

# System for an acoustic detection, localisation and classification

Jakub Svatos<sup>1</sup>, Jan Holub<sup>1</sup>, Jan Belak<sup>1</sup>

<sup>1</sup> Department of Measurement, Czech Technical University in Prague, Prague 166 27, Czechia

## ABSTRACT

Currently, acoustic detection techniques of gunshots (gunshot detection and its classification) are increasingly being used not only for military applications but also for civilian purposes. Detection, localisation, and classification of a dangerous event such as gunshots employing acoustic detection is a perspective alternative to visual detection, which is commonly used. In some situations, to detect and localise the source of a gunshot, an automatic acoustic detection system, which can classify the caliber, may be preferable. This paper presents a system for acoustic detection, which can detect, localise and classify acoustic events such as gunshots. The system has been tested in open and closed shooting ranges and tested firearms are 9 mm short gun, 6.35 mm short gun, .22 short gun, and .22 rifle gun with various subsonic and supersonic ammunition. As 'false alarms', sets of different impulse acoustic events like door slams, breaking glass, etc. have been used. Localisation and classification algorithms are also introduced. To successfully classify the tested acoustic signals, Continuous Wavelet and Mel Frequency Transformation methods have been used for the signal processing, and the fully two-layer connected neural network has been implemented. The results show that the acoustic detector can be used for reliable gunshot detection, localisation, and classification.

**Section:** RESEARCH PAPER

**Keywords:** Acoustic Detection; Gunshots; Localisation; Classification; Neural Network; Mel Frequency

**Citation:** Jakub Svatos, Jan Holub, Jan Belak, System for an acoustic detection, localisation and classification, Acta IMEKO, vol. 10, no. 2, article 10, June 2021, identifier: IMEKO-ACTA-10 (2021)-02-10

**Section Editor:** Giuseppe Caravello, Università degli Studi di Palermo, Italy

**Received** January 18, 2021; **In final form** April 15, 2021; **Published** June 2021

**Copyright:** This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Corresponding author:** Jakub Svatos, e-mail: [svatoja1@fel.cvut.cz](mailto:svatoja1@fel.cvut.cz)

## 1. INTRODUCTION

An acoustic detection (AD) of gunshots is a present topic that can help to detect hazardous and dangerous events, especially in public areas. In recent days, there is an increase in gunshot attacks in public areas such as schools, campuses, hospitals, and shopping centres. In some cases, it is challenging to identify a dangerous event from uncertain, inadequate data received by cameras or by security staff. The main asset of AD is based on the extraction of vital information from the recorded signal data and classify it as due to a given event (gunshot, a human scream, glass breaking, etc.) Due to this classification, AD can assist in Law enforcement to better discriminate dangerous events and intervene in the ongoing process on time and decrease the possibility of casualties. A fundamental goal of AD systems is to monitor acoustic signals around the area of interest and to detect a potentially hazardous event, record, localise the probable position of the event and classify the event into categories as an alert (gunshot) signals and non-treat, 'false alert', signals.

There are several experimental or commercial AD systems designed for gunshot events detection and localisation available on the market [1]-[5]. These systems are designed to localise the source of the gunshot and to estimate the type of treat. The more sophisticated systems can even identify the particular firearm type using advanced classification methods. A drawback of the sophisticated systems is usually very high purchase and operational cost, which makes it almost impossible to use for smaller non-governmental entities such as campuses or hospitals.

To successfully design a gunshot detecting and classifying algorithm, essential characteristics of its complex physics have to be understood. A comprehensive explanation of the problem can be found in [6], [7].

A gunshot pattern is characterised by the two phenomena, muzzle blast, and, if the bullet propagates at supersonic speed, shock wave. Muzzle blast is caused by an explosive charge where hot, high-pressured gases expand as acoustic energy from the centre of the gun barrel. The bullet travelling at supersonic speed generates a shockwave effect, which is propagating in a conic fashion behind the bullet trajectory. The shockwave is based on

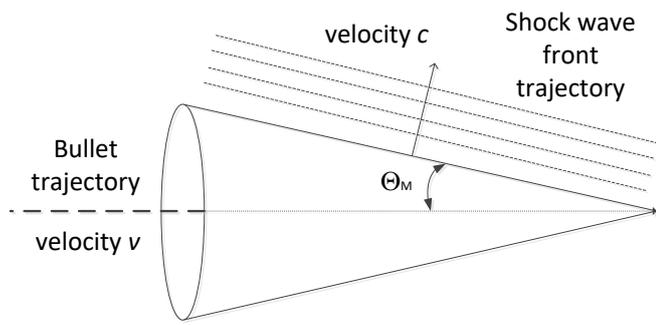


Figure 1. An acoustic trace of a supersonic bullet.

the combinations of compression and expansion shock, as shown in Figure 1.

The angle  $\Theta_M$  between the bullet trajectory and the shock wave trajectory is given by the Mach number  $M = v/c$ , where  $v$  is the velocity of the projectile and  $c$  is the speed of the sound:

$$\theta_M = \arcsin\left(\frac{1}{M}\right). \quad (1)$$

These factors can include valuable information that can be used for improving the detection capability of the AD system. Alongside this, the calibre of both bullet and barrel, the length of the latter, mechanical action caused by the gun itself, or even the chemical properties of the propellant cause different effects on the pattern of a gunshot. Last but not least, the temperature of the air, air humidity, wind speed, environment (e.g., foliage density, urban area) and soil characteristics also have an impact on the resulting gunshot pattern. Considering all these phenomena, to effectively detect and identify a gunshot, signal processing, including adaptive filtering and advanced data processing and classification have to be carried out [8] – [11].

