Journal of Theoretical and Applied Information Technology
31st May 2021. Vol.99. No 10

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

ABNORMAL EVENT DETECTION IN INDOOR ENVIRONMENT BASED ON ACOUSTIC SIGNAL PROCESSING

[1]RUSTAM ABDRAKHMANOV, [1]ABDIMUKHAN TOLEP, [2]ZHAZIRA KOZHAMKULOVA, [3]NURLAN NARBEKOV, [4]NURBAI DOSSANOV, [1]BAKYTGUL YESKARAYEVA

[1]Khoja Akhmet Yassawi International Kazakh-Turkish University, Turkistan, Kazakhstan

[2]Almaty University of Power Engineering and Telecommunication, Almaty, Kazakhstan

[3]Caspian University of Technology and Engineering named after Sh.Yessenov, Aktau, Kazakhstan

[4]Turkestan Regional Branch of the Academy of Public Administration under the President of the Republic of Kazakhstan, Turkistan, Kazakhstan

[5]Tenaga National University, 43000 Kajang, Selangor, Malaysia

E-mail: rustam.abdrakhmanov@ayu.edu.kz, tolepabdimukhan@gmail.com

ABSTRACT

Alert the public about emergencies is to bring to public alerts and emergency information on dangers arising from the threat or occurrence of emergency situations of natural and technogenic character, as well as the conduct of hostilities or owing to these actions, the rules of behavior of the population and the need for protection activities. The aim of the work is to develop a method for detecting the sounds of critical situations in the sound stream. In this paper, the term "critical situation" is understood as an event, the characteristic sound signs of which can speak of acoustic artifacts (a shot, a scream, a glass strike, an explosion, a siren, etc.).

The developed method allows you to classify events into two groups: normal (for example, street noise) and critical situations (for example, an explosion, a scream, a shot). To determine events, machine learning is used, namely the Support Vector Machine method, which solves classification and regression problems by constructing a nonlinear plane separating the solutions. SVM has a fairly wide application in data classification and shows good results in event detection problems. As part of the work, the minimum set of features for the machine learning model was determined, small training and test samples were formed, and a method was developed that classifies normal and abnormal events.

Keywords: *Audio Event, Impulsive Sounds, Signal Processing, Machine Learning, Acoustic Signals, Signal Processing*

1. INTRODUCTION

Features of the current stage of development of society is a sharp aggravation of the criminal situation, especially in places of mass congestion of people, the growth of terrorist threats, various natural disasters and man-made disasters caused by violation of the ecological balance of the natural environment [1].

These factors have led to the need to create a new generation of security systems built on the basis of the latest achievements of information technologies using modern means of removing and registering information, powerful mathematical and software tools for data processing [2].

One of the promising ways to solve this problem is to create systems for recognizing emergencies based on the so-called audio analytics [3].

The advantages of audio analytics systems include the low cost of microphones compared to video cameras, the absence of "blind spots", and a lower density of coverage of the territory than for video surveillance [4-5]. The audio stream from the microphone takes up much less space than the video stream, which means it is easier to transport and process [6]. This means that a system built based on

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

audio signal processing will be operational, cheaper when it is created and then used.

At the same time, the audio signal contains enough information to determine with a high degree of reliability not only the fact of an emergency (or non-emergency) situation, but also to identify its type and speed of propagation [7]. The advantages of this approach have led to the creation of a number of systems built based on audio signal processing [8].

Since emergencies usually occur in conditions of extraordinary situations, then the management in such conditions of emergencies should have flexibility, the ability to work with a lack of information, quickly adapt to the high rate of change in the situation, quickly form effective solutions with high efficiency and, as a result, minimize the loss of time during the elimination of emergencies [9-10].

In the course of the work, the main methods of sound processing that improve its perception by both humans and automatic pulse sound recognition systems are studied, thereby detecting dangerous events from audio; the influence of sound preprocessing is studied in order to improve the quality of sound recognition; algorithms are implemented that allow recognizing pulse sounds in real mode

2. BACKGROUND

The practical value and relevance of studying the problems of decision-making in the management process is determined by several reasons.

First, the decision-making process is a management function from a methodological point of view. It is more general than other management functions [11]. In addition, any control function can be technologically represented as a sequence of decisions.

Secondly, decision-making is the main function for such categories as managers of various levels of organizations related to the prevention, localization and elimination of the consequences of emergency situations [12]. Therefore, knowledge of the methods, technology and tools of the decision-making process is a necessary element of the qualification of these categories.

Third, this approach is focused on decision-making and creating a solid basis for further improvement of information technologies of automated information support and management systems (decision support systems, expert systems) [13].

Automated control system – a system that ensures the effective functioning of the object, in which the collection and processing of information necessary for the implementation of management functions is carried out using automation and computer technology [14].

The main purpose of automation of organizational management is to ensure optimal management and optimal functioning of the control object, which means the correct choice of goals and means to achieve them (taking into account the environment and the situation in the system), as well as the best definition of tasks and the distribution of tasks between parts of the system, the implementation of interaction of all parts [15].

The rapid development of information technologies in recent years has led to the emergence of large amounts of information, communication, audio and video data that need to be understood, structured and analyzed to make effective management decisions. At the same time, along with the acceleration of the pace of development of information technologies, the time allowed for making optimal management decisions, especially decisions taken in emergencies, is reduced [16].

Decision-making technologies in crisis situations can be designated as organizational and information technologies that have a number of common characteristics [17-18]:

- aimed at increasing the amount of information on the problem under consideration, allow you to get specific information that is currently missing from the point of view of the person making the decision;
- generate alternative solutions that can be compared;
- allow you to work in crisis situations, becoming a kind of anti-crisis tools;
- combine the efforts of entire teams, creating a corresponding synergistic effect.

