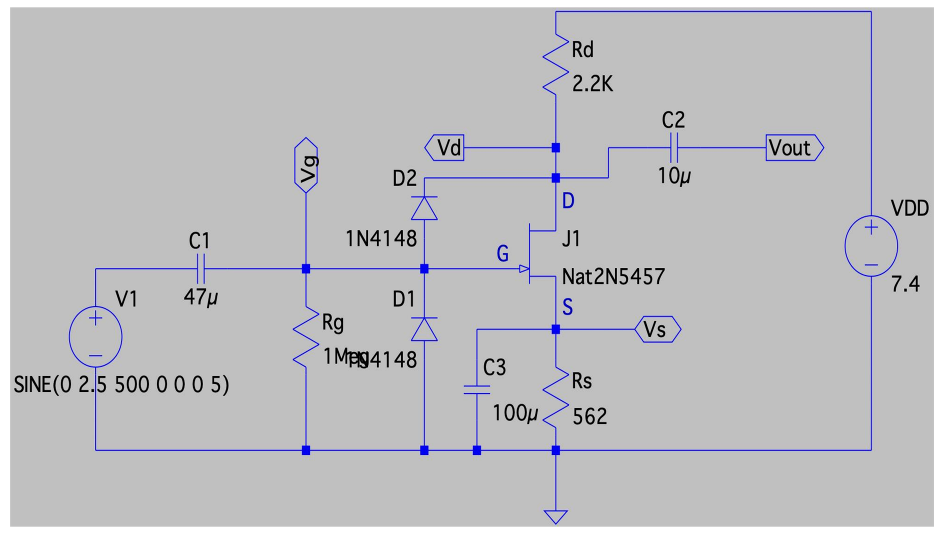

Figure 1 – JFET-based preamplifier circuit in LTSpice.

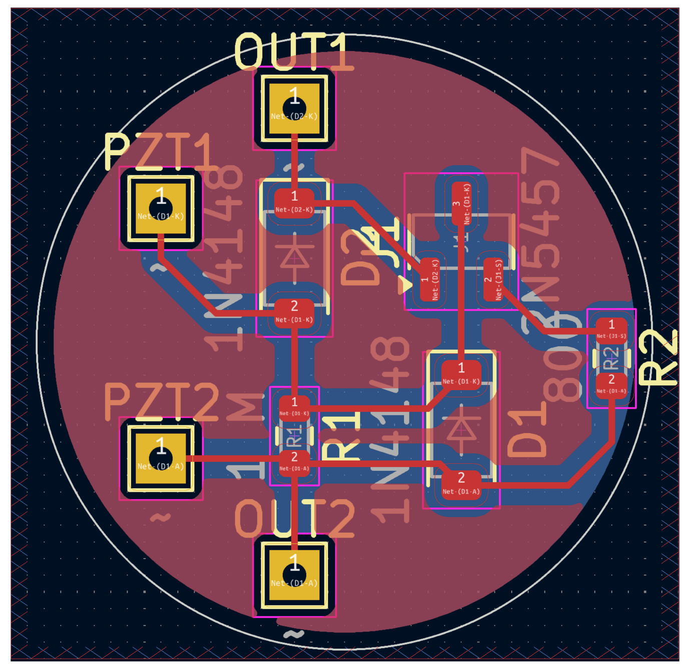

Figure 2 – PCB layout in KiCad.