An example of a typical subsonic pattern and a supersonic gunshot signal are shown in Figure 2 and Figure 3. Both bullets were shot in semi-open area (with multiple reflections from the ground and the walls) by the same 9 mm short gun, and its acoustic signature was recorded at the distance of approximately 10 m. In Figure 3, the shock wave pattern is not clearly visible due to its proximity to the muzzle blast pattern of the signal. It is caused by the relatively low speed of the supersonic bullet (Mach number  $M = 1.09$ ).

In this article, a system for acoustic detection, localisation and classification is introduced. The proposed system consists of stand-alone sensor units, which are placed around the monitored areas in sufficient numbers, to continuously monitor acoustic

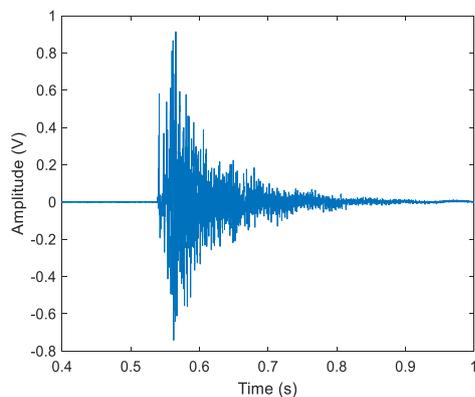


Figure 2. Signal corresponding to a 9 mm subsonic short gun gunshot.

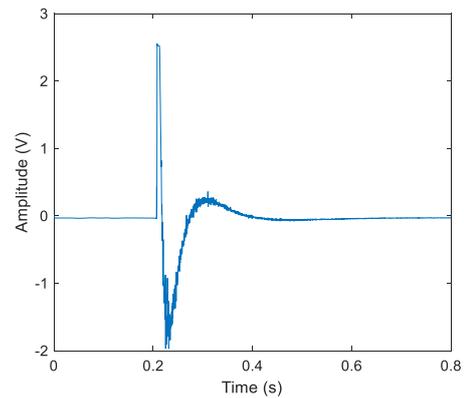


Figure 3. Signal corresponding to a 9 mm supersonic short gun gunshot.

events around the unit; if there is a possibility of a dangerous event, the stand-alone units send the data to the remote unit for advanced gunshot detection and classification. The remote central unit evaluates signals received from multiple sensor units and, using advanced signal processing and classification methods, determines the location of a gunshot and the probable firearm caliber used. To localise the place of the event, at least three units equipped with a microphone are needed. The localisation accuracy of the system depends on the density and the number of stand-alone sensor units used.

In comparison with other existing available systems (e.g. [3] or [5]), the presented system has a novel modular flexible structure. It can be deployed on a building or moving object (car, person, animal) while the central unit can be installed in a distant protected place. In the future, the presented acoustic detector can be used in public areas like schools, campuses, shopping centers to detect and localise gunshots.

The paper is organised as follows. In Section II, the sensor units, detection algorithm, localisation algorithm and signal processing, together with the classification methods, are introduced and described. In Section III, the experimental measurements and results are presented. The conclusion and future work directions are outlined in Section IV.

## 2. METHODS

The presented system for acoustic detection, localisation and identification consists of multiple (at least three to estimate the correct localisation of the event) sensor units and one central unit. Such a topology enables additional analyses at the central unit, i.e., advanced location of the acoustic event position using timestamped data from multiple sensor units receiving the acoustic signal related to the possible shooting event. Each sensor unit has to cover analogue signal pre-processing, digitisation, simple detection algorithm and simple evaluation. Based on the simple detection and evaluation algorithms, all the units with positive detection, sends the recorded data to the central remote unit. Every stand-alone unit works on the principle illustrated in Figure 4.

The stand-alone sensor detection algorithm works on the principle of a modified median filter introduced by the authors

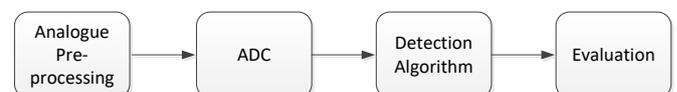


Figure 4. Unit function requirements.

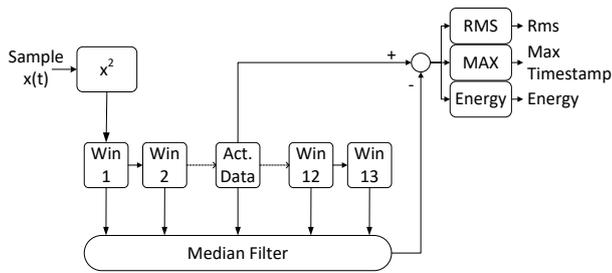


Figure 5. The principle of the median filter.

in [8]. The background signal is filtered by the median filter while an acoustic event is processed based on the algorithm presented in Figure 5.

It works on the principle of dividing the recorded acoustic signal, converted to energy (the input signal is squared), to the optimal number of data blocks - time windows, the defined odd number of windows, which are fed to the median filter. The middle window is the block with the actual data (actual data window), where the detection is performed. Each window represents 50 samples, i.e., 1 ms of recorded acoustic data around the unit. The data contained in the window before the actual data window represents the acoustic signal recorder after the actual data and vice versa, data recorded before are contained in windows that follows the actual data window. The resulting signal from the median filter is then subtracted from the actual data window, to suppress/filter the background acoustic noise. Thereafter the detection algorithm based on multiple thresholds distinguishes actual gunshot events from other 'false alarms'. More details about the algorithm can be found in [12] and can be described by the pseudocode in Figure 6.

