

ABSTRACT

for dissertation work on the topic

“RESEARCH OF AUDIO-BASED EMERGENCY DETECTION AND CLASSIFICATION”

submitted for the degree of Doctor of Philosophy (PhD)

in specialty "6D071900 - Radio engineering, electronics and telecommunications"

DOSBAYEV ZHANDOS MAKHSUTULY

The relevance of the work. Ensuring the safety of citizens is an important task for the state. However, throughout the world and in our country, there is an acute problem of finding effective approaches and solutions to prevent social emergencies such as domestic violence, theft, armed attacks, terrorism, etc.

Statistics of violence, especially related to children and women, are growing every year. It became known that during the covid-19 pandemic, the number of such cases increased. According to the UN, about 300 million children aged 2 to 4 years are subjected to violence. In 2017 alone, 40,150 children under the age of 18 died for this reason. According to UNICEF Kazakhstan, 44% of students were subjected to violence at school and 24% of students committed violent acts with other children. It is also reported that 150 million teenagers aged 13 to 15 years were subjected to violence on school territory.

As a result of a survey conducted by public activists in 13 countries, it was found that every second woman herself or a closer people became a victim of domestic violence. According to the information provided by the National Statistical Bureau of the Agency for Reforms and Strategic Management of the Republic of Kazakhstan, 63,447 cases in 2020, and 61,464 cases of domestic violence were registered against women in 2021.

According to the statistics of the Ministry of Internal Affairs of the Republic of Kazakhstan, 163,226 cases in 2020, 157,884 cases in 2021, and 140,592 criminal violations were registered in 10 months of 2022: 53,386 cases of thefts, 417 cases of violent crimes, 2,044 cases of hooliganism, etc. In the first half of 2021, 139 offenses related to extremism and terrorism were detected in the country. Also, 250,000 people die every year in the world because of weapons.

The results of the above statistics of social emergencies, such as weapons, violence among adolescents and against women, theft, damage to citizens' property, etc., indicates the presence of urgent problems and indicates the need for effective solutions to ensure the safety of citizens.

Nowadays, automated technologies are becoming an integral part of a person's daily life. Another area of use of AT is the identification of emergency events in public places and ensuring the safety of citizens while remaining vigilant. The study of autonomous surveillance systems is primarily related to computer vision for detecting events. The

combination of acoustic monitoring systems with video surveillance systems (CCTV) will increase the efficiency of activity recognition and provide an additional source of information.

Experts confirm that verbal aggression such as shouting, rude words, etc precedes the 90% of cases of physical aggression. Accordingly, valuable information and data about an emergency event can be obtained through audio data. Automated audio analysis systems operating in real time are used to detect, prevent and investigate emergencies.

Currently, video surveillance systems are widely used to ensure security. However, video surveillance systems cannot provide complete security due to some inherent disadvantages. For example, video surveillance systems cannot provide full surveillance at night in areas not equipped with lighting systems. The price of audio systems service, devices and components is lower than video surveillance systems. In addition, since security systems operate in real time, the volume of audio data is less than video recordings. Accordingly, the technical requirements of bandwidth will reduce. Acoustic sensors or microphones and even smartphones can be used as a recording device. These advantages of audio based systems indicate that such systems are in demand in the field of security and determine the relevance of studying automated emergency detection systems based on audio data in real time.

The number of studies conducted on the topic of the dissertation work is increasing every year. The detection and classification of emergency situations using deep learning methods based on audio signals attracts interest of researchers. This is due to the fact that the current period is the basis of many research works due to the completeness of the three technical driving forces of machine learning: high-performance equipment; due to the development of the Internet, a sufficient amount of data and the availability of testing; algorithmic achievements. For example, various seminars on audio processing are organized, and participants are given the opportunity to present their solutions and test them. One of the most famous of such seminars is the DCASE Challenge. The event is held annually to review various types of audio-related tasks of classification and detection, by updating tasks and datasets. Different type of tasks are considered such as sound events and their localization, sound detection, sound analysis of the environment, classification of acoustic events, as well as the detection of rare sound events. However, since the seminars are not aimed at identifying only emergency situations, audio data sets have not been created for a broader description of emergency events. And the results of the work of other researchers are limited to only a small number of classes. Therefore, systems for detecting several classes of social emergency situations based on audio signal processing need to be investigated.