In emergencies, management bodies need not only traditional systems for collecting, processing and presenting information, but also analytical models that allow them to quickly assess the real state of the current situation, anticipate development trends and analyze the possible consequences of management decisions [19].

To date, all this complex of tasks can be solved by situational centers - crisis management centers, which are automated information and analytical complexes where important strategic decisions are made and from which all aspects of

Journal of Theoretical and Applied Information Technology
31st May 2021. Vol.99. No 10

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

the region's or country's activities are managed as a whole.

The main purpose of such centers is to ensure effective consolidation, purposeful use and development of the company's organizational capabilities based on the widespread use of the latest information and analytical methods and technologies both for the operational management of large geographical areas and for their organizational construction and development.

The situation center also makes it possible to model the scenarios of events, to think through the consequences of certain actions in advance, without waiting for the onset of a crisis situation [20].

3. MATERIALS AND METHODS

3.1 Audio-event detection systems

To ensure the safety of the world's population actively planned to implement the "safe city", which with the help of network cameras and video analysis features allow you to quickly recognize and react to different emergency situations and cases of violations of law [21].

Recently, given that an increasing number of video cameras are equipped with built-in microphones, such a direction of recognition of abnormal or emergency situations as audio analytics is actively developing.

One of the most well-known commercial developments of audio analytics is the American ShotSpotter system [22]. This system has been installed in disadvantaged areas of Washington since 2006 and over the years has localized 39,000 shots from firearms, quickly warning police about the onset of this event.

Another example of the implementation of such a system is the development of Audio Analytical Ltd from the UK [23]. Sensors of this company, designed to create a "smart home", are able to register events such as gunshots, aggressive screams, crying baby, car alarm sounds, broken glass, etc. After registering the event, the system sends notifications to the user, security agencies for subsequent response.

Also known is the development of JSC "Innovative Technology Center "Sistema-Sarov" [23], which uses special acoustic microphones in certain points of the settlement. The system integrates with various video surveillance tools, allowing the operator to direct available video cameras to the area of the sound anomaly to identify the situation. Comparative characteristics of the functional capabilities of these systems are shown in the table.

Knowing the location of microphones and using triangulation methods [24], these systems accurately determine the location of the event. The undoubted benefits systems audibility include low, compared to camera cost of the microphones, no blind spots, low density coverage than CCTV. The audio stream from the microphone takes up much less space than the video stream, which means it is easier to transport and process.

An important point in the work of audio analytics systems is the means of recording signals, their geographical distribution (coverage of a wide area, since early notification is needed to reserve time), the availability of stable data transmission channels, ease of communication and interaction with users of the system.

Currently, the ideal tools that can perform these tasks are smartphones connected to the data transmission channels formed by the cellular networks of mobile operators. The main advantage of this approach is that smartphones are increasingly popular in Ukraine (as well as in the world as a whole). So, according to the latest data, in Ukraine more than 30% of the population aged 18 to 50 years use smartphones. Each smartphone has a microphone, an information exchange channel, a location sensor and can be equipped with various software [25].

All this makes the smartphone a potential candidate for both the role of a means of personal notification of the owner about the occurrence of an emergency situation, and a means of early detection and identification of an emergency situation. That is, in the presence of special software and a working data transmission channel, which is the mobile Internet, the smartphone can pick up sounds that identify an emergency situation, transmit the characteristics of these sounds, data about its location and exact time to a remotely located analytical system.

Detected emergencies may include situations related to a terrorist threat, a violation of public order, various man-made accidents that are accompanied by loud explosions, sirens and other acoustic artifacts.

The most widely used audio file analysis solutions in the world are shown in Table 1.

Journal of Theoretical and Applied Information Technology
31st May 2021. Vol.99. No 10

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 1: Systems of Audio based event recognition and analysis

Recognition systems	ShotSpotter	Sistema-Sarov	Audio Analytic
Year	1996	2010	2008
Country	USA	Russia	The UK
Application	when detecting shots from a firearm in an urban setting	in the implementation of natural, man-made, criminal, terrorist and other threats	when detecting dog barking, screaming/crying, breaking glass, smoke alarm
Advantages	High accuracy in detecting gunshot sounds	Wide range of recognizable audio situations	A housing-oriented system, with a relatively low cost
Disadvantages	The system does not provide recognition of other types of sounds	Huge data stream resource intensive to process	Small range of signals for recognition due to the scope of application

The stage of recognition of audio events is the selection of the training model. Currently, the following model classifiers are used: Bayesian Classifier, GM (Gaussian Mixture Model), SVM (supportvectormachine), HMM (hiddenmarkov Model), Neural networks. Their comparative analysis is presented in Table 2.

Table 2: Comparative analysis of training models for recognizing audio events

Classifier	A Bayesian classifier	GMM (GaussianMixture Model)	SVM (SupportVectorMachine)	HMM (Hidden Markov Model)	Neural networks
The principle	Maximum a posteriori probability	probability distribution of observations in the general population	translates the original feature vectors into a space with a larger dimension and finds the maximum separating hyperplane from the recognized classes	a sequence of random variables, where each value xt+1 depends only on the	built on a multi-layer perceptron, which allows you to classify the input signal

				previous th and, under the condition of th, is conditionally independent with the previous th-k)	in accordance with the pre-configuration of the network
Function	the probability density function of	a "mixture" of several Gaussians	linear function, polynomial, and RBF	Statistical model simulating a Markov process with unknown parameters	Mathematical model based on the principle of organization of biological neural networks
Advantages	to train and then to interrogate large data sets; the relative ease of interpretation	the recognition model has become more accurate, and the recognition result will improve accordingly	finds the maximum width of the strip, which provides greater accuracy	The model is quite simple and flexible for integration into many systems	The neural network itself detects uninformative noises and filters them out
Disadvantages	relatively poor classification quality in most real-world problems	high sensitivity to variations in the training data sample when selecting a large number of Gaussian distributions	sensitivity to the standardization of data and noise	High probability of erroneous inference, when several hidden states are connected to one observed result	problems of mathematical nature;

3.2 Applying machine learning in audio-event detection

Technical (software and hardware) aspect-related to the formulation of proposals (recommendations) for the practical implementation of algorithms for signal processing and decision-making.

