

Michał MAZUR ORCID 0000-0003-2208-3057

Michał WIŚNIOŚ ORCID 0000-0003-4642-9585

Jacek PAŚ ORCID 0000-0001-8900-1445

Military University of Technology (Wojkowska Akademia Techniczna)

## ANALYSIS OF FALSE ALARM CAUSES IN VIDEO FIRE DETECTION SYSTEMS

### Analiza przyczyn fałszywych alarmów występujących w wizyjnych systemach detekcji pożaru

**Abstract:** Video-based fire detection systems represent an innovative path in fire signalling. Thanks to a suitably designed algorithm, a system of this kind can enable the detection of a flame based on its characteristics such as colour or shape, which were not previously used in classical fire detection systems. Video-based detection systems, due to their early stage of development in the fire protection market, are not yet a certified, fully tested method for early fire detection. This paper focuses on the analysis of possible causes of false alarms occurring in video-based fire detection systems in relation to classical Fire Alarm Systems (FAS). For this purpose, a video-based flame detection algorithm is designed and implemented to further analyse the phenomena occurring in such systems.

**Keywords:** video-based flame detection, false alarms, neural network, machine learning, fire alarm systems

**Streszczenie:** Systemy wizyjnej detekcji pożaru stanowią innowacyjną ścieżkę w zakresie sygnalizacji pożarowej. Dzięki odpowiednio zaprojektowanemu algorytmowi, system tego rodzaju może umożliwiać detekcję płomienia na podstawie takich jego cech jak barwa lub kształt, które do tej pory nie były wykorzystywane w klasycznych systemach wykrywania pożaru. Systemy wizyjnej detekcji ze względu na wczesny okres ich rozwoju na rynku systemów ochrony przeciwpożarowej, nie są jeszcze certyfikowaną, w pełni sprawdzoną metodą wczesnego wykrywania pożarów. Niniejszy artykuł skupia się na analizie możliwych przyczyn fałszywych alarmów występujących w wizyjnych systemach detekcji pożaru w odniesieniu do klasycznych Systemów Sygnalizacji Pożarowej (SSP). W tym celu zaprojektowany i zaimplementowany zostały algorytm wizyjnej detekcji płomienia, który pozwoli dokładniej przeanalizować zjawiska zachodzące w tego rodzaju systemach.

**Słowa kluczowe:** wizyjna detekcja płomienia, fałszywe alarmy, sieci neuronowe, uczenie maszynowe, systemy przeciwpożarowe

## **1. Introduction**

Video fire detection is a new development in fire protection. This method is based on fire characteristics hitherto unused in classic FAS, such as colour, shape or even movement of the flame or smoke [1]. Such systems are most often based on intelligent algorithms that are developed based on machine learning techniques or neural networks, whose task is to process the image, extract the Desired features and, based on these, classify whether a given image, or an area of it, contains the phenomenon of flame [2].

A flame is a relatively common but very complex and dynamic phenomenon. Its shape, colour or flicker frequency can change significantly over time, depending on the prevailing weather conditions or the fuel being burned. These peculiarities of the flame mean that, despite the efforts of researchers studying video-based fire detection systems, they still have a significant number of false alarms.

To investigate the possible causes of false alarms in video-based fire detection systems, an algorithm based on the CIFAR-10 neural network was designed to recognise flame and objects or phenomena resembling flame.

The main objective of this paper is to analyse the most common causes of false alarms in video-based fire detection systems, and to present methods that can contribute to the complete or partial elimination of some of the causes of such errors.

## **2. Literature review**

The problem of false alarms in video-based fire detection systems has already been addressed in many articles. In the article [2], the authors noted that most of the algorithms for video-based fire detection, are based on analysing the image for colour, spectrum, motion and other spatial and temporal features of the phenomenon. The authors in this article focused on machine learning-based algorithms using classifiers for fire detection, which means that the extraction of fire features for these types of algorithms is the responsibility of the designer. A large part of these algorithms are based on the Support-Vector Machine (SVM) classifier. Their proposed method for reducing false alarms is to replace the SVM classifier, with a Naive Bayes (NB) classifier. The NB classifier needs less training data, is faster in terms of classification and is less sensitive to features that are not related to the fire characteristics (FC).

The authors in the paper [3] proposed an algorithm whose detection is based on the colour of the pixels and the movement of objects in a given image. The innovation in the proposed method is the use of the so-called Tracking Growth Object (TGO) instead of standard motion detection. A flame does not change its position in principle, but can spread. The proposed algorithm will ignore objects with a colour similar to the flame that change their position in the image, while effectively detecting a flame that will occupy more and

more of the area over time. This method, in theory, can eliminate false alarms caused by, for example, car headlights or reflective clothing items.

A different approach to the problem was proposed by the authors in a study [4]. In their paper, they namely focused on the aspect of flame motion. Using optical motion estimators such as Optical Mass Transport (OMT) and Non-Smooth Data (NSD), they managed to develop an algorithm model that is able to detect the flame mainly based on its motion. Such a model is relatively immune to interference from objects with a colour similar to the flame, but is susceptible to disturbances in the quality of the transmitted image and some moving objects.

The author of the article [1] proposes to use an algorithm based on a convolutional neural network, whose performance will be supported by a Channel Attention Module (CAM). This is a module that updates the feature map by weighting and summing all the channels, then captures significant relationships between channels in the deepest layers of the network and thus improves the generalisability of the model. The model itself was designed based on a YOLO v3-type neural network. Tests have shown that the model is able to detect fire with an efficiency of up to 91.74%, thus achieving a score 5.95% higher than the model without this module.

The authors in the study [5] created a model in which they use a Faster Region-based Convolutional Neural Network (R-CNN) to delineate areas in an image where fire may be present. They use the Structural Similarity Index Measure (SSIM) and the Mean Square Error (MSE) to distinguish images where flame or smoke may be present. The R-CNN then determines the area of the image where fire may be present. Detection is performed using spatial and temporal features such as wavelet transform, coefficient of variation, colour histogram, frame similarity and mean square error for the delineated area. The results included in the paper suggest that the fire detection performance of the algorithm is estimated to be in the region of 99%.

In order to reduce the number of false alarms, the authors of the article [6] decided to integrate the flame and smoke detection process with the object tracking process. The video-based flame detection model they developed is based on a YOLO v4-type neural network. The tracking process itself consists of multiple steps and is based on the centroids of objects detected as fire. When an object or phenomenon is interpreted as a flame or smoke, its boundaries are sent as input to the tracking algorithm. This algorithm then uses the received coordinates to calculate the centroids. The calculation is performed for each frame where the detector believes the flame or smoke is located. The use of the tracking algorithm reduces the number of false alarms by confirming whether the detected object has time-varying parameters.

