1.0 PROJECT DEFINITION

Guardian is an automatic notification system that uses smart sensors to detect gunshot sounds at street level and updates emergency services. By reducing response time, the system takes the city one step forward in tackling gun violence.

1.1 BACKGROUND INFORMATION

Gun violence is a prominent issue for the city of Toronto. In 2018, shootings in the city have more than doubled since 2014 [1]. In addition, incidents involving guns made up more than half of all homicides this year [1].

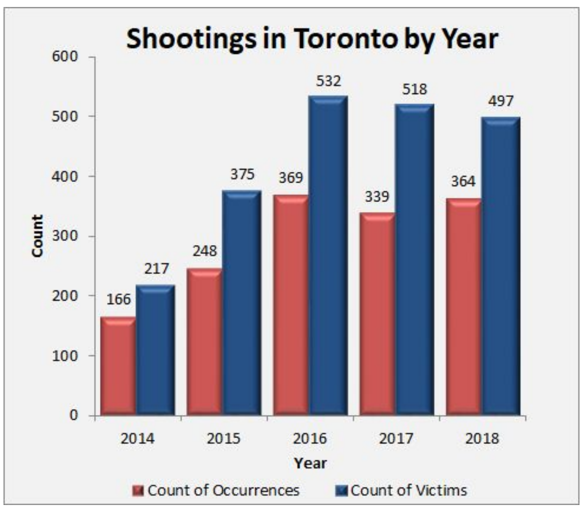

Figure 1: Shootings in Toronto from 2014-2018

Figure 2: Types of Homicides in Toronto in 2018

During these traumatic events, timing is crucial for saving human life. However, when responding to emergency situations, people can find themselves in a state of confusion, panic, and/or paralysis [2]. This results in individuals being unable to call police or assist victims. In some cases, it may be difficult for police to identify where the gunshots came from, especially in the example of the Las Vegas shooting. It took police about 72 minutes to locate the shooter [3], proving the importance in timing and emergency response.

1.2 BUSINESS OBJECTIVE

According to a study done by the Toronto Police Service (TPS), there is no one way to decrease gun violence in the city but multiple solutions to help solve the problem [4]. Our solution targets decreasing emergency response times through early detection and alerting. With smart city technologies, such as sensors, becoming more prominent, there is a large pool of data that can be used to our advantage.

Guardian will utilize acoustic sensors placed strategically throughout the city to immediately recognize gunshots. Reducing the time from when the incident occurs to when the emergency services are involved will result in many more lives saved.

1.3 BENEFITS AND LIMITATIONS

There are a number of benefits and limitations that can result from the implementation of a city wide sound classification system. Table 1 below outlines the benefits and limitations that can be found in different aspects.

Table 1: Benefits and Limitations of the System

ECONOMICAL	
+	Reduces the number of officers required for patrol and allows for already limited TP resources to be allocated to other functions
SOCIAL	

+	Encourages a sense safety within the community (although bystanders are still encouraged to call 911 when necessary)
-	Leaves residents feeling a sense of racial profiling, depending on which communities are selected to have the sensors
TECHNOLOGICAL	
+	Minimizes the lag in response time if there are no witnesses (late night/early morning) or no access to phones (children/elderly)
+	Extended application include notifying EMS of car accidents or fires
-	Reduced accuracy for shootings inside buildings
LEGAL	
+	Protects privacy, since the system currently identifies presence of gunshots and does not listen to conversations
-	Privacy will inevitably still be a concern for some residents

1.4 IMPLEMENTATION

Guardian will consist of acoustic sensors placed on city buildings and streetlights. Permits will be acquired for necessary, non-municipal buildings and property owners will be notified of the new installation. Installations will include any additional infrastructure that allows the program to work seamlessly with the TPS system. This includes mapping the locations of the acoustic sensors in reference to the police precincts.

An initial testing period of Guardian will be implemented in select Toronto neighbourhoods based on the frequency of gunshots in the area in 2018. Testing will be conducted in areas of high, medium and low frequencies of gunshot incidents to verify the accuracy of the algorithm (this also reassures the community regarding their concerns on profiling). After the testing period, the validity of the program will be verified. Community and police feedback, potential decrease in fatalities, and potential decrease in incidents will be analyzed. Any changes or improvements will be made before rolling out the program to the remainder of the city.