Machine learning is gradually changing the face of our world. We no longer need to teach computers how to perform complex tasks such as image recognition or text translation: instead, we create special systems that allow them to learn how to do it on their own [26].

The most powerful form of machine learning currently in use is called deep learning. On its basis, you can create a complex mathematical structure based on large amounts of data called a neural network.

For basic analysis, a deep neural network consisting of only two convolutional layers and two fully connected layers can be suitable. However, such a network is unable to provide data that is relevant enough to create a full-fledged system of music recommendations, since a lot of parameters will be overlooked [27].

Audio classification is a fundamental problem in the field of audio processing. This is mainly due to the fact that the results of deep learning are highly dependent on the source data, and it is quite difficult to present audio in an adequate form for machine perception, even if it is represented by a high-quality recording in lossless format [28].

It is not enough to use a simple representation of the audio wave equation - displaying the dependence of the amplitude on time.

Using the fast Fourier transform, we get a periodogram that shows the value of the signal that falls on each of the frequencies. In the case of musical passages, the signal strength is most often at low frequencies, despite the fact that the range from 1 Hz to the Nyquist frequency is initially taken, in this case 22 thousand Hz. The answer is simple - the human ear is most sensitive to changes in low frequencies, so it becomes obvious that these frequencies are used most often.

Next comes the window Fourier transform, where we use the Hamming window as a window function, inside which the fast Fourier transform will be performed [29].

4. RELATED WORKS

Audio classification is a fundamental problem in the field of audio processing. The task is to extract some markers from the audio and then determine which class the audio belongs to. Many useful applications related to audio classification have already been developed and perform their functions - such as genre classification, instrument recognition, and artist identification [30].

In the age of digital information, solutions for search and recognition in audio and video materials are an integral part of security systems and human-machine interfaces. There are many ways to analyze information, including using neural networks. Audio and video analytics are used by security agencies and security services to help those in trouble or prevent terrorist attacks [31].

Human activity is inextricably linked with the technogenic sphere. The quality of life of people is ensured, among other things, by safety in emergency situations of a man-made nature. It is important to improve the safety of people through the development and implementation of modern methodologies for protecting the population from various types of threats, as well as from fires [32].

Currently, there is an increase in the number of man-made emergencies. All this indicates that the modern development of society requires improvement of the existing risk management system in order to minimize the impact of adverse factors of the technogenic environment [33].

Only the creation and implementation of a new ideology in the process of emergency prevention, based on the principles of quality management, can reduce the risk of emergencies [34].

Currently, the ideal tools that can perform these tasks are smartphones connected to data channels formed by mobile operators' cellular networks. The main advantage of this approach is that smartphones are increasingly popular in Ukraine (as well as in the world). According to the latest data, more than 30 % of the population aged 18 to 50 years use smartphones. Each smartphone has a microphone, an information exchange channel, a location sensor, and can be equipped with various software [35-36].

This makes the smartphone a potential candidate as a means of personal notification of the owner about the occurrence of an emergency, as well as a means of early detection and identification of an emergency. That is, if the smartphone were equipped with special software and a working data transmission channel, which is the mobile Internet, it would be able to detect sounds that identify an emergency, transmit the characteristics of these sounds, data about its location and exact time to a remotely located analytical system [37].

Detected emergencies may include situations involving a terrorist threat, public order violations, or various man-made accidents that are accompanied by loud explosions, sirens, and other acoustic artifacts.

There are some research gaps in detecting impulsive sounds:

1. The absence of a database of impulse sounds. To address this research gap we developed a dataset of impulsive sounds. The dataset contains 10 000 sounds, split into eight categories: gunshot, broken glass, fire, siren, explosion, cry, dog barking, fire alarm bell. Sounds of the given categories allow to train machine learning models and detect impulsive sounds immediately with high precision.

2. Noises in input data. In order to get high precision in detection of impulsive sounds, the input data must not contain noises. We removed all the noises from our dataset to improve the quality of the detection process.

3. Machine Learning Model. Proposed machine learning model is essential to get high precision in detection.

5. EXPERIMENTS AND RESULTS

Figure 1 shows an acoustic monitoring system that allows for automatic and real-time detection of pulsed sounds and events. The system receives a sound and receives the characteristic of the sound event. Classifies audio signals using machine learning to detect pulsed sounds. Defines keywords in the event as "police", "ambulance", "bomb". Determining the location of dangerous events. The next step is to transmit information about the alarm event to the authorized bodies, store the event archive and add the database.