### **3. Algorithm for video-based fire detection**

In order to analyse the causes of false alarms occurring in video-based fire detection systems (VFDS), a proprietary model of such a system was developed for FC detection.

The algorithm was developed based on deep learning techniques, using a convolutional neural network. As the algorithm is designed to capture phenomena that may cause a false alarm, it will have a higher sensitivity than algorithms presented in previous literature on fire phenomena.

A CIFAR-10 neural network was used to develop the model, and the Transfer Learning (TL) method was used to train the network, which allows the algorithm to be trained more efficiently with less training data. This relies on the fact that only the last layers of the network, responsible for classifications, are modified during the training process. The CIFAR-10 neural network was developed and made available by MathWorks. It is a network designed to detect 10 different objects. Using TL, the last layers of the network were modified so that it detects the FC instead of 10 objects.

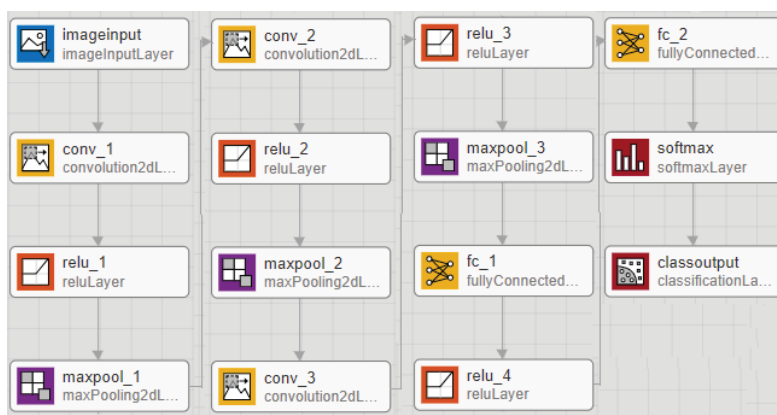


Fig. 1. Layers of neural network

Figure 1 shows the structure of the neural network used. The imageinput layer is the input layer, it applies normalisation to the input data. In this case, it reduces the resolution of the images to 32x32 pixels.

The second layer is the convolution layer. It performs a pixel value splicing operation with the help of a dimensioned filter. The filter moves through the pixel matrix with a defined step. The resulting output matrix is passed to the next layer. The size of the filter, the step and the size of the output matrix can be defined by the user, but their size values are likely to change during the training process.

The ReLu (Rectified Linear Unit) layer introduces non-linearity into the network. It changes all negative values to zero. This operation is performed for each element of the input matrix.

The maxpool layer reduces the dimension of the input matrix. The prefix "max" means that from each region of the matrix, a maximum value is extracted, which will be an element of the output matrix. In the case of the designed algorithm, the dimensions of the region from which the maximum value is extracted are 3x3 pixels. The step size in this case is 2 pixels.

The fullyConnected layer connects all neurons in the input layer to all neurons in the output layer. The input matrix, passing through this layer, is multiplied by a matrix of weight values. It also adds a so-called 'bias vector', which allows the activation function to be corrected, something that is necessary for the correct training process of the algorithm.

The softmax layer contains the previously mentioned activation function. Based on the information obtained, it determines the probability of whether an object is present in the image (in this case, a flame).

The final classoutput layer calculates the cross entropy loss of the classification problem. The layer determines how much the obtained probability is wrong. The loss value for an ideal neural network model is zero. The function that fulfils this task, often referred to in the literature as a cost function. The cost function in this case has the form of expression (1):

$$H(p, q) = - \sum_x p(x) \log q(x) \quad (1)$$

A stochastic gradient descent with momentum (SGDM) was used to train the original version of the network (CIFAR-10). The main difference between this method and other gradient methods, is that with a stochastic gradient, calculations at each step are made on the basis of only one observation, rather than on the entire set. As a result, the method is used with huge training sets, as it reduces the use of computing power and speeds up the training process. The disadvantage of this solution is the possibility of not finding the minimum of the function. Once the algorithm approaches the minimum, it may start 'circling' around it. However, from a machine learning point of view, this is not important, as minimisation to a value close to zero is sufficient and does not have to be zero either. Using additional momentum reduces the risk of the algorithm getting stuck in a local minimum and further speeds up the training process. The computer on which the model was trained has 16 Gb of RAM and a Windows 10 operating system.

The process of training the algorithm of the original version of the network (to detect 10 different objects) took 40 cycles. One cycle is defined as one complete pass over the entire training dataset. The total training process took approximately 60 minutes. The training data was 10,000 images representing the 10 different objects that this network is supposed to detect.

Training the final version of the network that will detect the flame is done almost identically, except for the number of epochs and the training dataset. The training dataset used to train the network was developed by the authors of the paper. The vast majority of it consists of images of the fire taken in the fire chamber of the Main School of Fire Service and images from the internet. The number of cycles is 30. The structures and weights of the whole network were mostly defined during the training of the first version of the network. The training process of the second version of the network took 60 minutes. 750 flame images were used for training.

## 4. Research methodology

The D-Fire database [6] was used for the study. It contains approximately 50 000 images depicting fire. From the database, 1000 images depicting fire and 1000 images without fire were drawn. A model was then developed and tested on the pool of drawn images. The algorithm was configured to signal detection only if the probability of correctly interpreting an object as a flame is more than 99%. We consider a correct detection to be the case where the algorithm correctly detects the flame and indicates its position in the image.

## 5. Findings

Table 1 shows the results of the tests carried out. Presenting them in the form of a confusion matrix will allow us to determine the sensitivity (True Positive Rate, TPR), precision (Positive Predictive Value, PPV) and accuracy (ACC) of the developed algorithm. The frames in which the algorithm correctly detected the flame were classified as True Positive (TP). The frames in which the algorithm correctly failed to detect the flame were classified as True Negative (TN). Images that contained a flame but no flame was detected were classified as False Negative (FN). Images that did not contain a flame, but triggered detection, were classified as False Positive (FP). The calculations of TPR, PPV and ACC are expressions 2 to 4, respectively.

