WatchDog: Real-Time Gunshot Detection & Response

Satyarth Arora, Scott Pyle, Matthew Sena, Zafiro Paredes, Dr. Nigel John
Department of Computer and Electrical Engineering, University of Miami

Background

Mass shootings in the United States represent a tragic and growing problem, significantly impacting public safety. Research by the National Institute of Justice indicates a sharp increase in such events, with 170 mass shootings recorded between 2013 and 2022, a stark contrast to the 12 incidents documented between 1966 and 1975 [1]. Furthermore, data from another NIJ study reveals that 20% of all shootings over the last half-century took place between 2014 and 2019, with the locations predominantly being workplaces (30.8%) and public venues such as retail spaces, bars, or restaurants (30.3%) [2].

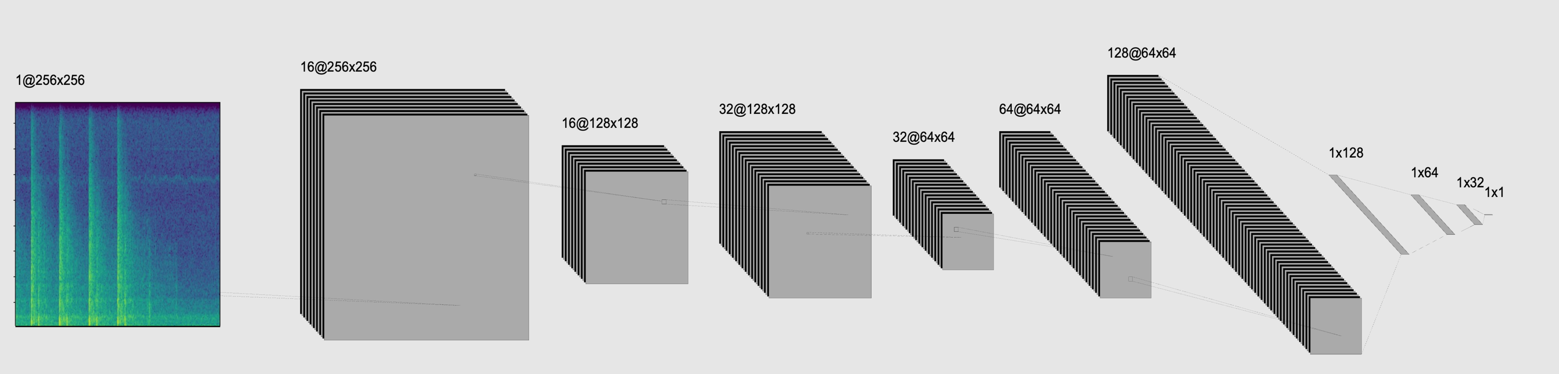
Existing gunshot detection systems, though vital, fall short in several areas; they often lack the ability to provide consistent, real-time data, are prohibitively expensive, and are not scalable to the needs of diverse environments ranging from small institutions to entire cities. By isolating and accurately identifying gunshot sources within institution-controlled environments, this system seeks to enhance the capacity of first responders to act swiftly and efficiently.

Methods

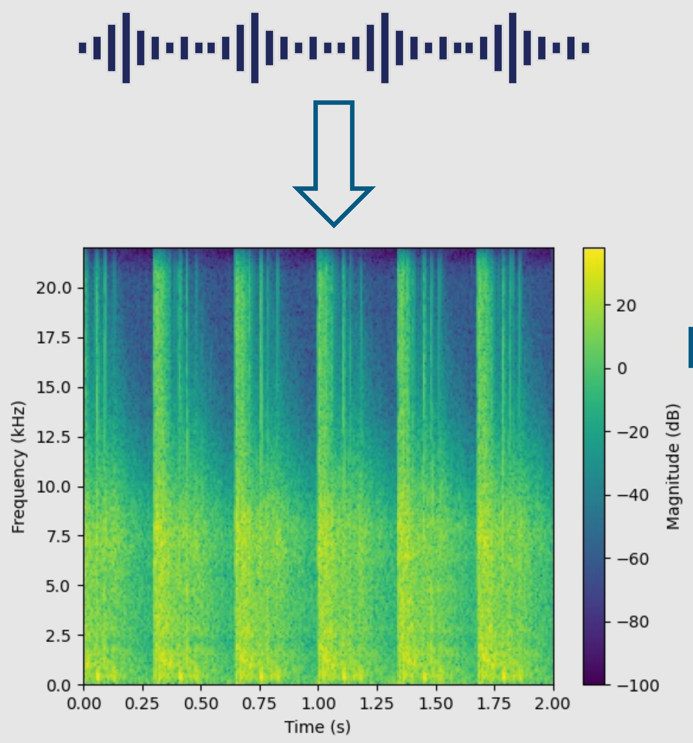
Data Collection and Preprocessing: Using publicly-available data and media, we aggregated a dataset of 10,000 unique gunshot samples from more than 30 types of firearms. Our negative samples were largely environmental sounds sourced from UrbanSound8k and ESC-50, supplemented by our own recordings around the university campus. Each audio file was standardized to a 2-second clip at a 44100 Hz sample rate and transformed into a 256x256 spectrogram. Common augmentation techniques were used for robustness.