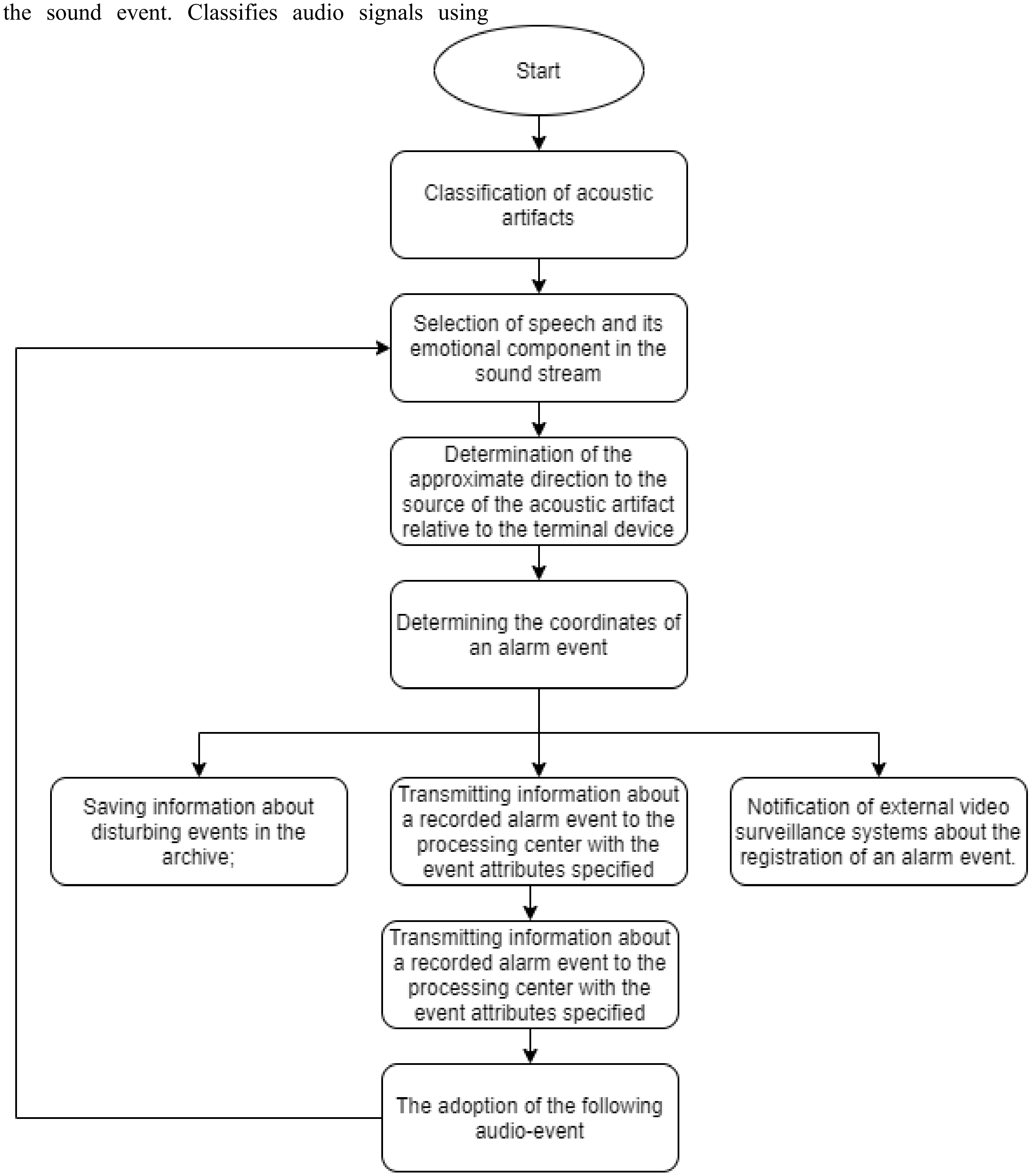


Figure 1: Classification of acoustic signals

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

5.1 Mathematical aspect

The solution of the above tasks is carried out in two stages:

- based on information received in real mode from acoustic signal sensors $S_1^j, S_2^j, ..., S_{mj}^j$, $(j = 1,2,...n)$, where mj – number of sensors on the j-th control object; n – the number of control objects., determine the intensity, position (coordinates) and nature (type) of the source of the dangerous signal (threat).

Features of setting and solving this problem include:

- consideration of directional patterns and other characteristics of the sensors;
- using the triangulation method to detect the coordinates of the location of the audio signal source;
- estimation of the error in determining the coordinates of the threat source;
- analysis (transformation) of the received signals in order to identify their characteristic features (attributes) of the threat source;
- assigning the threat source to one of the standard classes;
- making a decision in the centralized control about the necessary actions when the source of the threat is detected.

5.2 Technical aspect

Technical (software and hardware) aspect-related to the formulation of proposals (recommendations) for the practical implementation of algorithms for signal processing and decision-making.

At this stage, issues such as:

- selection of specific types of audio signal sensors, taking into account the characteristics of control objects;
- selection of pre-processing methods (conversion and compaction) and transmission of signals from the sensor to the PCC;
- selection of methods and software for processing the received signals in the PCC based on the proposed algorithms;
- solution of issues of integration of the proposed approach with the existing system of monitoring of geographically distributed objects based on video surveillance;
- feasibility study of the effectiveness of the proposed solutions.

5.3 Analysis of audio-signals

The solution of this problem involves:

- selection of directional microphones, analysis of their compatibility with the technical characteristics of outdoor surveillance systems;
- selection of the method of placing microphones on the object;
- use of standard (existing) interfaces and channels for transmitting information to the centralized control system;
- use of various types of transformations (BFP, wavelet) to analyze signals coming from sensors in real time.

Studies show that effective recognition of acoustic signals is possible with the help of neural network technologies based on the use of wavelet transforms [38]. It is known that the traditional device for representing signals in the form of Fourier series is ineffective for signals with local features, in particular for pulse signals. This is because the basis function of the Fourier series – a sinusoid – is defined in time and space from -∞ to +∞ and is a strictly periodic function. Such a function is fundamentally unable to describe arbitrary signals observed in real life [39].

If the process under study is characterized by non-stationarity, which is typical for real audio signals with a non-constant noise component, the use of spectral analysis does not give effective results, especially if the non-stationarity in time is significantly less than the analyzed time interval. This circumstance is clearly evident for the problem under study [40].