**Table 1**

### Tests results

TP	TN	FN	FP
697	705	303	295

$$TPR = \frac{TP}{TP+FN} = \frac{697}{697+303} = 0,697 \quad (2)$$

$$PPV = \frac{TP}{TP+FP} = \frac{697}{697+295} = 0,702 \quad (3)$$

$$ACC = \frac{TP+TN}{TP+FN+FP+TN} = \frac{1402}{2000} = 0,701 \quad (4)$$

During the test, the most false alarms were caused by shiny surfaces (those that shine due to light reflection) with a colour close to yellow or red, car headlights, front as well as rear lights during braking and light sources with a colour close to yellow or red. Figures 2 and 3 show examples of images where flame detection has been signalled.

Based on the nature of the phenomena that cause false alarms, there are two main characteristics that each of these phenomena possesses. The first is a sufficiently high light intensity. Even if the colour of the emitted light is identical to that of the flame, it will not be detected if the intensity of this light source is too low.



**Fig. 2.** True Positive detection



**Fig. 3.** False Negative detection

The second common feature for phenomena that cause alarm is the visible light spectrum, i.e. the wavelength  $\lambda$  of the electromagnetic wave, or colloquially the colour of light.

The phenomena with the highest probability of causing a false alarm were characterised by the emission of light whose wavelengths  $\lambda$  range from  $\lambda = 566$  nm to  $\lambda = 700$  nm, i.e. from yellow to red [7]. The algorithm was completely immune to phenomena emitting blue light, white light or a combination of several colours simultaneously.

## 6. Conclusions

Considering the effects of the measurements carried out, the following phenomena may be the main cause of false alarms in video-based detection systems:

- yellow, orange or red so-called glossy surfaces of movable or immovable interior furnishings in the room, e.g. furniture, paintings, painted walls, light fittings, TV and computer cases and screens, etc,
- light sources in the yellow to red spectral range,
- metallic surfaces occurring in objects subject to fire protection,
- moving, non-matt objects in the room, e.g. air-conditioning fan impeller, etc.

Glossy surfaces, by themselves, are not able to cause alarm. This requires light of sufficient intensity and wavelength  $\lambda$  incident on the surface in question.

During the measurements, it was also possible to see the effect of the ambient light levels of the indoor (artificial) and outdoor (natural) environments on the algorithm's performance. In a low to moderately illuminated environment, the algorithms were able to detect the flame without any problems and the number of false alarms was relatively low. As the light levels in a given environment increased, the number of false alarms also increased. Due to the increasing intensity of light reaching the camera from across the monitored area, it became more difficult for the algorithms to distinguish the characteristics of a flame. There were cases where the algorithm was unable to make a detection or the areas where it was interpreted as a flame were also the entire monitored scene (image) was incorrectly classified. In addition to this, a high level of light intensity in the room to be protected can cause objects or surfaces that previously did not trigger false alarms to change the value of this parameter in a meaningful way. As the intensity of light in a given environment increases, the intensity of light reflected by objects in that environment will also increase.

To prevent false alarms, it is important to position the camera appropriately in the room in question, as with detectors designed to detect FC. In addition, the camera should have the best possible technical parameters (resolution, viewing angle, focal length, etc.), as the quality of the transmitted image is important for correct detection. If the quality is not sufficiently high, the captured image will not be able to convey the relevant flame-related features to the algorithm. Furthermore, the correct design of the algorithm is a fundamental factor for its correct functioning, i.e. the detection of FC.

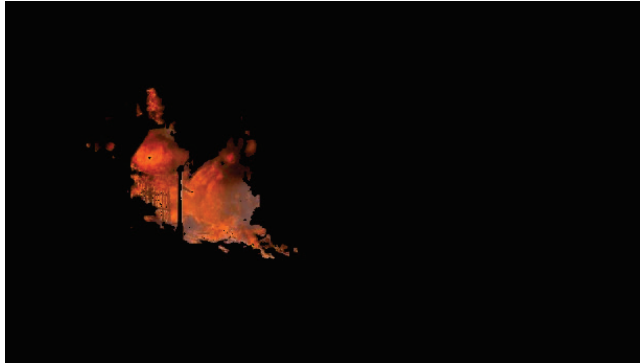
In order to reduce the number of false alarms, a number of solutions can be used to improve the effectiveness of such systems in various aspects.

The first method is the use of TGO [3]. This method will eliminate potential sources of false alarms from moving objects such as cars, motorbikes, bicycles, scooters, people or low-flying aircraft.

Another method may be to use a model that focuses on motion analysis [4,6]. This will allow a thorough analysis of a given slice of the image containing an object interpreted as fire and confirm whether the movement of this object is the same or at least similar to the dynamic movement of the flame or smoke.



Yet another method is colour filtering. By using this simple mechanism, it is possible to pre-filter objects (images) which, in shape and movement, may resemble a flame or smoke without being one. An example of such an object could be a flying national flag or blinds in windows.



**Fig. 4.** Processed Fig. 3 after colour filtering

It will also be effective to use logical conditions. A fire detection signal will only be sent to the user when an object interpreted as a fire has been detected for a given time interval, e.g. another 10 -15 seconds. This will eliminate short flashes of light that could cause a false alarm - such as welding.

Some models use spatial as well as temporal features such as the direction of flame propagation or flicker frequency [8].

## **7. Summary**

Video fire detection systems are based on optical signal analysis. The images 'captured' by the camera are transformed and analysed by an algorithm, which is the decision-making element in this type of system. These systems are susceptible to optical interference occurring in a given indoor environment subject to fire protection. In classical fire detection systems, which phenomenon a detector is susceptible to depends on the type of detector and the characteristics used for detection, which are usually related to the physical processes occurring during combustion (FC) [9,10]. The difference between the use of fire phenomena for detection, results in the causes of false alarms generally also being different.

Despite using completely different characteristics of fire for its detection, both systems (i.e. classic Fire Alarm Systems (FAS) and video FAS) share their susceptibility to certain phenomena that cause false alarms to a similar degree. These are phenomena that adversely affect most electronic devices, such as disturbance of the electromagnetic compatibility of the environment, inadequate training of service personnel, improper installation of detectors, the presence of light sources with high intensity and different wavelengths  $\lambda$ ,

errors at the design stage or deliberate actions of people using the FAS [11,12]. These phenomena are the only element in common between the two FC detection systems - the classic and the video-based system.

Early detection video-based systems are much more sensitive than classical systems [13,14]. This is due to the very fast detection time, in the order of several seconds. However, the high sensitivity also means that the risk of a false alarm is much higher. The use of only this type of system is therefore hardly recommended at the current stage of technical development. The best effect is achieved by combining video detection systems in support of the classic FAS (fusion of two systems). This will increase the sensitivity of the entire system due to the fast camera response time.

In the near future, developed research and the use of various signal processing algorithms are planned to further investigate methods for reducing false alarms in video-based fire detection systems.