Requirements for flexible modular sensor units are even more demanding than just to detect, record and send the acoustic event in case of a positive detection. Since the sensor units can be deployed everywhere, they have to be able to send their exact position. Moreover, all the sensor units have to be precisely time-synchronised to get a synchronised timestamp of the detected event and to send the data to the central unit wirelessly ensuring accurate localisation and classification of the event secured by the remote server.

To fulfil all these criteria, the sensor unit was designed according to Figure 7.

Every sensor unit uses a pre-polarised, electret condenser microphone with a flat frequency response from 20 Hz to 20 kHz and omnidirectional sensitivity, see Figure 8 and Figure 9.

```

If (Max > Sensitivity Threshold Level) &&
  (Max > (Rms * RMS Threshold Level))
{
  If (energy (Actual Data Window) > Energy Threshold
    Level)
  {
    If (energy (Actual Data Window + 1) > energy
      (Actual Data Window - 1))
    {
      Gunshot Detected (t1, t2, t3, t4)
    }
  }
}

```

Figure 6. The detection algorithm pseudocode.

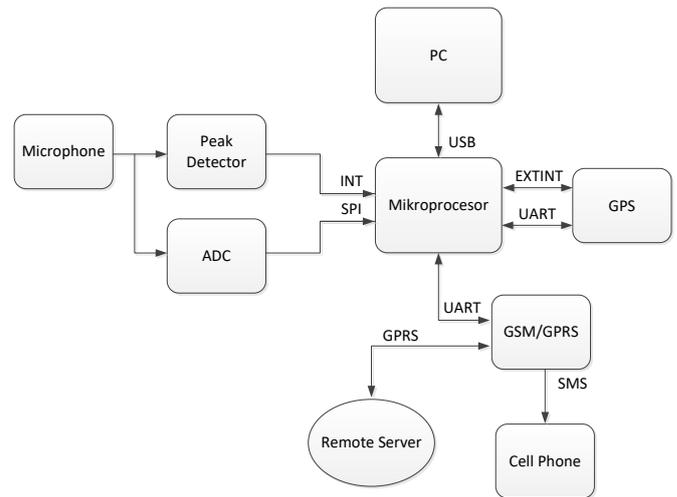


Figure 7. Block diagram of the sensor unit.

The heart of the acoustic sensor unit is a low power consumption, 32-bit LPC 1837, ARM Cortex-M3 based microcontroller, which processes the recorded data from the microphone. The analogue Op-amp based peak detector provides an interrupt for the microcontroller if the microphone records an acoustic impulse event. A 16-bit Analogue-to-Digital Converter (ADC) ADS8866 digitises the recorded analogue data. The successive approximation, low power consumption ADC samples the data with a sampling frequency of  $f_s = 44$  kHz. After the interrupt, the ADC sends the data to the microcontroller via Serial Peripheral Interface (SPI). The data is stored in a circular buffer as a 16-bit number. In this way, the data is ready for processing.

If a trigger from the peak detector occurs, the median filter algorithm described above evaluates if there is the possibility of a gunshot event. If the possible gunshot is evaluated and detected, low power consumption GSM chip GL865-QUAD V3

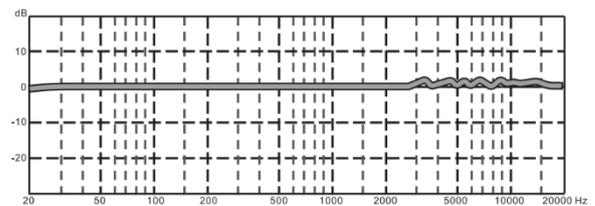


Figure 8. Frequency response of the microphone.

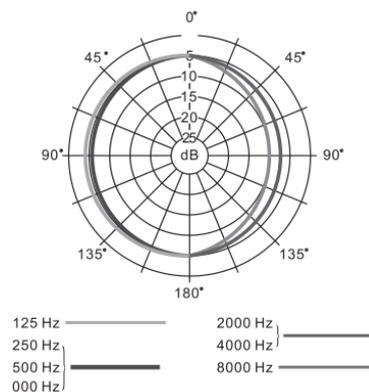


Figure 9. Omnidirectional sensitivity of the microphone.

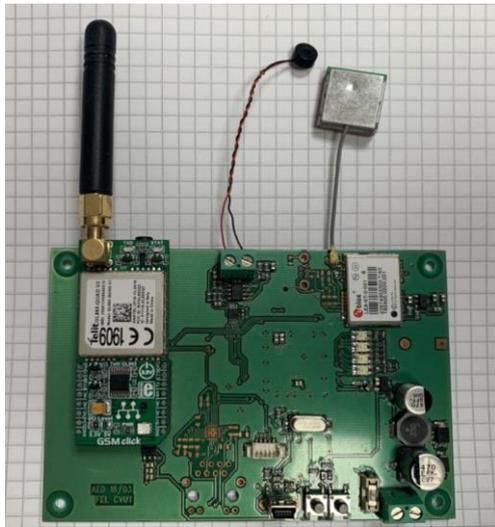


Figure 10. The sensor unit for acoustic detection.

then sends the recorded data, together with the exact position of the unit and a precise timestamp, when the event has been recorded, for further analysis to the remote central unit PC.

The precise localisation of the sensor unit position and synchronisation with all the other sensor modules secures GPS module NEO/LEA-6T. When the NEO/LEA-6T module operates stationary, GPS timing is possible with only one visible satellite. It means the time can be maintained even under adverse signal conditions or in an environment with poor sky visibility. The GPS module is also adding the precise timestamp to the recorded data to make it possible to assess the exact localisation of the gunshot event. The UCT timestamp and the exact position of each unit are determined by this way. For the synchronisation of the units, the NMEA protocol [13] is used. The time accuracy for the synchronisation of all units is 0.1 ms. The described sensor unit is power supplied from  $U_{cc} = +5V$ . The designed sensor unit with a testing microphone is shown in Figure 10.