To illustrate the above, consider two signals-information corresponding to a single shot (Figure 2, a) and noise (Figure 2, b), similar to the noise component of the signal in Figure 2,a. The obtained Fourier spectra (Figure 2, c, d) of these two signals are very similar. In the areas from 0 to 500 Hz, peaks related only to the noise component of the signal are detected, peaks with a frequency of less than 100 Hz are visible. The information component of the signals can be determined at frequencies (900-2500 Hz), where the difference between both files is maximum. At the same time, the analysis of the spectra obtained using the Fourier transform does not reveal an important feature of the original signal – the moment in time when the composition of the original signal changes.

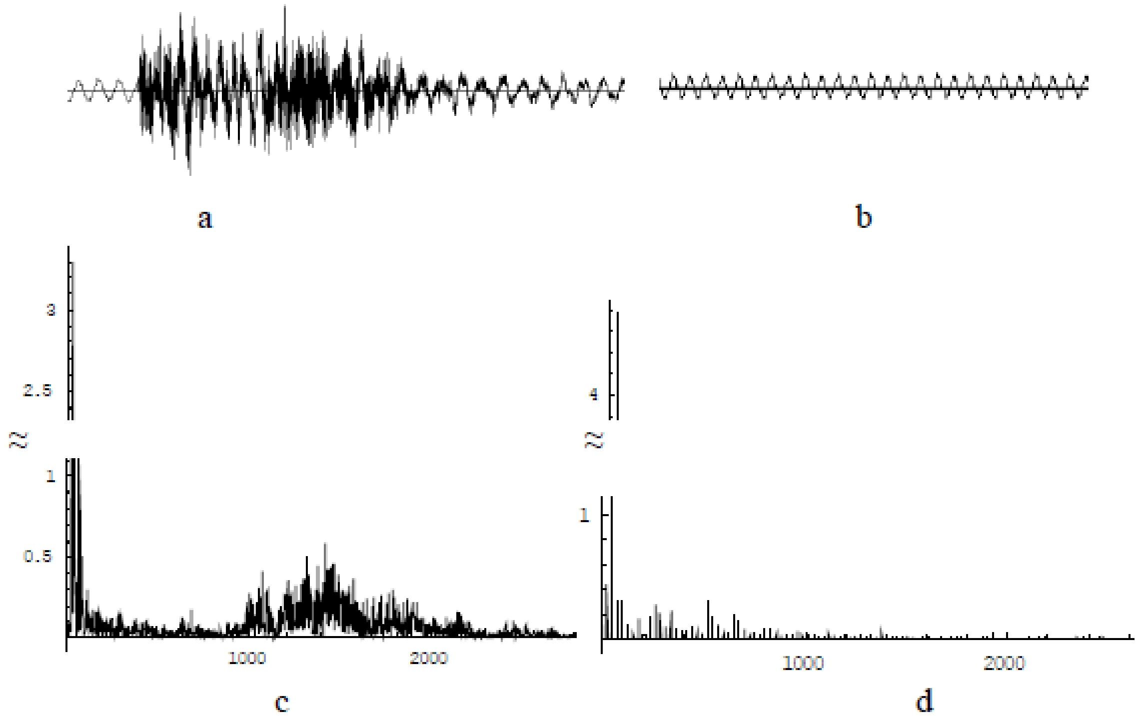


Figure 2: Signals and their Fourier spectra

5.4 Experiment results

The performance of the proposed system was evaluated for an automated surveillance application that should be able to recognize the following events (considered " abnormal" in the observed environment): shots, screams, broken windows.

For this purpose, we built a data set from several audio samples recorded in various railway station scenarios.

The data set consists of background noise signals: pick, shot, and broken glass. Background noise was obtained indoors and outdoors to account for the characteristics of various application scenarios.

For our experiments, the signals were divided into intervals of one second (the average time length of each event of interest), and then each interval was divided into frames of 200 MS, overlapping by 50%: in particular, each interval consists of nine frames.

Second stage interval classifier: this stage performs classification at the interval level, assigning each interval a final forecast based on a weighted majority voting strategy (WMV).

As a measure of reliability for each Classifier, we chose measure F, considering it as a good compromise between accuracy and recall:

$$recision = \frac{tp}{tp + fp} \quad (1)$$

Here tp is true positive classified samples; fp is false negative classified samples.

$$Recall = \frac{tp}{tp + fn} \quad (2)$$

$$F_{measure} = \frac{2 * Precision * Recall}{Precision + Recall} \quad (3)$$

Finally, as for the deviation threshold, it was determined at the setup stage using ROC analysis aimed at maximizing the accuracy of predictions.

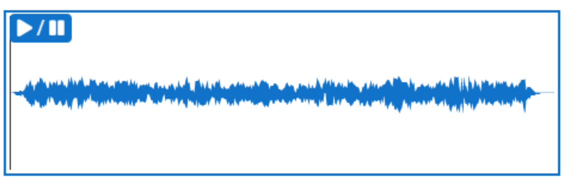
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 1: Results of impulsive audio event detection.

Event type	***Accuracy***	***Precision***	***Recall***	***F1 score***	***AUC ROC***
Gunshot	0.9178	0.9245	0.9427	0.8945	0.9748
Broken glass	0.9372	0.9765	0.9215	0.9154	0.9578
Fire	0.9435	0.9346	0.9215	0.9345	0.9576
Siren	0.9537	0.9462	0.9876	0.9642	0.9623
Explosion	0.8132	0.8254	0.8352	0.8124	0.9348
Cry	0.8635	0.8524	0.8864	0.8754	0.9467
Dog barking	0.8456	0.8325	0.8571	0.8254	0.9425
Fire alarm bell	0.8654	0.8452	0.8576	0.8457	0.9472