The subject of the research. Methods of audio signal analysis and audio data preprocessing, neural network architectures and algorithms for emergency detection.

The object of the research. Audiosignals.

The purpose of the work is to create a model for detecting and classifying emergencies with high accuracy based on audio signal pulses, working in real time.

Research tasks. the following tasks were set to achieve the goal of the research work:

- Conducting a comparative analysis of audio data processing methods and emergency detection models based on audio analysis;
- Creation of an audio data set that can be used to train an emergency detection model based on audio signals;
- Creation of a deep learning model for detecting emergency events and performing multiclassification based on audio signal pulses;
- Testing and evaluation of the constructed model.

Research methods. To perform the fixed tasks of the research work, the methods of machine learning, digital processing of audio signals, and acoustic method of recording audio data were used.

Scientific novelty and significance of research work. A combined architecture of the CNN-BiLSTM neural network for audio data processing has been created. The model was trained on a data set collected for several classes of emergencies, for detecting emergencies based on audio signal pulses and performing multiclassification in real time. The number of trained parameters is increased by using the CNN-BiLSTM combination.

Practical significance of the study. The proposed model detects eight classes of emergency situations based on аудио signal processing in public places such as railway stations, airports, transport stations, shopping and entertainment centers, educational institutions, residential complexes and other infrastructures.

The impact of the obtained results on the development of science and technology.

The data set and the results obtained in the research work can be used by other researchers as a methodological guide to conduct a research on detecting emergency events based on audio signals, collecting and pre-processing data, selecting and training neural network models.

The abstract prepared based on the results of the dissertation was introduced into the educational process of the “Almaty Academy atov of the Ministry of Internal Affairs of the Republic of Kazakhstan named after M. Esbolatov”.

Personal contribution of the author. All the results of the dissertation research work were obtained by the author himself. The approving of tasks to achieve the research goal, the analysis of research methods and the implementation of the proposed system, the analysis of the results of scientific research were conducted under the guidance of the author, his supervisor and a foreign supervisor. The results of other researchers used in the study were indicated by references to the relevant literature.

The relationship of research work with other research works. The project "AR149715555 Design and implementation of a security system inside a building in real time using machine learning methods" is being implemented as part of competition grant funding of the research of young scientists under the "Zhas Galym" project for 2022-2024.

Results submitted for protection.

1. A data set of eight different classes describing emergency situations.

2. A model that detects emergencies in eight different classes based on audio data processing.

3. Combined neural network architecture CNN-BiLSTM, which allows to increase the accuracy of emergency detection based on the audio signal.

Publications. According to the results of the research work, eight scientific articles were published, two articles in international peer-reviewed scientific journals with a non-zero impact factor included in the Scopus database (79% percentile, 31%), two articles in conferences included in the Scopus database, three articles in publications recommended by the CQFSHE of the Ministry of Science and Higher Education of the Republic of Kazakhstan, 1 article in other international conferences.

Approbation of the results of the study. The main results of the dissertation research were presented and discussed at seminars of the Department of "Radio Engineering, Electronics and Space Technologies" of KazNRTU named after K.I.Satpayev (Kazakhstan, Almaty); International Scientific and Practical Conference "Satbayev Readings -2020", Satbayev University (Kazakhstan, Almaty, 2020); 2020 8th International Conference on Information Technology and Multimedia (ICIMU) – IEEE, (Selangor, Malaysia 2020); 13th International Conference, ICCCI 2021, (Kallithea, Rhodes, Greece, 2021)

The structure and scope of the dissertation. The dissertation work consists of an introduction, 4 parts, a conclusion, and a list of references. The total volume of the work consists of 104 pages, 56 figures, 6 tables, the list of references consists of 99 titles.