## **8. References**

1. X. Zhang, K. Qian, K. Jing, J. Yang, H. Yu, "Fire Detection based on Convolutional Neural Networks with Channel Attention", Chinese Automation Congress, China 2020.
2. T. E. Abidha, P. Mathai Paul, "Reducing False Alarms in Vision Based Fire Detection with NB Classifier in EADF Framework", Federal Institute of Science and Technology, Angamaly 2020.
3. A. Khalil, S.U. Rahman, F. Alam, I. Khalil, "Fire Detection Using Multi Color Space and Background Modeling", University of Malakand, Pakistan 2020.
4. P. Karasev, I. Kolesov, A. Tannenbaum, "Optical Flow Estimation for Flame Detection in Videos", IEEE Transactions on Image Processing, 2013.
5. Y. Lee, J. Shim, "False Positive Decrement Research for Fire and Smoke Detection in Surveillance Camera using Spatial and Temporal Features Based on Deep Learning", Andong National University, Andong 2019.
6. P. V. Almeida, T. M. Rezende, A. C. Lisboa, A. V. Barbosa, "Fire Detection based on a Two-Dimensional Convolutional Neural Network and Temporal Analysis", 7th IEEE LA-CCI, Temuco 2021.
7. M. P. Thekaekara, "Solar energy outside the earth's atmosphere", Solar Energy Vol. 14, p.109-127, Maryland 1971.
8. T. Chen, X. Guo, J. Jia, J. Xiao, "Frequency and Phase Characteristics of Candle Flame Oscillation", Nature 2019.
9. J. Antos, M. Basiak, D. Kręciwilk, „Wpływ fałszywych alarmów pożarowych na skuteczność pracy monitoringu pożarowego”, Częstochowa 2016.
10. A. Żyłuk, M. Zieja, A. Szelmanowski, J. Tomaszewska, M. Perlińska, K. Głyda, „Electrical Disturbances in Terms of Methods to Reduce False Activation of Aerial Fire Protection Systems”, 2022, <https://doi.org/10.3390/s22208059>.

11. T. Klimczak, J. Paś, S. Duer, A. Rosiński, P. Wetoszka, K. Białek, M. Mazur, “Selected Issues Associated with the Operational and Power Supply Reliability of Fire Alarm Systems”, *Energies* 2022, 15(22), 8409; <https://doi.org/10.3390/en15228409>.
12. K. Głyda, A. Szelmanowski, J. Sulkowski, A. Pazur, „False actions of the aviation fire protection systems caused by disturbances in power circuits”, 2022, *Journal of KONBiN*, volume 52, DOI 10.2478/jok-2022-0037.
13. J. Paś, T. Klimczak, A. Rosiński et al. „The analysis of the operational process of a complex fire alarm system used in transport facilities”, *Build. Simul.* 15, 615–629 2022. <https://doi.org/10.1007/s12273-021-0790-y>.
14. A. Szelmanowski, M. Zieja, K. Głyda, „Dynamic properties modeling of the thermoelectric fire sensors in the aircraft fire suppression system”, 2017, *Journal of KONBiN*, vol. 44, DOI 10.1515/jok-2017-0074.

# **ANALIZA PRZYCZYŃ FAŁSZYWYCH ALARMÓW WYSTĘPUJĄCYCH W WIZYJNYCH SYSTEMACH DETEKCJI POŻARU**

## **1. Wprowadzenie**

Wizyjna detekcja pożaru jest nowym rozwiązaniem w dziedzinie ochrony przeciwpożarowej. Metoda ta bazuje na niewykorzystywanych do tej pory w klasycznych systemach SSP cechach pożaru takich jak barwa, kształt czy nawet ruch płomienia lub dymu [1]. Takie systemy najczęściej opierają się na inteligentnych algorytmach, które są opracowywane w oparciu o techniki uczenia maszynowego lub sieci neuronowych, których zadaniem jest przetworzenie obrazu, wyodrębnienie pożądaných cech i na ich podstawie sklasyfikowanie, czy dany obraz lub jego obszar, zawiera zjawisko płomienia [2].

Płomień jest względnie pospolitym, lecz bardzo skomplikowanym i dynamicznym zjawiskiem. Jego kształt, kolor lub częstotliwość migotania mogą ulegać znacznym zmianom w czasie, w zależności od panujących warunków atmosferycznych lub spalanego paliwa. Te charakterystyczne cechy płomienia sprawiają, że pomimo wysiłków naukowców badających wizyjne systemy detekcji pożaru, nadal charakteryzują się one znaczną liczbą fałszywych alarmów.

W celu zbadania możliwych przyczyn fałszywych alarmów w wizyjnych systemach detekcji pożaru, zaprojektowany został algorytm na bazie sieci neuronowej CIFAR-10, którego zadaniem jest rozpoznawanie płomienia oraz obiektów lub zjawisk, przypominających płomień.

Głównym celem artykułu jest analiza najczęstszych przyczyn fałszywych alarmów w wizyjnych systemach detekcji pożaru oraz przedstawienie metod, które mogą przyczynić się do całkowitej lub częściowej eliminacji niektórych przyczyn powstawania tego rodzaju błędów.

## **2. Przegląd literatury**

Problematyka zjawiska fałszywych alarmów w wizyjnych systemach detekcji pożaru była już poruszana w wielu artykułach. W artykule [2] autorzy zauważyli, iż większość algorytmów wizyjnej detekcji pożaru, opiera się na analizie obrazu pod kątem barwy, widma, ruchu i innych przestrzennych oraz czasowych cech tego zjawiska. Autorzy w tym artykule skupili się na algorytmach na bazie uczenia maszynowego, wykorzystujących klasyfikatory do detekcji pożaru, co oznacza, że za ekstrakcje cech pożaru w przypadku

tego rodzaju algorytmów odpowiada jego projektant. Znaczna część tych algorytmów opiera się na klasyfikatorze SVM (ang. Support-Vector Machine). Zaproponowana metoda na ograniczenie fałszywych alarmów polega na zastąpieniu klasyfikatora SVM klasyfikatorem NB (ang. Naive Bayes). Klasyfikator NB potrzebuje mniejszej liczby danych uczących, jest szybszy pod względem klasyfikacji oraz jest mniej wrażliwy na cechy, które nie są związane z wielkością charakterystyczną pożaru (WCHP).