The central remote unit process received data using advanced signal processing. When at least three sensor units detect the event and send the data together with its timestamps, an algorithm uses these timestamps from the network of sensor units to triangulate the location of the event. The remote server also secures the classification to state if there is a gunshot detected and identify/estimate the caliber of the used firearm or if it is only a 'false alarm'.

Figure 11 describes the layout of the localisation problem using three sensor units/microphones.

Measuring the time delay  $\tau_n$  between the event and each microphone  $M_n$  and knowing their positions, it is possible to calculate the position of the source  $S$  inside of the domain  $D$ . The relationship between the source  $S$  of the gunshot and three

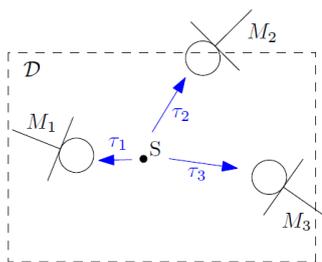


Figure 11. The layout of the situation with three microphones.

microphones  $M_n$ , that detect it, is described by equations (2) to (4)

$$(\tau_1 - \tau_0) \cdot c = (x - a_1)^2 + (y - b_1)^2 \quad (2)$$

$$(\tau_2 - \tau_0) \cdot c = (x - a_2)^2 + (y - b_2)^2 \quad (3)$$

$$(\tau_3 - \tau_0) \cdot c = (x - a_3)^2 + (y - b_3)^2, \quad (4)$$

where  $\tau_0$  is the time where the gunshot has occurred;  $\tau_1$ ,  $\tau_2$  and  $\tau_3$  are the times where the gunshot has been detected by unit 1, 2 and 3;  $(a_s, b_s)$  are the coordinates of the relevant unit; and  $(x, y)$  are the coordinates of the gunshot. The resulted coordinates are necessary to recalculate to the Cartesian system. The unit that detects the gunshot first ( $\tau_1$ ) is considered as it is placed in coordinate  $(0, 0)$ . The coordinates in meters of other units are then calculated through (5):

$$d = \text{acos}(\sin(\Phi_1) \sin(\Phi_2)) + \cos(\Phi_1) \cos(\Phi_2) \cos(\delta\lambda) R, \quad (5)$$

where  $\Phi_1$  and  $\Phi_2$  are the coordinates in meters and  $\delta\lambda$  is the difference of the longitudes and  $R$  is the mean Earth radius (6378 km). Once the coordinates of the Cartesian system are calculated, it is necessary to recalculate the latitude and the longitude coordinates. Since the distance of one degree of longitude is different at the North Pole and the Equator, it is necessary to know the relationship between degree and meter at a given latitude [14]. This is done by applying the simplified formulas (6) and (7), which recalculate meters to one degree of latitude of longitude, resulting in uncertainty values in the order of centimetres:

$$x_{\text{latitude}} = 111123.92 - 559.82 \cos(2 \lambda) + 1.175 \cos(4 \lambda) - 0.0023 \cos(6 \lambda) \quad (6)$$

$$x_{\text{longitude}} = 111412.84 \cos(\Phi) - 93.5 \cos(3 \Phi) - 1.175 \cos(5 \Phi) \quad (7)$$

An example of the received timestamps from the three sensor units (AED 1 – AED 3) with their exact position in longitude and latitude coordinates and calculated source of a detected acoustic event by the remote server is in Figure 12.

A custom software application for the remote unit PC to communicate with sensor units, processing the sent data and displaying the location of the event was developed as a part of the acoustic detection system (Figure 13). The application was developed in C# programmable language. The program shows a

```

--- AED 1 ---
Latitude: 50.0
Longitude: 14.00
Timemark [ms]: 5193
Last captured: 26 Dec 18 02:01:12. PM
-----
--- AED 2 ---
Latitude: 50.01
Longitude: 14.0025
Timemark [ms]: 5193
Last captured: 26 Dec 18 02:02:01. PM
-----
--- AED 3 ---
Latitude: 50.0
Longitude: 14.005
Timemark [ms]: 5203
Last captured: 26 Dec 18 02:00:54. PM
-----
--- Source of acoustic event ---
Latitude: 50.004866764791
Longitude: 14.002504288864

```

Figure 12. An example of a message with calculated source of an acoustic event sent by the sensor unit.

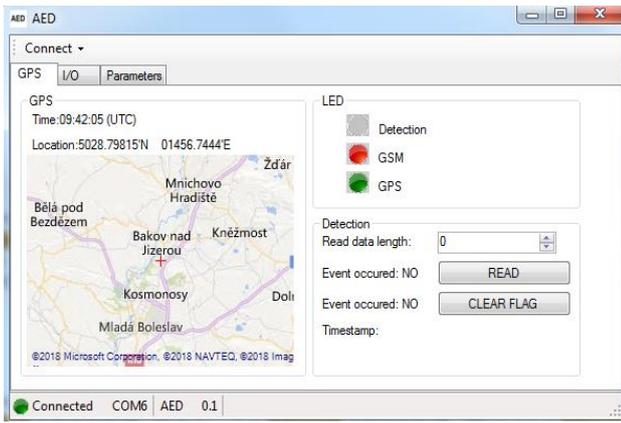


Figure 13. PC application.

map with the exact location of the chosen acoustic unit and precise/synchronised UTC time. In case of detection, the application signals the event, displays an exact timestamp in UTC and stores the data for further processing. Through the application, the sampling frequency, parameters for the detection algorithm based on the median filter described in the beginning of the Section II, and the connection of additional detection units can be done. Parallel to localisation, the remote central unit applies the algorithms to obtain the features for classification and classify the event into the classes.

