

Research Article

Research on Application of the Feature Transfer Method Based on Fast R-CNN in Smoke Image Recognition

Xi Cheng 

Sichuan University of Arts and Science, Dazhou 635000, China

Correspondence should be addressed to Xi Cheng; 20030032@sasu.edu.cn

Received 20 October 2021; Revised 2 November 2021; Accepted 6 November 2021; Published 23 November 2021

Academic Editor: Qiangyi Li

Copyright © 2021 Xi Cheng. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Most of the existing smoke detection methods are based on manual operation, which is difficult to meet the needs of fire monitoring. To further improve the accuracy of smoke detection, an automatic feature extraction and classification method based on fast regional convolution neural network (fast R-CNN) was introduced in the study. This method uses a selective search algorithm to obtain the candidate images of the sample images. The preselected area coordinates and the sample image of visual task are used as network learning. During the training process, we use the feature migration method to avoid the lack of smoke data or limited data sources. Finally, a target detection model is obtained, which is strongly related to a specified visual task, and it has well-trained weight parameters. Experimental results show that this method not only improves the detection accuracy but also effectively reduces the false alarm rate. It can not only meet the real time and accuracy of fire detection but also realize effective fire detection. Compared with similar fire detection algorithms, the improved algorithm proposed in this paper has better robustness to fire detection and has better performance in accuracy and speed.

1. Introduction

In the process of rapid economic development in today's cities, people's lives and property as well as the normal operation of enterprises are often threatened by fire. Fire has the characteristics of suddenness and great harm [1, 2], and there is a clearly realistic demand for fire detection and early warning. At present, there are three fire detection and alarm methods: sensor based, image processing, and depth learning.

The causes and places of fire are diverse, which hinder the fire early warning and fighting. The detection method of traditional sensors is to use a variety of sensors to detect the smoke, flame, heat, and other signals generated during fire. The purpose of distinguishing fire is achieved by analyzing different phenomena, such as the temperature sensor for detecting heat change and the smoke sensor for detecting smoke concentration in the air. These signals are processed by sensing the changes of different parameters, to detect whether there is a fire. At present, the main types of fire detectors are heat-sensing detectors, light-sensing detectors, and smoke-sensing detectors. These traditional detectors are

cheap and accurate, but they generally have some defects that are difficult to solve. For example, due to the relatively long time required for the occurrence of smog propagation and temperature rise, the response delay of traditional sensors will inevitably occur. In addition, the sensor is usually installed close to the fire point and exposed to a large amount of dust for a long time, which makes the traditional sensors vulnerable to failure. Moreover, they are especially unsuitable for fire detection in places with high fire hazards, such as tall space or outdoor scenes. The development of more effective and reliable detection methods has always been the direction of fire control efforts. Effective identification of smoke in the early stage of fire has important theoretical significance and application value [3].

The traditional detection method based on image processing is to extract the dynamic and static features of flame or smoke from video by image processing technology and then determine whether there is a fire by recognition algorithm. To make effective use of the existing video monitoring hardware resources, the research on video-based fire detection technology has important theoretical and application value. The fire prevention method based on video

monitoring is a noncontact fire detection way based on machine vision, which was especially suitable for solving fire detection problems in large space, outdoor, and other places [4]. This kind of method not only has strong anti-interference ability and fast response speed but also has the advantages of wide application range and low cost. It has become an important interdisciplinary research field in fire detection methods.

The traditional smoke identification methods usually use physical signals for monitoring. For example, Yamada [5] proposed a smoke sensor based on a layer-by-layer self-assembled electrolyte membrane for smoke perception recognition. Keller et al. [6] proposed the use of photo-acoustic sensors for flame smoke monitoring. Cheon [7] proposed the use of temperature sensors and smoke sensors for flame smoke identification. However, this method has a strong dependence on the environment. If the surrounding environment changes, the recognition accuracy will decrease sharply or even fail.

The detection method based on deep learning is to train the fire detection model using the marked fire image and then input the image to be tested into the model for recognition. The achievements of the image processing and pattern recognition method in recent years provide a new solution for smoke recognition. Yu et al. [8] used the video smoke monitoring method on the basis of optical flow and determined the moving pixels and areas in the video through the background estimation method. The Lucas-Kanade method is used to extract optical flow characteristics. Abadi et al. [9] proposed the smoke recognition using the machine vision. In this method, the Gaussian mixture model is used to extract the preselected smoke region, and the dynamic and static characteristics of smoke are obtained. Finally, support vector machine was used to train and predict the model. Compared with the existing smoke recognition technology based on physical signal, this method effectively reduces the cost and improves the accuracy and stability of recognition.

To further improve the effectiveness and accuracy of the smoke recognition method, we propose a smoke recognition method based on fast R-CNN. That is, by reducing space complexity and time complexity, the neural network training process does not need to be graded, and the efficiency and accuracy in the detection process are improved.