Autorzy w opracowaniu [3] zaproponowali algorytm, którego detekcja opiera się na kolorze pikseli oraz ruchu obiektów w danym obrazie. Innowacją w zaproponowanej metodzie jest wykorzystanie tzw. współczynnika wzrostu TGO (ang. Tracking Growth Object) zamiast standardowej detekcji ruchu. Płomień z zasady nie zmienia swojego położenia, może natomiast się rozprzestrzenić. Zaproponowany algorytm będzie ignorował obiekty o barwie zbliżonej do płomienia, które zmieniają swoje położenie na obrazie, natomiast skutecznie wykrywa płomień, który z upływem czasu będzie zajmował coraz więcej obszaru. Ta metoda w teorii pozwala na eliminację fałszywych alarmów powodowanych przez np. światła samochodów lub odblaskowe elementy ubioru osób znajdujących się w danym obiekcie.

Inne podejście do problemu zaproponowali autorzy w opracowaniu [4]. W swojej pracy skupili się mianowicie na aspekcie ruchu płomienia. Korzystając z estymatorów ruchu optycznego takich jak OMT (ang. Optical Mass Transport) i NSD (ang. Non-Smooth Data), opracowali model algorytmu, który jest w stanie wykryć płomień głównie na podstawie jego ruchu. Taki model jest stosunkowo odporny na zakłócenia związane z obiektami o barwie zbliżonej do płomienia, jednak jest podatny na zaburzenia w jakości przesyłanego obrazu oraz niektórych poruszających się obiektów.

Autor artykułu [1] proponuje zastosowanie algorytmu opartego na konwolucyjnej sieci neuronowej, którego działanie będzie wsparte przez CAM (ang. Channel Attention Module). Jest to moduł, który aktualizuje mapę obiektów poprzez ważenie i sumowanie wszystkich kanałów, następnie wychwytuje znaczące zależności między kanałami w najgłębszych warstwach sieci i tym samym poprawia zdolność generalizacji modelu. Sam model został zaprojektowany w oparciu o sieć neuronową typu Yolo v3. Testy pokazały, że model jest w stanie wykryć pożar ze skutecznością sięgającą 91,74%, tym samym osiągając wynik o 5,95% wyższy od modelu bez tego modułu.

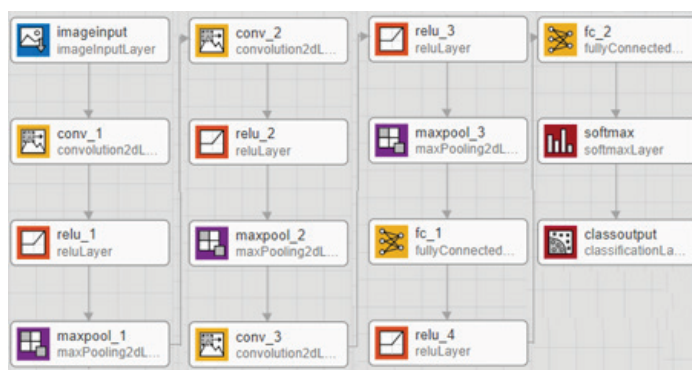
Autorzy w opracowaniu [5] stworzyli model, w którym wykorzystują konwolucyjną sieć neuronową R-CNN (Faster Region-based Convolutional Network) do wyznaczania obszarów na obrazie, w których może znajdować się pożar. Wykorzystują wskaźnik podobieństwa struktur SSIM (ang. Structural Similarity Index Measure) oraz błąd średniokwadratowy MSE (ang. Mean Square Error) do wyróżnienia obrazów, na których może znajdować się płomień lub dym. Następnie sieć R-CNN wyznacza obszar obrazu, na którym może znajdować się pożar. Detekcja dokonywana jest z wykorzystaniem cech przestrzennych i czasowych, takich jak transformata falkowa, współczynnik zmienności, histogram koloru, podobieństwo ramek i błąd średniokwadratowy dla wyznaczonego obszaru. Wyniki zawarte w artykule sugerują, że skuteczność wykrywania pożaru algorytmu szacowana jest w okolicach 99%.

Aby zmniejszyć liczbę fałszywych alarmów, autorzy artykułu [6] postanowili zintegrować proces detekcji płomienia i dymu z procesem śledzenia obiektów. Stworzony przez nich model wizyjnej detekcji płomienia opiera się na sieci neuronowej typu YOLO v4. Sam proces śledzenia składa się z wielu etapów i jest on oparty na centroidach obiektów wykrytych jako pożar. Gdy obiekt lub zjawisko zostanie zinterpretowane jako płomień lub dym, jego granice zostają przesłane jako wejście do algorytmu śledzącego. Następnie ten algorytm wykorzystuje otrzymane współrzędne do obliczenia centroidy. Obliczenia są wykonywane dla każdej klatki, w której według detektora znajduje się płomień lub dym. Zastosowanie algorytmu śledzącego pozwala na zmniejszenie liczby fałszywych alarmów, poprzez potwierdzenie, czy wykryty obiekt posiada parametry zmienne w czasie.

### 3. Algorytm wizyjnej detekcji pożaru

W celu analizy przyczyn fałszywych alarmów występujących w wizyjnych systemach detekcji pożaru (WSDP), opracowano własny model takiego systemu do wykrywania WCHP. Algorytm został stworzony na bazie technik głębokiego uczenia, wykorzystując konwolucyjną sieć neuronową. Ze względu na to, że algorytm ma za zadanie uchwycić zjawiska, które mogą wywołać fałszywy alarm, będzie on charakteryzował się większą czułością niż algorytmy przedstawione we wcześniejszej literaturze dotyczącej zjawiska pożaru.

Do opracowania modelu została wykorzystana sieć neuronowa CIFAR-10, a do nauki sieci została wykorzystana metoda „uczenie przez przeniesienie” TL (ang. Transfer Learning), która pozwala na skuteczniejszą naukę algorytmu, w przypadku mniejszej ilości danych uczących. Polega to na tym, że podczas procesu uczenia modyfikowane są tylko ostatnie warstwy sieci, odpowiedzialne za klasyfikację. Sieć neuronowa CIFAR-10 została opracowana i udostępniona przez MathWorks. To sieć, której zadaniem jest detekcja 10 różnych obiektów. Wykorzystując TL, zmodyfikowane zostały ostatnie warstwy sieci tak, aby zamiast 10 obiektów wykrywała WCHP.



Rys. 1. Warstwy sieci neuronowej

Rysunek 1 przedstawia strukturę wykorzystanej sieci. Warstwa `imageinput` to warstwa wejściowa, stosuje normalizację danych wejściowych. W tym przypadku zmniejsza rozdzielczość obrazów do 32x32 pikseli.