The Mel Frequency Transformation (MFT) and Continuous Wavelet (CW) methods are used for the processing of received data to obtain features for the further classification. To identify the components in an acoustic signal, it is preferable not to use the linear frequency scale, but to detect the differences in lower frequencies rather than higher frequencies. Incorporating this can be preferable in acoustic detection. MFT algorithm uses band-pass filters to get the energy of the signal for each defined band. Then, the algorithm uses frequency distribution to create Mel Frequency Coefficients (MFC). MFC is computed using cosine transformation on the logarithm of bank energies. It can be described as follows. The waveform is divided into the Frames, then the discrete Fourier Transform is applied, Logs of amplitude spectrum is taken, the scale is converted using the Mel-scaling and, finally, the discrete cosine transform is performed. To convert the frequencies into the Mel Scale, equation (8) is used. More details on MFC can be found in [15], [16]

$$M(f) = 1125 \ln \left( 1 + \frac{f}{700} \right). \quad (8)$$

Through experiments, it was set the optimal number 20 MFC and filters with frequency bands from 500 Hz to 5kHz. These MFCs served as features for the future classification. To limit the influence of an echo on a classification and increase its validity and reliability, more features have to be added.

For this reason, the CW algorithm was considered. Unlike Discrete Fourier Transform DFT, CW transform uses defined waves to create a frequency spectrum with significant time resolution. The time resolution allows limiting the influence of echo, because the echo does not usually interact with the beginning of the impulse acoustic event/gunshot pattern.

The best results for a gunshot recognition showed the Bump wavelet served as a mother wavelet. The used Bump wavelet is a bandlimited function defined in the frequency domain by (9), and its shape is shown in Figure 14.

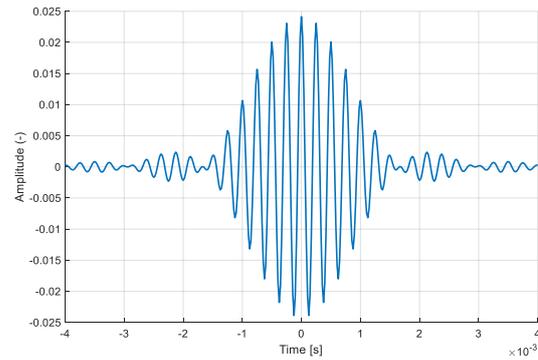


Figure 14. Bump mother wavelet.

$$\Psi(s\omega) = e^{-\frac{1}{1-(s\omega-\mu)^2/\sigma^2}} 1_{\left[\frac{\mu-\sigma}{s}, \frac{\mu+\sigma}{s}\right]} \quad (9)$$

where  $1_{\left[\frac{\mu-\sigma}{s}, \frac{\mu+\sigma}{s}\right]}$  is the indicator function,  $s$  is the scale,  $\omega$  is angular frequency,  $\sigma$  is standard deviation, and  $\mu$  is mean value. More details about the wavelet can be found in [17], [18].

Both presented algorithms are used as features for the advanced classification of received signals by central remote unit from sensor units. Various sets of gunshots, as well as several diverse false signals, similar to a gunshot signal were used for classification. Each false signal was chosen to be challenging to differentiate it from the real gunshots by a human operator.

For classification, two independent Neural Networks (NN) were created. First, a NN was designed to classify signals based on MFC. Considering the low dimensionality of these coefficients, a fully two-layer connected neural network has been used. For the CW algorithm, a convolution neural network was used. As the spectrum is a two-dimensional matrix, the convolution network is significantly better than a fully connected system. Convolution network allows finding local features in multidimensional, location-dependent input data. Spectrogram features are location-dependent, and therefore, the convolution network leads to better results than a typical fully connected network. Also, the convolution network is typically much smaller in dimension, thus significantly reducing the computation time. A single network combining both designs was implemented and used for the classification. The network had two inputs, one for the MFC and one for the CW algorithm.

Along with the NN, two more classifiers, Support Vector Classifier (SVC) and Naive Bayes Classifier (NBC) [19] - [21] are presented herein for a results comparison. Both classifiers were implemented in Matlab software with classification pattern recognition toolbox PRTools ver. 5 [22]. Naive Bayes method is based on Bayesian theorem (10).

$$P(A/B) = \frac{P(B/A) P(A)}{P(B)} \quad (10)$$

where  $P(A)$ ,  $P(B)$  is the prior probability of  $A$ ,  $B$  respectively, and  $P(B/A)$  is the conditional probability, the probability of  $A$  given that  $B$  is true. The Naive Bayes Classifier estimates for every feature and every class separately. Naive Bayes method is simple, but despite its simplicity can often outperform other sophisticated classification methods. It is widely used for results comparison with other classifiers. SVC works based on the construction of non-linear decision boundaries. SVM creates a decision plane, so-called Hyperplane, which in Feature space separate training data optimally. The decision plane then

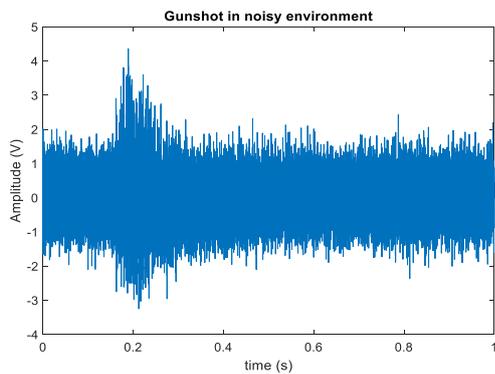


Figure 15. Recorded signal corresponding to 9 mm subsonic short gun shoot noised by a diesel engine.

separates sets of objects, which have different class memberships. The input data are transformed (mapped) to Feature space using a mathematical function known as the kernel. Support vector models can be linear, polynomial, sigmoidal and radial basis function.