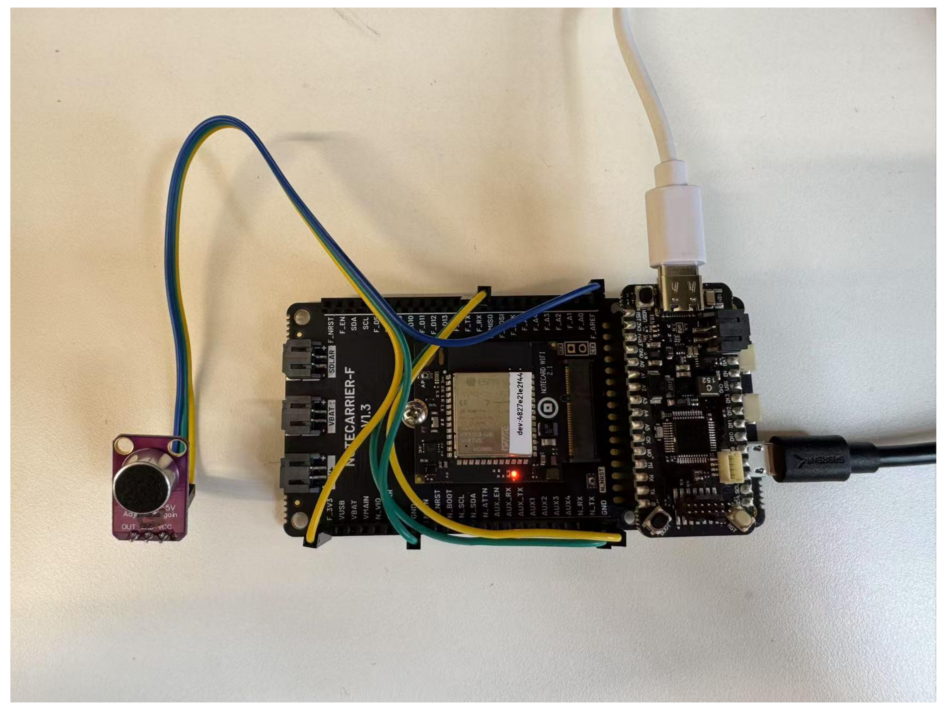

Figure 3 – Device layout: Blues Cygnet + Blues Notecard connected to a condenser microphone and PC.

- This project implements an integrated hardware-software architecture to enable robust acoustic sensing. The hardware subsystem was initially designed around a custom-built studio condenser microphone, which utilizes a piezoelectric transducer for high-fidelity signal reception and a JFET-based preamplifier circuit (simulated in LTSpice with a 5× voltage gain) to amplify transient acoustic events like gunshots. A PCB layout was developed using KiCad, though unexpected delays in shipping necessitated the temporary use of a commercial microphone to maintain project timelines.

- The deployed setup pairs the microphone with a Blues Cygnet microcontroller board, which digitizes and streams audio data to a PC via USB. The wireless Blues Notecard operates by connecting to local Wi-Fi, enabling real-time environmental sound capture and automatic cloud uploads for centralized processing. Acoustic features, such as Mel-Frequency Energy (MFE) coefficients, are extracted from the raw signals to train machine learning models that are capable of distinguishing gunshots from ambient noise. This hybrid approach balances hardware resilience with algorithmic precision, ensuring scalability for real-world public safety applications.

DSP & ML

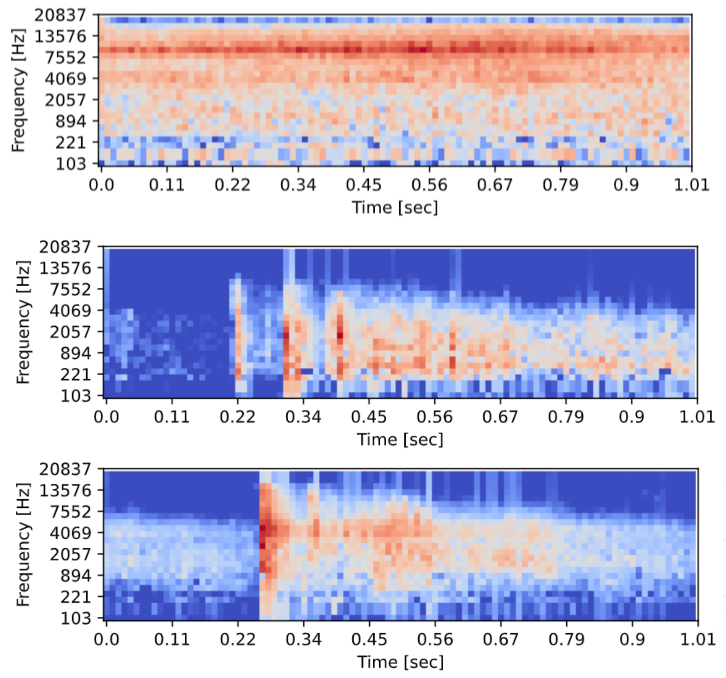

Figure 4 – Mel energy spectrograms of ambient noise, fireworks, and gunshot, in order.