2.0 PROJECT APPROACH

Several steps were taken to train and test Guardian, this section will outline these steps and the rationale for the decisions made.

2.1 DATA SELECTION

Audio clips of various sounds were used to train the algorithm. Sounds such as fireworks, jackhammer, and airplanes were selected to mimic natural sounds of the city, as well as sounds that can be mistaken for gunshots. Table 2 below lists the various sounds used along with their respective class ID.

Table 2: Classifications IDs of the sound data

CLASS ID	SOUND CLASSIFICATION
1	Thunderstorm
2	Fireworks
3	Chainsaw
4	Drilling
5	Engine Idling
6	Gunshot

7	Jackhammer
8	Airplane
9	Train
10	Church Bells
11	Car Horn

Figure 3 shows the sounds of gunshots and fireworks in terms of their amplitude over time. The plots indicate how similar the two sounds are.

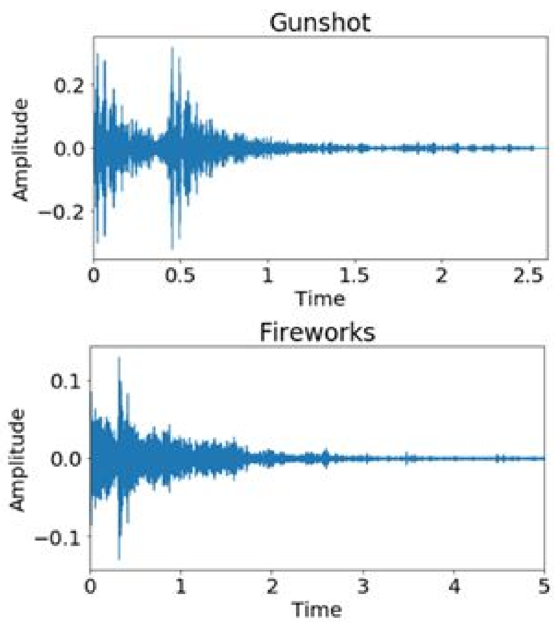

Figure 3: Sound wave plot comparing gunshots and fireworks

2.2 FEATURE EXTRACTION

Two features were extracted from the audio files to train and test the algorithm.

The first feature is Mel-frequency cepstral coefficients (MFCCs) which takes the human ear's perception to sensitivity with respect to frequencies into consideration and calculates the outputs which are coefficients unique to a particular sample. These coefficients represent the overall shape of the sound envelope which outlines the frequencies of the individual audio files.

The second feature used is the spectrogram which forms visual images by mapping out the frequency of audio files with respect to time as seen in figure 4 below. Amplitude is also depicted with colour (different spectrograms have their own unique colour schemes). These three parameters describe the timbre, or quality, of the sound.

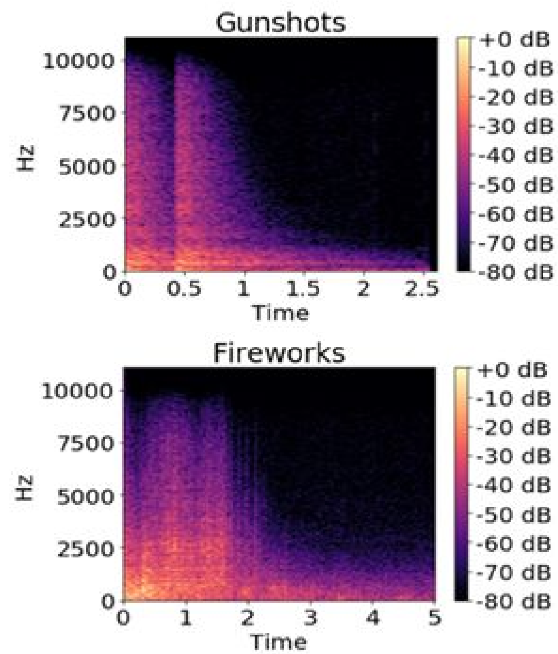

Figure 4: Spectrograms comparing gunshots and fireworks

2.3 TRAINING METHOD

The LibROSA library is used throughout Guardian. LibROSA handles audio analysis, loads and extracts the necessary features from the audio files. This process prepares the inputs for Convolutional Neural Network (CNN) training.