### 3. RESULTS

To acquire the test data, multiple gunshots measurements in different shooting ranges, closed – with many reflections and open, were done. For increasing the diversity of gunshots, three different calibres were measured [23]. Tested firearms were 9 mm short gun, 6.35 mm short gun, .22 short gun, .22 rifle gun. Various subsonic and supersonic (up to Mach number of  $M = 1.1$ ) ammunition was used with the 9 mm short gun. Each type of gun was measured at least 60 times. In total, approximately 400 samples corresponding to gunshots were and recorded by sensor units described in Section II (in many cases, one sample has been recorded by more sensor units). For false signals, acoustic impulse events similar to gunshots were measured. Approximately 200 samples of various false alarms were measured. As examples of false signals, glass breaking, different slams (such a door slam), handclaps, or close big bubble wrap popping were recorded. At least three sensor units placed in nearby locations recorded all the tested signals. The example of a gunshot of a 9 mm short gun in a noisy environment (diesel engine) recorded at a distance of 10 m is shown in Figure 15.

Figure 16 shows another recorded gunshot, shoot by .22 short gun in closed shooting range where many reflections are visible. The gunshot was recorded at a distance of 12 m.

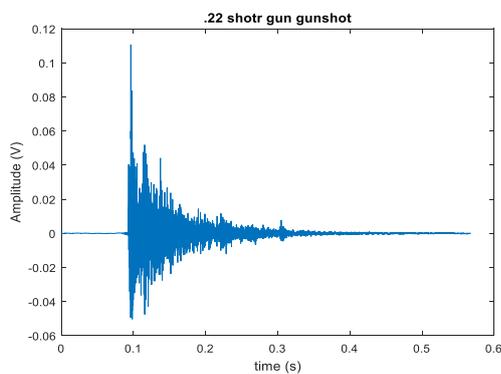


Figure 16. Recorded signal corresponding to .22 mm subsonic short gun shoot with reflections.

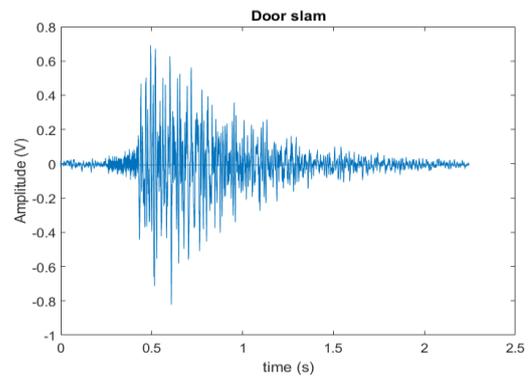


Figure 17. Recorded signal corresponding to a door slam.

Two examples of recorded false alarms represented by a door slam and bubble wrap popping are presented in Figure 17 and Figure 18.

Both presented false signals exhibit similar patterns as gunshots. Therefore, the recorded data have to be processed in a way to emphasise the differences. This can be done by MFT and CW methods presented in Section II.

This work presents the classification into ‘false alarms’ and gunshot event and in the case of a gunshot, the classifier classifies the recorded signal into an individual caliber.

To train the classifiers in an optimal way, the recorded acoustic events data have to be divided into two groups – to training and to validation sets. There are many articles dealing with an optimal number of training data. One of the proposed methods is the use of the learning curve. Use of learning curves for optimal size of training data is described in [24]. From the Learning curve, dependency of the model performance on the training data size can be obtained. It usually depends on the classification method, the complexity of the classifier or how well can be the classes separated. Optimal size of training data can be determined from the maximum of the Learning curve. Recorded ‘false alarms’ and gunshots samples were chosen randomly chosen for the training and the validation sets. In this case, the number of training data is optimal for approximately 30% of the samples.

The training set had approximately 200 independent measurements and the validation set had approximately 400 samples. The experimental results for the validation data to be classified into false alarms and gunshots events are shown in Table 1.

The results show the validation success rate of classifying into the gunshots is 100 %. The system can classify the gunshot from

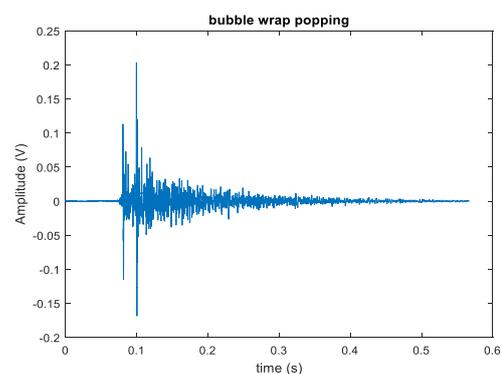


Figure 18. Recorded signal corresponding to a bubble wrap popping.

Table 1. Validation data classification into gunshot and false alarms results.

Classifier	NN	SVC	NBC
Gunshot	100%	100%	89.4%
False Alarm	73.9%	79.4%	60.7%

the ‘false alarm’ with a 100% success rate but, on the other hand, approximately 30% of the false alarms are identified as gunshots. The main reason for this is the limited number of signals corresponding to the false alarms and their diversity. To improve the classifier learning, a significant number of impulse acoustic events similar to gunshots have to be recorded in real scenarios and used it for classifier training. On the other hand, if the acoustic detection system will be complemented by visual detection an operator can restrain a False Alarms Occurrences. The results also show that SVC achieves slightly better results than NN in false alarm detection, while NBC’s success rate is below 90% even for a proper classification into gunshots classes. For this reason, NBC is not used for the classification into individual calibers.