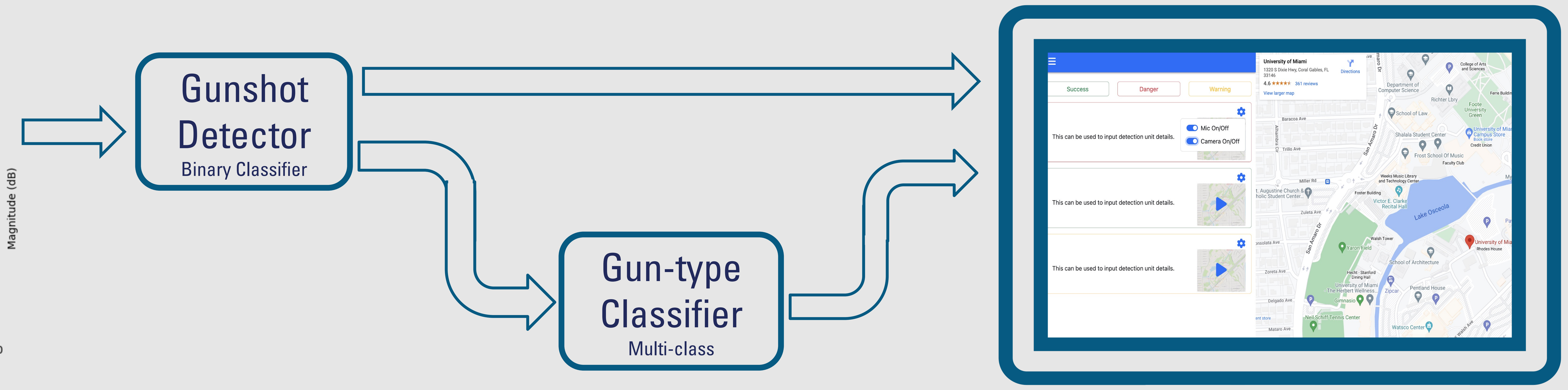
Network Architectures: We evaluated four CNN models: a baseline 3-layer CNN, ResNet18, ResNet34, and our custom model (see below) featuring dilated convolutions to optimize for real-time performance with minimal computational resources.

Training and Evaluation: All models were trained and cross-validated on the same data splits, with identical hyper-parameters. Our methods of evaluation include testing performance, classification metrics, and computational metrics.



Data Pipeline





Containerized deployment pipeline: audio data is streamed in real-time, converted to a spectrogram and processed by the binary classifier. If a gunshot is detected, information is sent to the user and the audio sample is forwarded to the multi-class model to identify the firearm. Containerization allows the system to continue operation throughout this response with no downtime and no gaps in the audio stream, so every threat is known.

Results

Figure 1 to the right compares the training and validation loss curves of the four networks throughout identical training sessions. The custom network appears to outperform the others, showing far more stability than the ResNets and almost no divergence. The table below compares testing performances of the networks after training. Again, the custom network outperforms the others in terms of accuracy, loss, and F1-score despite being 100x smaller in terms of parameters with much faster inference. The custom network is best-suited for the given task.

Network	# Params	Test Accuracy	Test Loss (BCE)	F1-Score	Inference Rate*
Simple CNN	203K	98.901%	3.39e-2	0.9878	580.56 FPS
ResNet18	11M	99.773%	4.66e-3	0.9974	44.94 FPS
ResNet34	21M	99.774%	9.90e-3	0.9975	28.26 FPS
Our Network	**280K**	**99.899%**	**3.02e-3**	**0.9988**	97.42 FPS

*Average from 10,000 forward passes with gradient calculation disabled. Conducted on Apple M2 Ultra – 24 core CPU