Druga warstwa to warstwa konwolucyjna. Za pomocą filtra o określonych wymiarach dokonuje ona operacji splotu wartości pikseli. Filtr porusza się po macierzy pikseli z określonym krokiem. Powstała w ten sposób macierz wyjściowa jest przekazywana do kolejnej warstwy. Rozmiar filtra, krok i wielkość macierzy wyjściowej mogą być zdefiniowane przez użytkownika, jednak wartości ich rozmiarów najprawdopodobniej ulegną zmianom podczas procesu uczenia.

Warstwa `ReLU` (ang. Rectified Linear Unit) wprowadza nieliniowość do sieci. Zmienia ona wszystkie wartości ujemne na zero. Operacja ta jest wykonywana dla każdego elementu macierzy wejściowej.

Warstwa `maxpool` redukuje wymiar macierzy wejściowej. Przedrostek „max” oznacza, że z każdego regionu macierzy ekstrahowana jest maksymalna wartość, która będzie elementem macierzy wyjściowej. W przypadku projektowanego algorytmu, wymiary regionu, z którego ekstrahowana jest maksymalna wartość, to 3x3 piksele. Wielkość kroku w tym przypadku to 2 piksele.

Warstwa `fullyConnected` łączy wszystkie neurony w warstwie wejściowej z wszystkimi neuronami w warstwie wyjściowej. Macierz wejściowa, przechodząc przez tę warstwę zostaje pomnożona przez macierz wartości wag. Dodaje ona także tzw. bias vector, który umożliwia korygowanie funkcji aktywacji, czyli rzecz niezbędną do poprawnego przeprowadzenia procesu uczenia algorytmu.

Warstwa `softmax` zawiera wcześniej wspomnianą funkcję aktywacji. Na podstawie uzyskanych informacji, określa ona prawdopodobieństwo tego, czy dany obiekt znajduje się na obrazie (w tym przypadku jest to płomień).

Ostatnia warstwa `classoutput` oblicza straty entropii krzyżowej problemu klasyfikacji. Warstwa określa, jak bardzo uzyskane prawdopodobieństwo jest błędne. Wartość strat dla idealnego modelu sieci neuronowej wynosi zero. Funkcja spełniająca to zadanie, często w literaturze określania jest jako funkcja kosztów (ang. cost function). Funkcja kosztów w tym przypadku ma postać wyrażenia (1):

$$H(p, q) = - \sum_x p(x) \log q(x) \quad (1)$$

Do nauczenia oryginalnej wersji sieci (CIFAR-10) wykorzystano gradient stochastyczny z pędem `SGDM` (ang. stochastic gradient descent with momentum). Podstawową różnicą między tą metodą a innymi metodami gradientowymi jest fakt, że w przypadku gradientu stochastycznego obliczenia w każdym kroku dokonywane są na podstawie tylko jednej obserwacji, a nie na całym zbiorze. Dzięki temu, metoda wykorzystywana jest przy ogromnych zbiorach uczących, ponieważ pozwala na ograniczenie użycia mocy obliczeniowej oraz przyspiesza proces uczenia. Wadą tego rozwiązania jest możliwość niezalezienia minimum funkcji. Algorytm po zbliżeniu się do minimum może zacząć „krążyć” wokół niego. Jednak z punktu widzenia uczenia

maszynowego nie jest to istotne, ponieważ minimalizacja do wartości bliskiej zeru jest wystarczająca i nie musi być też równa zeru. Zastosowanie dodatkowego pędu zmniejsza ryzyko utknięcia algorytmu w minimum lokalnym oraz dodatkowo przyspiesza proces uczenia. Komputer, na którym uczono model, ma 16 Gb pamięci RAM oraz system operacyjny Windows 10.

Proces uczenia algorytmu pierwotnej wersji sieci (do wykrywania 10 różnych obiektów) trwał 40 cykli. Jeden cykl zdefiniowany jest jako jedno kompletne przejście po całym zbiorze danych uczących. Całkowity proces uczenia trwał w przybliżeniu 60 minut. Dane uczące to 10 000 zdjęć przedstawiające 10 różnych obiektów, które ta sieć ma wykrywać.

Uczenie finalnej wersji sieci, która będzie wykrywać płomień, odbywa się niemal identycznie, z wyjątkiem liczby epok i zbioru danych uczących. Zbiór danych uczących wykorzystany do nauczenia sieci został opracowany przez autorów artykułu. W znacznej większości są to zdjęcia pożaru wykonane w komorze pożarowej Szkoły Głównej Służby Pożarniczej oraz obrazy z internetu. Liczba cykli wynosi 30. Struktury i wagi całej sieci zostały w większości zdefiniowane podczas uczenia pierwszej wersji sieci. Proces uczenia drugiej wersji sieci trwał 60 minut. Do nauki wykorzystano 750 obrazów płomienia.

## **4. Metodyka badań**

Do badania wykorzystano bazę danych D-Fire [6]. Zawiera ona w przybliżeniu 50 000 obrazów przedstawiających pożar. Z bazy danych wylosowano 1 000 zdjęć przedstawiających pożar oraz 1 000 zdjęć bez zjawiska pożaru. Następnie opracowano model, który został przetestowany na puli wylosowanych obrazów. Algorytm został skonfigurowany tak, aby sygnalizował detekcje tylko w przypadku, gdy prawdopodobieństwo poprawnej interpretacji obiektu jako płomienia wynosi ponad 90%. Jako poprawną detekcję przyjmujemy przypadek, w którym algorytm poprawnie wykryje płomień oraz wskaże jego położenie na obrazie.

## **5. Wyniki badań**

Tabela 1 przedstawia wyniki przeprowadzonych badań. Przedstawienie ich w postaci tablicy pomyłek (błędów) pozwoli nam określić czułość (ang. True Positive Rate, TPR), precyzję (ang. Positive Predictive Value, PPV) oraz dokładność (ang. Accuracy, ACC) opracowanego algorytmu. Klatki, w których algorytm poprawnie wykrył płomień, zostały sklasyfikowane jako TP (ang. True Positive). Klatki, w których algorytm nie wykrył poprawnie płomienia, sklasyfikowano jako TN (ang. True Negative). Obrazy, które zawierały płomień, jednak nie został on wykryty, sklasyfikowane zostały jako FN (ang. False Negative). Obrazy nie zawierające płomienia jednak wywołujące detekcje zostały



oznaczone jako FP (ang. False Positive). Obliczenia wskaźników TPR, PPV oraz ACC to odpowiednio wyrażenia od 2 do 4.