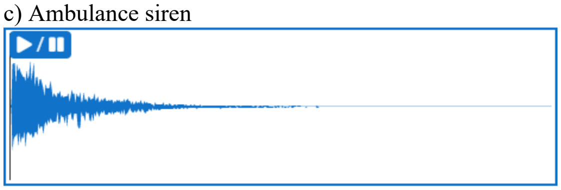
Figure 3 illustrates some samples of audio-signals that detected by the system.

a) Automobile glass shattering

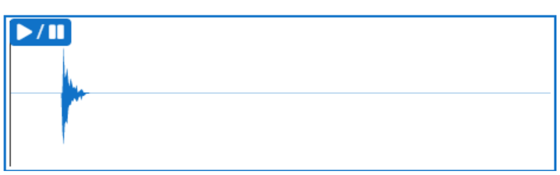


b) Police siren

c) Ambulance siren



d) Explosion



e) Single gun shot

Figure 3: Impulsive audio-signal

4. DISCUSSION

As a result of the work, an approach to solving the main research tasks was chosen: the stages of obtaining audio signal characteristics were determined, the Essentia library was chosen as a tool for extracting characteristics and the Python programming language for developing system components. Further solutions require two global tasks. The first is the detection of sharp pulse signals from background noise in the audio data stream. To do this, you will need to determine the characteristics of the signal that will allow them to be distinguished. And the second is the classification of the detected signal to one of the types of audio events using machine learning algorithms.

When comparing, priority was given to the cross-platform nature of the tool, so that it was a plug-in library, and additional features were also noted. As a result, the Essentia library was chosen to analyze the audio signal. It meets all the selected criteria, supports the use of the Python language, which has a developed infrastructure for solving various research tasks (statistics, machine learning, visualization, etc.), has a Vamp plugin that can be connected to other software products, for example, to Sonic Visualiser, and, unlike many other analyzed tools, is still supported by developers.

Essentia is an open source library in C++. It contains an extensive collection of algorithms, including algorithms that implement audio file input/output functionality, standard digital signal processing algorithms, algorithms for obtaining statistical descriptions of data, and a large set of spectral, temporal, tonal, and high-level music descriptors. Essentia was created as a library of signal processing units. Each processing block is called an algorithm and has three different types of attributes: inputs, outputs, and parameters.

Now, the field of application of neural networks is very diverse. In many areas, quite high results have been achieved, for example, in pattern recognition, contour selection, and so on. The use of neural networks in speech recognition began to be used not so long ago, because until recently, neural networks did not give the desired percentage of accuracy.

The first thing you need to do when working with sound is to represent the sound as an analog signal in digital form for processing it on the computer.

Sound is a wave propagating through space. A sound wave is a signal that has a dimension of one. At each moment of time, the value of the sound signal is the amplitude of the sound wave.

Modeling of various types of sounds and their spatio-temporal effects becomes important for assessing the sound environment both in work areas and in recreation areas, bus stations, airports, and in general in the city when the term smart city has a high relevance. The development of a model that reflects the characteristics of sounds, their sources and rules governing their distribution in different environments will help to track changes in sounds and predict their future changes for spatio-temporal states and determine the location and source of the event. There are similar works abroad, but they are of a private nature. There are many functions that can be used to describe audio signals. We look at a wide range of objects to evaluate the effect of each object and select the appropriate set of objects to distinguish between classes.

The research of methods and approaches for audio analysis is carried out. The forms of representation of the audio signal for its transformation before analysis are considered. The possibility of distinguishing features from the audio signal in the time and frequency domains is analyzed. Examples of frequently used and most informative features are given. Methods of calculation are described, their appearance in the form of graphs is presented. Three different classification methods are investigated: the kneading model, the Gaussian mixture model, and the support vector machine.

4. CONCLUSION

Measures aimed at increasing the financial support of specialists-analysts will contribute to reducing the occurrence of emergency situations. A differentiated approach to the payment system, aimed at increasing the monetary content, will help reduce the outflow of these specialists to other areas of economic activity.

Thus, the creation and integration of quality systems for predicting and preventing man-made emergencies will help reduce the risk of these emergencies. At the same time, the insignificant financial costs necessary for the creation of the system will be compensated by the economic benefits of industrial enterprises, where there may be abnormal situations that can lead to irreparable human and technological losses.

The system allows you to register signals about a potential threat, identify the nature of the threat, the dynamics and direction of its spread, and notify the user of the system with the provision of recommendations and rules of conduct in the event of a specific emergency. The system consists of several subsystems and is built on a modular principle. The goals and tasks performed by individual subsystems are described, and the logic of their work is described. All this will make it possible to warn the population, authorities, enterprises, organizations, institutions and educational institutions in advance about the occurrence of emergencies, therefore, to respond adequately to the prevailing conditions. In the end – it is necessary to reduce losses in people and material values to the maximum extent.

REFRENCES:

[1] Abou Elassad, Z. E., Mousannif, H., Al Moatassime, H., & Karkouch, A. (2020). The application of machine learning techniques for driving behavior analysis: A conceptual framework and a systematic literature review. Engineering Applications of Artificial Intelligence, 87, 103312.

[2] Kauppinen, I. (2002, July). Methods for detecting impulsive noise in speech and audio signals. In 2002 14th International Conference on Digital Signal Processing Proceedings. DSP 2002 (Cat. No. 02TH8628) (Vol. 2, pp. 967-970). IEEE.

[3] Esquef, P. A., Biscainho, L. W., Diniz, P. S., & Freelanci, F. P. (2000, September). A double-threshold-based approach to impulsive noise detection in audio signals. In 2000 10th European Signal Processing Conference (pp. 1-4). IEEE.

[4] Chandra, C., Moore, M. S., & Mitra, S. K. (1998, May). An efficient method for the removal of impulse noise from speech and audio signals. In 1998 IEEE International Symposium on

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Circuits and Systems (ISCAS) (Vol. 4, pp. 206-208). IEEE.