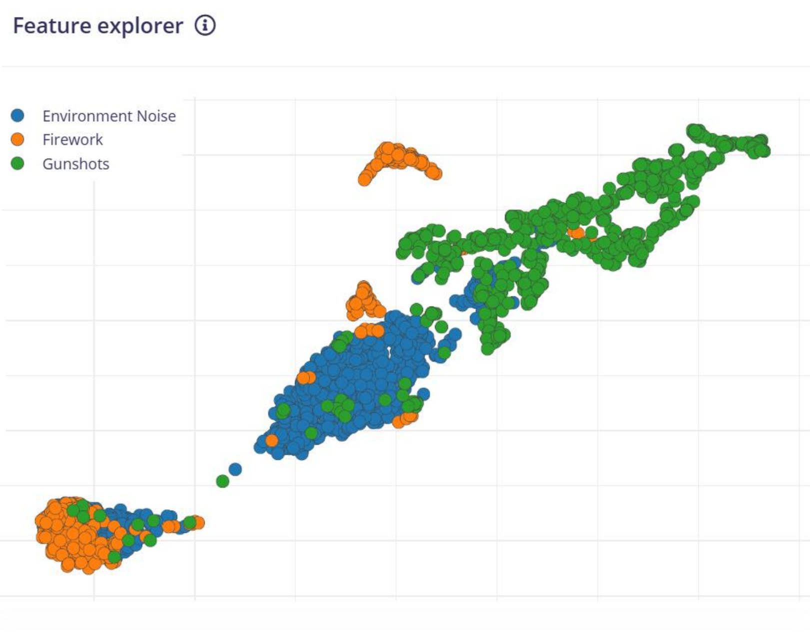

Figure 5 – Training data distribution.

- **Digital Signal Processing (DSP):** By applying Mel Energy spectrogram extraction, the time-domain audio signals converts into a compact time-frequency representation that highlights the energy distribution across different frequency bands. As shown in Figure 4, ambient noise, fireworks, and gunshots exhibit distinct Mel spectral patterns. Gunshots typically produce sharp bursts of energy across both low and mid-frequency bands, whereas fireworks and environmental sounds have more dispersed or fluctuating energy.

- **Machine Learning (ML):** The extracted Mel Energy features were used to train a machine learning model to distinguish between different noise. In Figure 5, gunshots (green) form a well-separated cluster from fireworks (orange) and environmental noise (blue), indicating that the model successfully learned discriminative features. This separation validates that our DSP pipeline preserved important class-specific information and that the model can generalize to new, unseen audio events.

Discussion

- The resulting performance underscores the viability of machine learning in reducing false alarms and enabling rapid, precise threat identification—a critical advancement over legacy acoustic detection systems.
- Limitations include the potential for unseen environment noise types (e.g., construction noise) to cause misclassification.
- Some future works include:
  - Expand the dataset to include more diverse and complicated urban noises, such as sirens, construction, and so on.
  - Implement on-device model updating so the system can learn and adapt over time.
  - Develop a multi-microphone network to allow gunshot localization (triangulation of sound source).

Results

- The trained machine learning model achieved a 91.3% classification accuracy in distinguishing gunshots from ambient urban noises (e.g., fireworks, traffic, crowd).
- The confusion matrix shows high true positive and true negative rates, indicating some but minimal misclassifications.
- The training performance plot demonstrates rapid convergence with only a little signs of overfitting (training and validation accuracy/loss are relatively matched).
- The final model was successfully deployed onto the Blues Cygnet microcontroller, maintaining high accuracy in real-world testing conditions.
- Using Blues Notecard, each detection of gunshot was successfully transmitted to Notehub cloud in real-time for easier and more direct access.

ACCURACY 91.3%

LOSS 0.26

Confusion matrix (validation set)

	ENVIRONMENT NOISE	FIREWORK	GUNSHOTS
ENVIRONMENT NOISE	99.3%	0.7%	0%
FIREWORK	12.5%	81.8%	5.7%
GUNSHOTS	11.3%	2.2%	86.5%
F1 SCORE	0.93	0.87	0.92

Figure 6 – Training performance and confusion matrix.