A CNN was used as the classification algorithm for the different sound classes. CNN's offer the following advantages to the proposed solution [5]:

- They scale well with size of input, thus reducing the number of variables and minimizing the risk of overfitting
- They are useful when there is structure in data and patterns are spatially related, such as in audio or image recognition applications
- They have the ability to learn and extract high-order relevant features from data without human aid

Tensorflow was used in the training of the convolutional neural network. The rectified linear unit (ReLU) was selected as the activation function for the hidden layers. ReLU was used due to it ability to reduce the likelihood of vanishing gradient (neurons in an earlier layer learning slower than neurons in a later layer, reducing the model accuracy) [6]. The softmax function was used to compute the probability of the classes in the output layer.

2.4 FINDINGS AND RESULTS

An accuracy of 83.1% was achieved following the testing and training of the algorithm. As seen in Figure 5 below, the accuracy and loss from training and evaluating Guardian was visualized on Tensorboard.

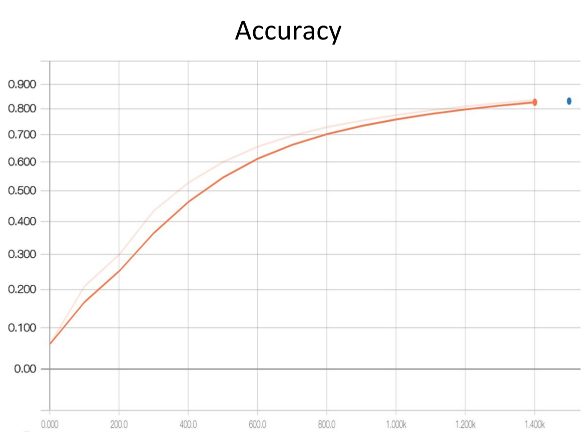

Figure 5. Accuracy and Loss of Guardian

A demonstration of Guardian's algorithm is seen in Figure 6 below. An audio file of each of the eleven classes were inputted and the resulting confidence was the output. Any confidence that falls below 80% will be classified as a sound not recognized by Guardian and dropped.

```
Prediction is "Gunshot"(100.0%),
correct sound label is"Gunshot".

Prediction is "Jackhammer"(100.0%),
correct sound label is"Jackhammer".

Prediction is "Drilling"(100.0%),
correct sound label is"Drilling".

Prediction is "Engine Idling"(100.0%),
correct sound label is"Engine Idling".

Prediction is "Thunderstrom"(99.9%),
correct sound label is"Thunderstrom".

Prediction is "Church Bells"(100.0%),
correct sound label is"Church Bells".

Prediction is "Jackhammer"(95.3%),
correct sound label is"Airplane".
```

Figure 6

2.5 MODEL IMPROVEMENTS

Currently, the algorithm is designed such that when the sound of a gunshot is heard, the police service will be notified to investigate the general area around the sensor. An improvement to the system would be to have a network of sensors in an area with coordinates of their position. This ensures that the exact location of the incident is triangulated when a gunshot is identified.

In addition, the model has been trained to classify sounds in isolation (only a gunshot and no other noise in the background). This may result in the possibility of a false positive or worse, a false negative occurring. A system improvement would be to introduce audio files that have a mix of unique sounds, making a more robust model that can detect the sound of a gunshot among other noises.

The final improvement to the model can be the addition of more functions. Sounds of car crashes and fires can be introduced to the algorithm and it can be trained to notify the EMS and firefighters, respectively, of the incident.

3.0 FINANCIAL PROSPECTIVE

The algorithm will work in conjunction with existing city initiatives to mitigate gun violence and the resulting negative outcomes. In 2018, the provincial government allocated $7.6 million for legal SWAT teams in Toronto courts [7]. The city allocated an additional $12 million to improve existing city-run programs for youth [8]. $3 million was also provided to the TPS to fund an additional 200 officers to work overtime for an 8 week pilot run [4]. The program was proven to be ineffective, as it reduced gun violence by 1 incident [4]. Adding a gunshot detection system, however, requires about $4 million for 2 years [9], and is a smaller cost by comparison.