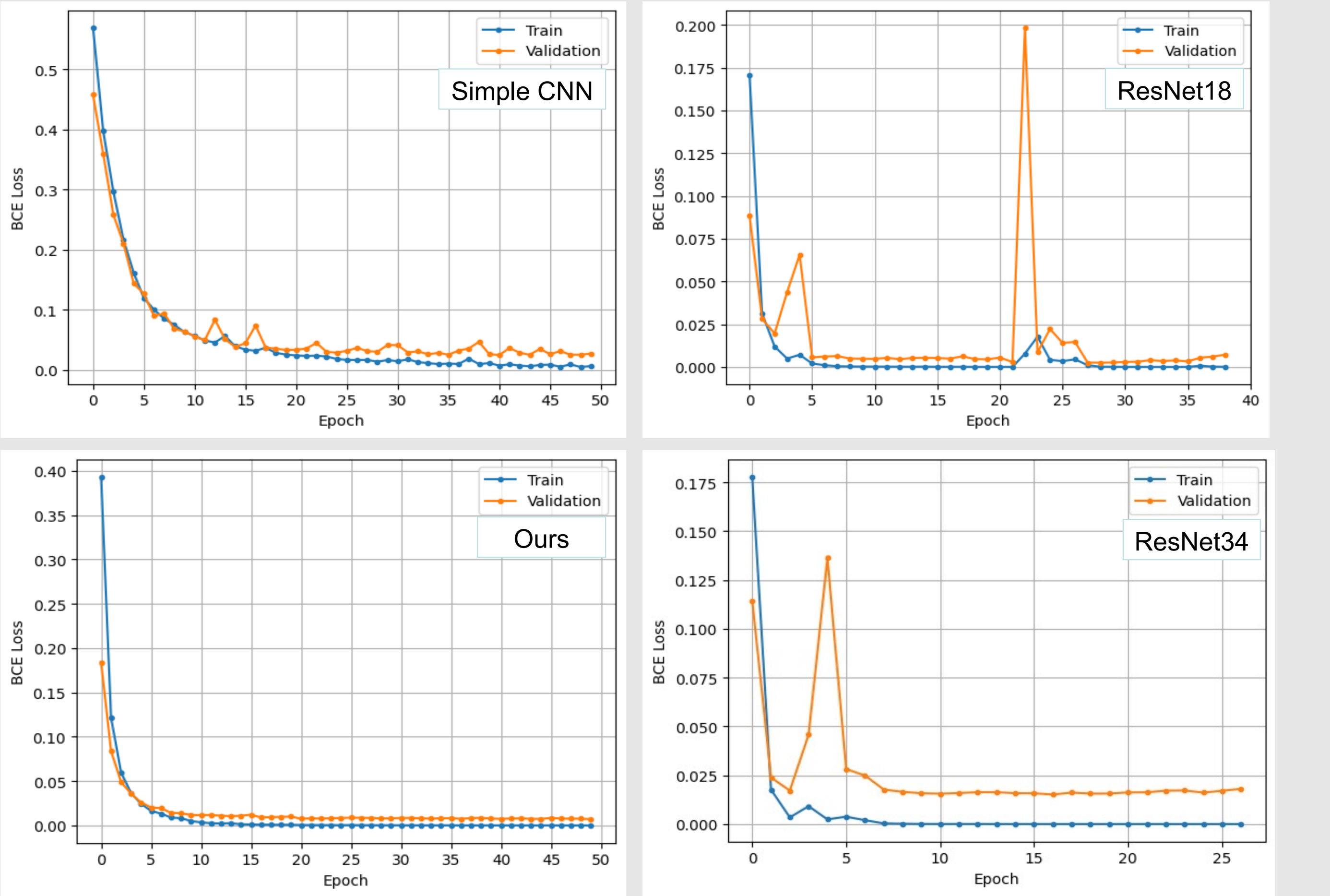
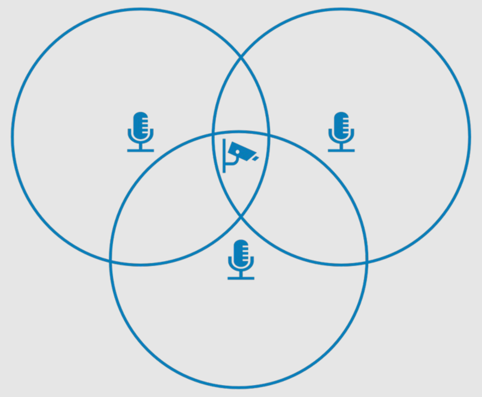


Figure 1: Training and validation loss while training for 50 epochs (or until loss divergence)

Future Direction

- Optimize the multi-classification system to accurately predict weapon source type and model.
- This integration with current first responder protocols to mitigate the impact of incidents.
- Complete hardware implementation to test modular system
- Expand modular system into a public space for further analysis and testing



Acknowledgments

We would like to extend our gratitude to our advisor Dr. Nigel John in advice and encouragement during the ideation, implementation, and deployment phases of this project.

References

[1]"Mass Shooting Factsheet | Rockefeller Institute of Government," Rockefeller Institute of Government, Jun. 07, 2018. https://rockinst.org/gun-violence/mass-shooting-factsheet/ (accessed Apr. 16, 2024).
[2]"Public Mass Shootings: Database Amasses Details of a Half Century of U.S. Mass Shootings with Firearms, Generating Psychosocial Histories," National Institute of Justice, 2024. https://nij.ojp.gov/topics/articles/public-mass-shootings-database-amasses-details-half-century-us-mass-shootings (accessed Apr. 16, 2024).
[3]A. Morehead, L. Ogden, G. Magee, R. Hosler, B. White, and G. Mohler, "Low Cost Gunshot Detection using Deep Learning on the Raspberry Pi," IUScholarWorks (Indiana University), Dec. 2019, doi: https://doi.org/10.1109/bigdata47090.2019.9006456.
[4]Moez Krichen, "Convolutional Neural Networks: A Survey," Computers, vol. 12, no. 8, pp. 151–151, Jul. 2023, doi: https://doi.org/10.3390/computers12080151.