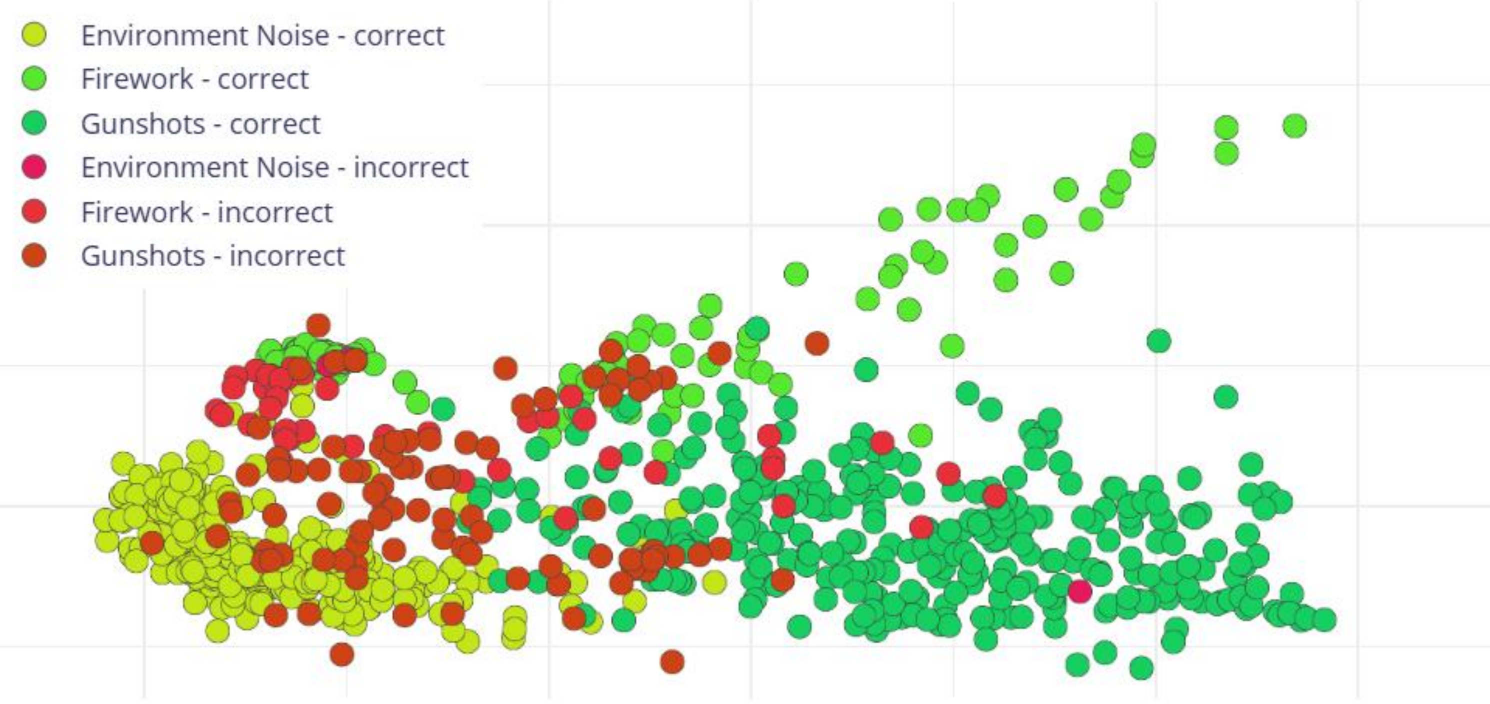

Figure 7 – Training performance feature.

Transport	Uploaded	Best Location	Best UID	File	Body
	Mon 04:31:00 PM	Baltimore Maryland	dev:4827e21e2f44	gunshot.qo	{"GUNSHOT DETECTED! "
	Mon 04:30:30 PM	Baltimore Maryland	dev:4827e21e2f44	gunshot.qo	{"GUNSHOT DETECTED! "

Figure 8 – Notehub syncing result.