**Tabela 1**

**Wyniki testu**

TP	TN	FN	FP
697	705	303	295

$$TPR = \frac{TP}{TP+FN} = \frac{697}{697+303} = 0,697 \quad (2)$$

$$PPV = \frac{TP}{TP+FP} = \frac{697}{697+295} = 0,702 \quad (3)$$

$$ACC = \frac{TP+TN}{TP+FN+FP+TN} = \frac{1402}{2000} = 0,701 \quad (4)$$

Podczas badania najwięcej fałszywych alarmów powodowały powierzchnie połyskliwe (takie, które błyszczą dzięki odbijaniu się światła) o barwie zbliżonej do żółtego lub czerwonego koloru, światła samochodów, zarówno przednie, jak i tylne podczas hamowania, oraz źródła światła o barwie zbliżonej do koloru żółtego lub czerwonego. Na rys. 4–10 przedstawiono przykładowe obrazy, na których zasygnalizowana została detekcja płomienia.

Na podstawie charakteru zjawisk wywołujących fałszywe alarmy, można wyróżnić dwie główne cechy, które charakteryzują każde z tych zjawisk. Pierwszą jest odpowiednio wysokie natężenie światła. Nawet jeśli barwa emitowanego światła jest identyczna z barwą płomienia, nie zostanie ona wykryta, jeśli natężenie tego źródła jest zbyt niskie.



**Rys. 2.** Detekcja True Positive



Rys. 3. Detekcja False Negative

Drugą wspólną cechą dla zjawisk powodujących alarm jest zakres spektrum światła widzialnego, czyli długości  $\lambda$  fali elektromagnetycznej lub potocznie barwa światła.

Zjawiska z największym prawdopodobieństwem wywołania fałszywego alarmu cechowały się emisją światła, którego długość fal  $\lambda$  zawiera się w zakresie od  $\lambda = 566$  nm do  $\lambda = 700$  nm, czyli od barwy żółtej do czerwonej [7]. Algorytm był całkowicie odporny na zjawiska emitujące światło niebieskie, białe lub połączenie kilku barw równocześnie.

## 6. Wnioski

Biorąc pod uwagę efekty przeprowadzonych pomiarów, główną przyczyną fałszywych alarmów w systemach wizyjnego wykrywania WCHP mogą być następujące zjawiska:

- żółte, pomarańczowe lub czerwone tzw. połyskliwe powierzchnie zgromadzonych w pomieszczeniu ruchomych lub nieruchomych środków wyposażenia wewnętrznego, np. mebli, obrazów, pomalowanych ścian, opraw źródeł światła, obudów i ekranów telewizyjnych i komputerowych itd.,
- źródła światła o barwie z zakresu spektrum od żółtego do czerwonego,
- powierzchnie metaliczne występujące w obiektach podlegających ochronie pożarowej,
- ruchome, niemiatowe obiekty znajdujące w pomieszczeniu, np. wirnik wiatraka do klimatyzacji itd.

Połyskliwe powierzchnie same w sobie nie są w stanie wywołać alarmu. Niezbędne jest do tego światło o odpowiednim natężeniu i długości  $\lambda$  fali padające na daną powierzchnie.

Podczas pomiarów można było także zauważyć wpływ poziomu oświetlenia otoczenia wewnętrznego (sztucznego) i zewnętrznego (naturalnego) na funkcjonowanie algorytmu. W środowisku słabo lub umiarkowanie oświetlonym, algorytmy bez problemu były w stanie wykryć płomień, a liczba fałszywych alarmów była względnie niska.

W momencie, gdy poziom oświetlenia w danym środowisku wzrastał, liczba fałszywych alarmów również wzrastała. Z powodu zwiększającego się natężenia światła docierającego do kamery z całego obszaru monitorowanego, algorytmy mają większe trudności z wyróżnieniem cech charakterystycznych dla płomienia. Występowały przypadki, gdzie algorytm nie był w stanie dokonać detekcji, lub obszary, w których został zinterpretowany jako płomień, była też cała monitorowana scena (obraz) był nieprawidłowo sklasyfikowana. Oprócz tego, wysoki poziom natężenia światła w ochranianym pomieszczeniu może powodować, że obiekty lub powierzchnie, które wcześniej nie wywoływały fałszywych alarmów, mogą zmieniać w sposób znaczny wartość tego parametru. Wraz ze wzrostem natężenia światła w danym środowisku, wzrastać będzie także natężenie odbijanego światła przez obiekty znajdujące się w tym środowisku.

Aby zapobiec fałszywym alarmom, istotne jest odpowiednie umiejscowienie kamery w danym pomieszczeniu, podobnie jak z czujkami przeznaczonymi do wykrywania WCHP. Oprócz tego, kamera powinna mieć najlepsze możliwe parametry techniczne (rozdzielczość, kąt obserwacji, ogniskowa itd.), ponieważ jakość przekazywanego obrazu jest istotna do poprawnej detekcji. Jeśli jakość nie jest odpowiednio wysoka, uchwycony obraz nie będzie w stanie przenieść do algorytmu odpowiednich cech związanych z płomieniem. Ponadto, poprawne zaprojektowanie algorytmu jest fundamentalnym czynnikiem wpływającym na jego prawidłowe funkcjonowanie, tj. wykrywanie WCHP.

Aby ograniczyć liczbę fałszywych alarmów, można zastosować szereg rozwiązań, które w różnych aspektach mogą przyczynić się do zwiększenia skuteczności takich systemów.

Pierwszą metodą jest zastosowanie TGO [3]. Metoda ta pozwoli na eliminację potencjalnych źródeł fałszywych alarmów ze strony poruszających się obiektów, takich jak: samochody, motory, rowery, hulajnogi, ludzie czy nisko przelatujące statki powietrzne.

Inną metodą może być zastosowanie modelu, który skupia się na analizie ruchu [4, 6]. Pozwoli to na dokładną analizę danego wycinka obrazu, na którym znajduje się obiekt zinterpretowany jako pożar, i potwierdzenie, czy ruch tego obiektu jest taki sam lub przynajmniej zbliżony do dynamicznego ruchu płomienia lub dymu.

Jeszcze inną metodą jest filtracja koloru. Zastosowanie tego prostego mechanizmu pozwoli na wstępną filtrację obiektów (obrazów), które kształtem i ruchem mogą przypominać płomień lub dym, jednocześnie nim nie będąc. Przykładem takiego obiektu może być powiewająca flaga państwowa lub żaluzje w oknach.