Table 2 shows the results of the classification into individual calibers by the NN. The NN performs well in classification into the caliber of a gunshot, where it was able to classify more than 95% correctly for a 9 mm caliber, more than 82 % for 6.35 mm caliber and more than 96% for .22 caliber. On the other hand, NN classifies almost 26% of false alarms as gunshots (9 mm caliber mostly). For the sake of comparison, results obtained through SVC are summarised in Table 3.

The SVC classified almost 80% of false alarms correctly, but the classification of gunshots into correct caliber classes is less successful than for NN. For example, 6.35 mm caliber successful classification is less than 80% in comparison with 96% for the NN. The efficiency of NN classification strongly depends on the amount and variability of the false alarm signals. Based on the tested results, only the NN classifier was implemented to the ADS. In the future, more false alarm scenarios have to be recorded and used for the classifier training.

The tested measurements show that the presented ADS can successfully detect, locate and classify the impulse event. The ADS consists of sensor units operates on the principle of the modified median filter, and central remote unit, which calculates the location of the acoustic event and classify it using NN and CW and MFC algorithms. The tested acoustic signals were taken at different open and closed shooting ranges, where at least three acoustic units recorded the detected impulse acoustic event. The NN classifier can classify an individual caliber of a used firearm with a very high success rate, which is for some calibers more than 95%. As a next step, the System for Acoustic Detection will be deployed in an environment (residential area) to test it in real conditions.

#### 4. CONCLUSION

The system for acoustic detection, localisation, and classification into a firearm caliber was presented. The system

Table 2. Gunshot data classification into individual calibres by NN.

original/ predicted	false alarm	6.35 mm caliber	9 mm caliber	.22 caliber
false alarm	68	2	17	5
6.35 mm caliber	0	55	10	2
9 mm caliber	0	0	78	4
.22 caliber	0	3	0	80

Table 3. Gunshot data classification into individual calibres by SVC.

original/ predicted	false alarm	6.35 mm caliber	9 mm caliber	.22 caliber
false alarm	73	2	14	3
6.35 mm caliber	0	46	11	10
9 mm caliber	0	6	70	6
.22 caliber	0	6	3	74

consists of sensor units that continuously monitor acoustic events around the unit and the central remote unit. The sensor units use a modified median filter algorithm to state if there is a possibility of a gunshot. The central remote unit – PC, then evaluates the signal through advanced signal processing and classification to determine if there is a gunshot or a false alarm event. The system can localise the event on the principle of measuring the time delay between the event and each microphone of the sensor unit and knowing their positions. The accuracy of the localisation depends on the number and the density of the sensor units. The central remote units use Continuous Wavelet and Mel Frequency Transformation methods to get the features for a neural network classifier.

Gunshots of different calibers and various false alarms similar to gunshots were recorded on shooting ranges to test the system. More than 600 signals were recorded and tested.

The system shows the ability to detect the gunshot with 100% accuracy and to correctly classify the caliber of a gun with a high accuracy depending on the individual caliber. Considering the limited size of false alarms training dataset, such results are impressive. However, all measurements were measured in a similar environment. To practically employ the ADS in real conditions, a significantly larger dataset in a real environment, such as urban areas, should be examined for future tests and unit improvements.

In the future, the proposed Systems for Acoustic Detection can be used as a standalone unit placed in schools, campuses, shopping centers or other public areas in general, to detect, localise and classify gunshot events and to increase the safety of the civil population.

#### ACKNOWLEDGEMENT

This research was supported by the “Energy for Smart Objects” grant provided by Electronic Components and Systems for European Leadership Joint Undertaking in collaboration with the European Union's H2020 Framework Programme (H2020/2014-2020) and National Authorities, under grant agreement n° 692482.

#### REFERENCES

- [1] B. Kaushik, D. Nance, K. K. Ahuja, A Review of the Role of Acoustic Sensors in the Modern Battlefield, 11th AIAA/CEAS Aeroacoustics Conference, Monterey, California, 23-25 May 2005. DOI: [10.2514/6.2005-2997](https://doi.org/10.2514/6.2005-2997)
- [2] H. E. Bree, The microflow, dissertation thesis, 1997. <https://ris.utwente.nl/>
- [3] SHOTSPOTTER. Online [Accessed 04 June 2021] <http://www.shotspotter.com>
- [4] J. Millet, B. Baligand, Latest Achievements in Gunfire Detection Systems, Battlefield Acoustic Sensing for ISR Applications (pp. 26-1–26-14) NATO RTO-MP-SET-107, 2006. Online [Accessed 04 June 2021] <http://www.rto.nato.int/abstracts.asp>
- [5] G. L. Duckworth, D.C. Gilbert, J.E. Barger, Acoustic counter-