[5] Avila, F. R., & Biscainho, L. W. (2012). Bayesian restoration of audio signals degraded by impulsive noise modeled as individual pulses. IEEE transactions on audio, speech, and language processing, 20(9), 2470-2481.

[6] Esquef, P., Karjalainen, M., & Valimaki, V. (2002, July). Detection of clicks in audio signals using warped linear prediction. In 2002 14th International Conference on Digital Signal Processing Proceedings. DSP 2002 (Cat. No. 02TH8628) (Vol. 2, pp. 1085-1088). IEEE.

[7] Omarov, B., ANARBAYEV, A., TURYSKULOV, U., ORAZBAYEV, E., ERDENOV, M., IBRAYEV, A., & KENDZHAEVA, B. (2020). Fuzzy-PID based self-adjusted indoor temperature control for ensuring thermal comfort in sport complexes. J. Theor. Appl. Inf. Technol, 98(11).

[8] Castillo-Escario, Y., Ferrer-Lluis, I., Montserrat, J. M., & Jané, R. (2019, July). Automatic silence events detector from smartphone audio signals: A pilot mHealth system for sleep apnea monitoring at home. In 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 4982-4985). IEEE.

[9] Sadrizadeh, S., Zarmehi, N., Asadi, E., Abin, H., & Marvasti, F. (2020). A Fast Iterative Method for Removing Impulsive Noise from Sparse Signals. IEEE Transactions on Circuits and Systems for Video Technology.

[10] Landa, I., Díez-Bravo, A., Velez, M., & Arrinda, A. (2019). Influence of impulsive noise from lifts on orthogonal frequency-division multiplexing digital radio mondiale signals. IET Communications, 14(1), 111-116.

[11] Buhuş, E. R., Grama, L., & Rusu, C. (2017, July). Several classifiers for intruder detection applications. In 2017 International Conference on Speech Technology and Human-Computer Dialogue (SpeD) (pp. 1-6). IEEE.

[12] ARSLAN, Y. (2018). Detection and Recognition of Sounds from Hazardous Events for Surveillance Applications (Doctoral dissertation, Ankara Yıldırım Beyazıt Üniversitesi Fen Bilimleri Enstitüsü).

[13] PIMPLEKAR, D. M., & PATANGE, A. (2017). Removal of Impulse Noise from Speech Signal by Preserving the Original Speech Signal Quality using Wavelets Transformation Approach.

[14] Grama, L., Buhuş, E. R., & Rusu, C. (2017, July). Acoustic classification using linear predictive coding for wildlife detection systems. In 2017 International symposium on signals, circuits and systems (ISSCS) (pp. 1-4). IEEE.

[15] Wan, Y., Chen, Y., Sim, K. Y. H., Wu, L., Geng, X., & Chau, K. (2019, October). Precise Temporal Localization of Sudden Onsets in Audio Signals Using the Wavelet Approach. In Audio Engineering Society Convention 147. Audio Engineering Society.

[16] Omarov, B., Baisholanova, K., Abdrakhmanov, R., Alibekova, Z., Dairabayev, M., Narykbay, R., & Omarov, B. (2017). Indoor microclimate comfort level control in residential buildings. Far East Journal of Electronics and Communications, 17(6), 1345-1352.

[17] Nguyen, T. N. T., Jones, D. L., & Gan, W. S. (2020, May). A Sequence Matching Network for Polyphonic Sound Event Localization and Detection. In ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 71-75). IEEE.

[18] Bevinamarad, P. R., & Shirldonkar, M. S. (2020, June). Audio Forgery Detection Techniques: Present and Past Review. In 2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184) (pp. 613-618). IEEE.

[19] Ivry, A., Cohen, I., & Berdugo, B. (2020, May). Evaluation of Deep-Learning-Based Voice Activity Detectors and Room Impulse Response Models in Reverberant Environments. In ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 406-410). IEEE.

[20] Omarov, B., Altayeva, A., Turganbayeva, A., Abdulkarimova, G., Gusmanova, F., Sarbasova, A., ... & Omarov, N. (2018, November). Agent based modeling of smart grids in smart cities. In International Conference on Electronic Governance and Open Society: Challenges in Eurasia (pp. 3-13). Springer, Cham.

[21] Omarov, B., ANARBAYEV, A., TURYSKULOV, U., ORAZBAYEV, E., ERDENOV, M., IBRAYEV, A., & KENDZHAEVA, B. (2020). Fuzzy-PID based self-adjusted indoor temperature control for ensuring thermal comfort in sport complexes. J. Theor. Appl. Inf. Technol, 98(11).

[22] Alam, J., & Kenny, P. (2017, August). Spoofing detection employing infinite impulse response—constant Q transform-based feature

representations. In 2017 25th European Signal Processing Conference (EUSIPCO) (pp. 101-105). IEEE.

[23] Abayomi-Alli, A., Abayomi-Alli, O., Vipperman, J., Odusami, M., & Misra, S. (2019, July). Multi-class Classification of Impulse and Non-impulse Sounds Using Deep Convolutional Neural Network (DCNN). In International Conference on Computational Science and Its Applications (pp. 359-371). Springer, Cham.

[24] Arslan, Y., & Canbolat, H. (2018, October). Performance of deep neural networks in audio surveillance. In 2018 6th International Conference on Control Engineering & Information Technology (CEIT) (pp. 1-5). IEEE.

[25] Rossi, A., Montefoschi, F., Rizzo, A., Diligenti, M., & Festucci, C. (2017, October). Auto-associative recurrent neural networks and long term dependencies in novelty detection for audio surveillance applications. In IOP Conference Series: Materials Science and Engineering.

[26] Oltean, G., Grama, L., Ivanciu, L., & Rusu, C. (2015, October). Alarming events detection based on audio signals recognition. In 2015 International Conference on Speech Technology and Human-Computer Dialogue (SpeD) (pp. 1-8). IEEE.