Skutecznym rozwiązaniem będzie również zastosowanie warunków logicznych. Sygnał o detekcji pożaru będzie wysyłany do użytkownika dopiero wtedy, gdy obiekt zinterpretowany jako pożar będzie wykrywany przez dany interwał czasowy, np. kolejne 10-15 sekund. Pozwoli to wyeliminować krótkie błyski światła, które mogłyby spowodować fałszywy alarm – np. spawanie.

Niektóre modele korzystają z przestrzennych oraz czasowych cech, takich jak kierunek rozprzestrzeniania się płomienia czy częstotliwość migotania [8].



Rys. 4. Obraz z rys. 3 po filtracji kolorów

## 7. Podsumowanie

Systemy wizyjnego wykrywania pożarów opierają się na analizie sygnału optycznego. Obrazy „uchwycone” przez kamerę są poddawane transformacji i analizie przez algorytm, który jest elementem decyzyjnym w tego rodzaju systemach. Systemy te są podatne na zakłócenia optyczne pojawiające się w danym środowisku wewnętrznym, podlegającym ochronie pożarowej. W klasycznych systemach detekcji pożaru, to na jakie zjawisko jest podatna czujka, zależy od jej rodzaju i cech wykorzystywanych do detekcji, które zazwyczaj są powiązane z procesami fizycznymi zachodzącymi podczas spalania (WCHP) [9, 10]. Różnica między wykorzystaniem zjawisk pożaru do jego detekcji, skutkuje tym, że przyczyny fałszywych alarmów z reguły również są inne.

Pomimo wykorzystania całkowicie innych cech pożaru do jego detekcji, oba systemy (czyli klasyczne Systemy Sygnalizacji Pożaru (SSP) oraz wizyjne SSP) dzielą swoją podatność na niektóre zjawiska powodujące fałszywe alarmy w podobnym stopniu. Są to zjawiska, które wpływają niekorzystnie na większość urządzeń elektronicznych, takie jak np. zaburzenie kompatybilności elektromagnetycznej środowiska, brak odpowiedniego przeszkolenia personelu serwisującego, nieprawidłowy montaż czujek, występowanie źródeł światła o dużym natężeniu i różnych długościach fali  $\lambda$ , błędy na etapie projektowania lub umyślne działania osób wykorzystujących użytkownicy SSP [11, 12]. Te zjawiska to jedyny element wspólny między dwoma systemami detekcji WCHP – system klasyczny i wizyjny.

Systemy wizyjne wczesnej detekcji są znacznie bardziej czułe od systemów klasycznych [13, 14]. Wiąże się to z bardzo szybkim czasem detekcji, rzędu kilkunastu sekund. Jednak wysoka czułość powoduje również, że ryzyko fałszywego alarmu jest znacznie większe. Stosowanie tylko tego rodzaju systemów nie jest więc praktycznie na obecnym etapie rozwoju technicznego zalecane. Najlepszy efekt uzyskuje się, łącząc systemy wizyjnej detekcji jako wsparcie dla klasycznych SSP (fuzja dwóch systemów).

Spowoduje to zwiększenie czułości całego systemu ze względu na szybki czas reakcji kamery.

W zakresie opracowanych badań naukowych i wykorzystania różnych algorytmów obróbki sygnałów planowana jest w niedalekiej przyszłości dalsza analiza metod redukcji fałszywych alarmów w systemach wizyjnej detekcji pożaru.

## **8. Literatura**

1. X. Zhang, K. Qian, K. Jing, J. Yang, H. Yu, "Fire Detection based on Convolutional Neural Networks with Channel Attention", Chinese Automation Congress, Chiny 2020.
2. T. E. Abidha, P. Mathai Paul, "Reducing False Alarms in Vision Based Fire Detection with NB Classifier in EADF Framework", Federal Institute of Science and Technology, Angamaly 2020.
3. A. Khalil, S. U. Rahman, F. Alam, I. Khalil, "Fire Detection Using Multi Color Space and Background Modeling", University of Malakand, Pakistan 2020.
4. P. Karasev, I. Kolesov, A. Tannenbaum, "Optical Flow Estimation for Flame Detection in Videos", IEEE Transactions on Image Processing, 2013.
5. Y. Lee, J. Shim, "False Positive Decremental Research for Fire and Smoke Detection in Surveillance Camera using Spatial and Temporal Features Based on Deep Learning", Andong National University, Andong, 2019.
6. P. V. Almeida, T. M. Rezende, A. C. Lisboa, A. V. Barbosa, "Fire Detection based on a Two-Dimensional Convolutional Neural Network and Temporal Analysis", 7<sup>th</sup> IEEE LA-CCI, Temuco 2021.
7. M. P. Thekaekara, "Solar energy outside the earth's atmosphere", Solar Energy Vol. 14, p.109-127, Maryland, 1971.
8. T. Chen, X. Guo, J. Jia, F. Xiao, "Frequency and Phase Characteristics of Candle Flame Oscillation", Nature, 2019.
9. J. Antos, M. Basiak, D. Kręciwilk, „Wpływ fałszywych alarmów pożarowych na skuteczność pracy monitoringu pożarowego”, Częstochowa 2016.
10. A. Żyłuk, M. Zieja, A. Szelmanowski, J. Tomaszewska, M. Perlińska, K. Głyda, „Electrical Disturbances in Terms of Methods to Reduce False Activation of Aerial Fire Protection Systems”, 2022, <https://doi.org/10.3390/s22208059>.
11. T. Klimczak, J. Paś, S. Duer, A. Rosiński, P. Wetoszka, K. Białek, M. Mazur, "Selected Issues Associated with the Operational and Power Supply Reliability of Fire Alarm Systems", Energies 2022, 15(22), 8409; <https://doi.org/10.3390/en15228409>.
12. K. Głyda, A. Szelmanowski, J. Sulkowski, A. Pazur, „Fałszywe zadziałania lotniczego systemu przeciwpożarowego wywołane zakłóceniami w obwodach zasilania”, 2022, Journal of KONBiN, volume 52, DOI 10.2478/jok-2022-0037.

13. J. Paś, T. Klimczak, A. Rosiński et al. „The analysis of the operational process of a complex fire alarm system used in transport facilities”, *Build. Simul.* 15, 615–629 (2022). <https://doi.org/10.1007/s12273-021-0790-y>.
14. A. Szelmanowski, M. Zieja, K. Głyda, „Modelowanie dynamicznych właściwości termoelektrycznych czujników pożaru w lotniczym systemie przeciwpożarowym”, 2017, *Journal of KONBiN*, volume 44, DOI 10.1515/jok-2017-0074.