- sniper system, command control, communications, and intelligent systems for law enforcement, SPIE Proceedings 2938 (1996). DOI: [10.1117/12.266747](https://doi.org/10.1117/12.266747)
- [6] R. C. Maher, Modeling and signal processing of acoustic gunshot recordings, 2006 IEEE 12th Digital Signal Processing Workshop & 4th IEEE Signal Processing Education Workshop, Teton National Park, WY, pp. 257-261, 24-27 September 2006. DOI: [10.1109/DSPWS.2006.265386](https://doi.org/10.1109/DSPWS.2006.265386)
- [7] R. C. Maher, Acoustical characterization of gunshots, 2007 IEEE Workshop on Signal Processing Applications for Public Security and Forensics, Washington, DC, USA, pp. 1-5, 2007. DOI: [10.1109/IEEECONF12259.2007.4218954](https://doi.org/10.1109/IEEECONF12259.2007.4218954)
- [8] K. Łopatka, J. Kotus, A. Czyżewski, Detection, classification and localization of acoustic events in the presence of background noise for acoustic surveillance of hazardous situations, *Multimedia Tools and Applications* 75 (2016), pp. 10407-10439. DOI: [10.1007/s11042-015-3105-4](https://doi.org/10.1007/s11042-015-3105-4)
- [9] J. Sallai, W. Hedgecock, P. Völgyesi, András Nádas, György Balogh, Akos Ledecz, Weapon classification and shooter localization using distributed multichannel acoustic sensors, *Journal of Systems Architecture - Embedded Systems Design* 57(10) (2011), pp. 869-885. DOI: [10.1016/j.sysarc.2011.04.003](https://doi.org/10.1016/j.sysarc.2011.04.003)
- [10] M. Lojka, M. Pleva, E. Kiktová, J. Juhár, A. Čížmár, Efficient acoustic detector of gunshots and glass breaking, *Multimedia Tools and Application* 75 (2016), pp. 10441-10469. DOI: [10.1007/s11042-015-2903-z](https://doi.org/10.1007/s11042-015-2903-z)
- [11] J. Tomlain, O. Teren, Wireless solutions for outdoor training polygons, *Proceedings of the International Conference on New Trends in Signal Processing, Armed Forces Academy*, 12-14 Oct. 2016, Demanovska, Slovakia, pp. 87-90. DOI: [10.1109/NTSP.2016.7747791](https://doi.org/10.1109/NTSP.2016.7747791)
- [12] J. Svatos, J. Holub, Smart acoustic sensor, 5th International Forum on Research and Technologies for Society and Industry: Innovation to Shape the Future, Florence, Italy, 9-12 Sept. 2019, pp. 161-165. DOI: [10.1109/RTSI.2019.8895591](https://doi.org/10.1109/RTSI.2019.8895591)
- [13] K. Betke, The NMEA 0183 Protocol, 2010. <http://www.tronico.fi/OH6NT/docs/NMEA0183.pdf>
- [14] F. Ivis, Calculating Geographic Distance: Concepts and Methods, NESUG 2006, Data Manipulation, Canada, 2006. Online [Accessed 04 June 2021] <https://lexjansen.com/nesug/nesug06/dm/da15.pdf>
- [15] V. Fernandes, L. Mascarehnas, C. Mendonca, A. Johnson, R. Mishra, Speech emotion recognition using Mel Frequency Cepstral Coefficient and SVM classifier, 2018 International Conference on System Modeling & Advancement in Research Trends (SMART), Moradabad, India, pp. 200-204, 2018. DOI: [10.1109/SYSMART.2018.8746939](https://doi.org/10.1109/SYSMART.2018.8746939)
- [16] A. D. P. Ramirez, J. I. de la Rosa Vargas, R. R. Valdez, A. Becerra, A comparative between Mel frequency Cepstral Coefficients (MFCC) and Inverse Mel Frequency Cepstral Coefficients (IMFCC) features for an automatic bird species recognition system, 2018 IEEE Latin American Conference on Computational Intelligence (LA-CCI), 7-9 November 2018, Guadalajara, Mexico, pp. 1-4, 2018. DOI: [10.1109/LA-CCI.2018.8625230](https://doi.org/10.1109/LA-CCI.2018.8625230)
- [17] Mathworks, Matlab documentation. Online [Accessed 04 June 2021] <https://www.mathworks.com/help/wavelet/ref/cwtft.html>
- [18] M. F. Ghazali, Irregularity detection in artificial signal using time-frequency analysis, *Journal of Engineering and Applied Sciences* 11(2) (2016), pp. 3593-3597. Online [Accessed 04 June 2021] [http://www.arpnjournals.org/jeas/research\\_papers/rp\\_2016/jeas\\_0316\\_3850.pdf](http://www.arpnjournals.org/jeas/research_papers/rp_2016/jeas_0316_3850.pdf)
- [19] P. Bilski, B. Polok, Analysis of the RBF ANN-based classifier for the diagnostics of electronic circuit, *Acta IMEKO* 7(1) (2018), pp. 42-49. DOI: [10.21014/acta\\_imeko.v7i1.516](https://doi.org/10.21014/acta_imeko.v7i1.516)
- [20] D. L. Carni, E. Balestrieri, I. Tudosa, F. Lamonaca, Application of machine learning techniques and empirical mode decomposition for the classification of analog modulated signals, *Acta IMEKO* 9(2) (2020), pp 66-74. DOI: [10.21014/acta\\_imeko.v9i2.800](https://doi.org/10.21014/acta_imeko.v9i2.800)
- [21] T. Hastie, R. Tibshirani, J. Freedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Second edition, Springer-Verlag, New York, USA, 2009, ISBN: 9780387848570. DOI: [10.1007/978-0-387-84858-7](https://doi.org/10.1007/978-0-387-84858-7)
- [22] Delft University of Technology, 2014. Online [Accessed 04 June 2021] <http://prtools.org/>
- [23] F. Pavese, A. Charki, Data inter-comparisons in the context of the knowledge-gaining process: an overview, *Acta IMEKO* 7(2) (2018), pp. 73-83. DOI: [10.21014/acta\\_imeko.v7i2.541](https://doi.org/10.21014/acta_imeko.v7i2.541)
- [24] C. Beleites, U. Neugebauer, T. Bocklitz, C. Krafft, J. Popp, Sample size planning for classification models, *Anal. Chim. Acta*, 760 (2013), pp. 25-33. DOI: [10.1016/j.aca.2012.11.007](https://doi.org/10.1016/j.aca.2012.11.007)