[27] Łopatka, K., Zwan, P., & Czyżewski, A. (2010). Dangerous sound event recognition using support vector machine classifiers. In Advances in Multimedia and Network Information System Technologies (pp. 49-57). Springer, Berlin, Heidelberg.

[28] Segal, G., Segev, A., Brom, A., Lifshitz, Y., Wasserstrum, Y., & Zimlichman, E. (2019). Reducing drug prescription errors and adverse drug events by application of a probabilistic, machine-learning based clinical decision support system in an inpatient setting. Journal of the American Medical Informatics Association, 26(12), 1560-1565.

[29] Omarov, B., Omarov, B., Issayev, A., Anarbayev, A., Akhmetov, B., Yessirkepov, Z., & Sabdenbekov, Y. (2020, November). Ensuring Comfort Microclimate for Sportsmen in Sport Halls: Comfort Temperature Case Study. In International Conference on Computational Collective Intelligence (pp. 626-637). Springer, Cham.

[30] Nousias, Stavros, Aris S. Lalos, Gerasimos Arvanitis, Konstantinos Moustakas, Triantafillos Tsirelis, Dimitrios Kikidis, Konstantinos Votis, and Dimitrios Tzovaras. "An mHealth system for monitoring medication adherence in obstructive respiratory diseases using content based audio classification." IEEE Access 6 (2018): 11871-11882.

[31] Abou Elassad, Z. E., Mousannif, H., Al Moatassime, H., & Karkouch, A. (2020). The application of machine learning techniques for driving behavior analysis: A conceptual framework and a systematic literature review. Engineering Applications of Artificial Intelligence, 87, 103312.

[32] Vozarikova, E., Lojka, M., Juhar, J., & Cizmar, A. (2012, May). Performance of basic spectral descriptors and mrmr algorithm to the detection of acoustic events. In International Conference on Multimedia Communications, Services and Security (pp. 350-359). Springer, Berlin, Heidelberg.

[33] Rehman, Z. U., Khan, S. N., Muhammad, K., Lee, J. W., Lv, Z., Baik, S. W., ... & Mehmood, I. (2018). Machine learning-assisted signature and heuristic-based detection of malwares in Android devices. Computers & Electrical Engineering, 69, 828-841.

[34] Lopatka, K., Kotus, J., & Czyzewski, A. (2016). Detection, classification and localization of acoustic events in the presence of background noise for acoustic surveillance of hazardous situations. Multimedia Tools and Applications, 75(17), 10407-10439.

[35] Łopatka, K., Kotus, J., & Czyżewski, A. (2014, June). Evaluation of sound event detection, classification and localization in the presence of background noise for acoustic surveillance of hazardous situations. In International Conference on Multimedia Communications, Services and Security (pp. 96-110). Springer, Cham.

[36] Syafrudin, M., Alfian, G., Fitriyani, N. L., & Rhee, J. (2018). Performance analysis of IoT-based sensor, big data processing, and machine learning model for real-time monitoring system in automotive manufacturing. Sensors, 18(9), 2946.

[37] Anemüller, J., Bach, J. H., Caputo, B., Havlena, M., Jie, L., Kayser, H., ... & Hermansky, H. (2008, October). The DIRAC AWEAR audio-visual platform for detection of unexpected and incongruent events. In Proceedings of the 10th international conference on Multimodal interfaces (pp. 289-292).

[38] Gulmezoglu, B., Eisenbarth, T., & Sunar, B. (2017, April). Cache-based application

detection in the cloud using machine learning. In Proceedings of the 2017 ACM on Asia Conference on Computer and Communications Security (pp. 288-300).

[39] Narynov, S., Mukhtarkhanuly, D., & Omarov, B. (2020). Dataset of depressive posts in Russian language collected from social media. Data in brief, 29, 105195.

[40] Omarov, B., Suliman, A., & Tsoy, A. (2016). Parallel backpropagation neural network training for face recognition. Far East Journal of Electronics and Communications, 16(4), 801-808.

[41] El Attaoui, A., Largo, S., Kaissari, S., Benba, A., Jilbab, A., & Bourouhou, A. (2020). Machine learning-based edge-computing on a multi-level architecture of WSN and IoT for real-time fall detection. IET Wireless Sensor Systems, 10(6), 320-332.

[42] Sazonov, E. S., Makeyev, O., Schuckers, S., Lopez-Meyer, P., Melanson, E. L., & Neuman, M. R. (2009). Automatic detection of swallowing events by acoustical means for applications of monitoring of ingestive behavior. IEEE Transactions on Biomedical Engineering, 57(3), 626-633.

[43] Bobade, P., & Vani, M. (2020, July). Stress Detection with Machine Learning and Deep Learning using Multimodal Physiological Data. In 2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA) (pp. 51-57). IEEE.

[44] Raschke, Robert A., Bea Gollihare, Thomas A. Wunderlich, James R. Guidry, Alan I. Leibowitz, John C. Peirce, Lee Lemelson, Mark A. Heisler, and Cynthia Susong. "A computer alert system to prevent injury from adverse drug events: development and evaluation in a community teaching hospital." Jama 280, no. 15 (1998): 1317-1320.

[45] Blackwell, A. F. (2015, August). Interacting with an inferred world: the challenge of machine learning for humane computer interaction. In Proceedings of The Fifth Decennial Aarhus Conference on Critical Alternatives (pp. 169-180).

[46] Kendler, K. S., Hettema, J. M., Butera, F., Gardner, C. O., & Prescott, C. A. (2003). Life event dimensions of loss, humiliation, entrapment, and danger in the prediction of onsets of major depression and generalized anxiety. Archives of general psychiatry, 60(8), 789-796.