The system cost and financial prospects can be approximated based on GE's CityIQ Nodes. In 2017, the smart sensors were installed on 5% of San Francisco's streetlights during GE's energy saving efforts to upgrade street lights with LED lights [10]. The sensors were used to monitor traffic, predict real time maintenance on the LEDs, and detect gunshot sounds, and are built for future applications including contacting emergency response and weather monitoring [10]. Accounting for the current and future applications of our system, the cost is estimated to be 2200 USD to develop and install each sensor [10].

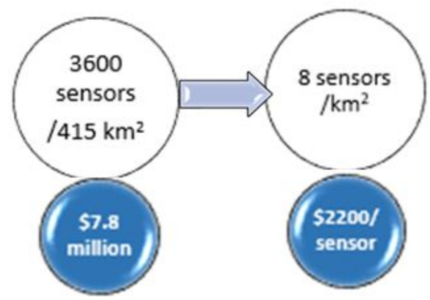

Figure 7: Cost and distribution of sensors for GE's CityIQ Nodes

A similar GE pilot program in Jacksonville in 2014 proved that upgrading to smart LED lights was an additional 48.7 million USD compared to converting streetlights to regular LEDs [11]. However, the market value for gunshot detection was predicted to be 790 million USD in 2017 [12]. The market value in 2025 is expected to reach 3.4 billion USD, with a compound annual growth rate of 20% between 2018 to 2025 [12]. This indicates that despite the current high cost for implementation, quick advances in technology will make it possible to develop the sensors at a lower cost. This is further driven by current gunshot detection systems having high customer renewal rates and partnering with cities and companies, such as GE, Intel and AT&T [13].

4.0 Conclusion

Guardian is one of many strategies that bring Toronto closer to efficiently tackling gun violence. It notifies EMS as soon as it detects a gunshot sound, and can distinguish gunshots from similar sounds, such as fireworks. Reducing detection time ensures that EMS respond in a timely manner, especially when there are no bystanders around or the shooter is difficult to locate. Guardian encourages a sense of safety in the community, and is a catalyst to saving lives.

Future additions to the sound data can enhance the system performance to tackle other issues in the city. These include:

1	Dispatching EMS during a car accident
2	Dispatching fire services during a fire
3	Monitoring noise pollution as per municipal code chapter 591

5.0 References

[1] https://app.powerbi.com/view?r=eyJrljoiNGI4YzgwYzMtNjk4MC00ZTdjLTljZGQtYzA0NzRjMDQwMzc0IiwidCI6Ijg1MjljMjI1LWFjNDMtNDc0Yy04ZmI0LTBmNDA5NWFlOGQ1ZCIsImMiOjN9

[2]https://www.forbes.com/2008/09/02/crisis-reaction-health-forbeslife-cx_avd_0902health.html#ca4c6c95391f

[3]https://www.iotworldtoday.com/2017/10/04/how-smart-building-technology-could-address-mass-shootings/

[4]https://www.cbc.ca/news/canada/toronto/police-summer-shooting-toronto-1.4821119

[5]https://www.oreilly.com/library/view/deep-learning/9781491924570/ch04.html#convo_neu_net

[6] https://towardsdatascience.com/activation-functions-and-its-types-which-is-better-a9a5310cc8f

[7]https://globalnews.ca/news/4378036/toronto-police-funding-gun-violence/

[kw2]https://www.cbc.ca/news/canada/toronto/gun-violence-town-hall-1.4835312

[8] https://globalnews.ca/news/4344093/controversial-gunshot-detector-shotspotter-toronto-police/

[9] https://www.marketwatch.com/press-release/gunshot-detection-system-market-2018-global-a

[10] https://www.ge.com/reports/light-bright-san-diego-leads-way-future-smart-cities/

[11] https://www.bizjournals.com/jacksonville/news/2016/05/12/remember-those-smart-streetlights-ge-announced.html

[12] https://www.marketwatch.com/press-release/gunshot-detection-system-market-2018-global-analysis-opportunities-and-forecast-to-2025-2018-07-27

[13] https://observer.com/2017/05/shotspotter-ipo/

EMS response time (to include if space): https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6156551/