2. Related Work

The ever-changing shapes and colors of smoke and the difficulty in controlling the movement rule bring great challenges to the video smoke detection. Many researchers have made full use of the various properties of smoke, such as turbulence and fluttering, and devoted themselves to analyzing the essential characteristics of smoke.

The movement pattern of the target provides important information for smoke detection. Guillemant et al. [10] used grayscale embedding and other methods to generate the linked table, then extracted the moving characteristics of the target based on the table to detect whether there is smoke, and proposed a fire monitoring method. Kopilovic et al. [11]

extracted the distribution entropy in the direction of moving optical flow to explore the motion characteristics of smoke, so as to realize smoke detection. The fuzziness and fluctuation of smoke with time were studied by wavelet transform. Tung et al. [12] studied a four-level video tracking and smoke monitoring algorithm. Firstly, the approximate median method is used to detect the moving area, and then, the fuzzy c-means method is used to realize the clustering analysis of the moving area to obtain the area where smoke may appear, and then, the spatiotemporal characteristics of the area are extracted. Finally, the support vector machine (SVM) is used to analyze and judge and get the results. Yuan Feiniu et al. [13] explored a rapid detection model of translucent shielding with high-pass filtering. Through the study of smoke motion law, a fast smoke detection method integrating image and color saturation monitoring was put forward. On this basis, the elimination technology of isolated noise and chaotic motion interference is further studied.

How to extract the color and texture features of smoke in the process of smoke detection is very important. In the RGB model, the smoke gray values of the three channels are very close, mainly distributed between 80 and 220. Krstini et al. [14] compared the RGB, YCbCr, CIE Lab, and HSI color model and put forward the characteristics which reflect the smoke of the HSI model. Gubbi et al. [15] studied a smoke detection method based on wavelet transform and support vector machine. Geometric mean, inclination, arithmetic mean, kurtosis, and entropy are extracted from the obtained subimages to describe the variation characteristics of smoke. The smoke detection method based on fused images by means of image separation was proposed. The method requires the fusion image of smoke and background to be calculated, and the smoke opacity is solved by the optimization method. Yuan [16] studied a smoke monitoring method based on pyramid multiscale feature fusion. It adopted the method of regular partition of detection window to reduce the shape dependence generated by the AdaBoost method, so as to propose a robust video smoke feature.

According to different applications, the existing smoke monitoring methods based on video images are mainly divided into two categories. One type of smoke detection algorithm is combined with flame detection, which focuses more on the framework suitable for both flame detection and smoke detection. The other type focuses on smoke detection and puts forward more revealing methods to improve detection accuracy and reduce false positive rate. No matter which of the two algorithms is used, the basic detection framework can be summarized, which is divided into video image preprocessing, extraction of suspected smoke area, description of smoke characteristics, and smoke recognition.

3. Algorithm Implementation

Deep learning theory has also gained extensive attention in the research field of fire recognition. Convolution neural network (CNN) extracts image features, uses multiple convolution cores to realize the construction of low to high and local to global features, fuses feature information in the

final full connection layer, and recognizes fire images with the help of Softmax algorithm [17]. Compared with the traditional flame recognition algorithm of artificially selecting image features, this method can obtain more diverse and comprehensive features, improve the accuracy of fire recognition, and reduce the false alarm rate of the algorithm model [17, 18].

It is known from the existing research that smoke identification provides an important basis for fire early warning. Traditional machine learning and deep learning methods require more data and cannot be directly applied to smoke recognition. The delay data of smoke recognition based on fixed scene is relatively single and the model generalization ability is weak when the environment such as smoke scene is changed. Therefore, in this paper, the feature extraction layer in the pretrained VGG-16 model on the ImageNet dataset, which was also image data, is migrated to the classification task of the target dataset. The feature extraction capability of the model is migrated, so as to expand the application scope of the smoke recognition method. The feature extraction capability is contains edge feature extraction capability, texture feature extraction capability, shape, and other high-level abstract feature extraction capability.

Although there is a certain difference between ImageNet data and target smoke recognition data, there are some invariable universal features at the feature level. High-level abstract features are edges, textures, and shapes. These features are common to both ImageNet datasets and target smoke datasets. Therefore, feature transfer based on isomorphism space can be carried out.

3.1. The Model Flowchart of Transfer Learning. The definition of transfer learning is very broad. There are many aliases of transfer learning technology in relevant research, such as learning to learn, life-long learning, multitask learning, knowledge transfer, and metalearning. Multitask learning technology is most closely related to transfer learning. This technology attempts to train multiple unrelated tasks at the same time, so as to find the similarities between tasks and guide the learning of a single task according to the same characteristics. Since 2005, transfer learning has given a new connotation transfer learning to learn knowledge from one or more source tasks and then use knowledge to guide the learning of target tasks [19]. This definition further defines the purpose of transfer learning. Different from multitask learning, transfer learning only trains the source task and learns knowledge from it, rather than training the source task and target task at the same time. The following figure shows the difference between traditional machine learning and transfer learning. For different tasks, traditional machine learning needs separate training models, while transfer learning only needs training source task models, as shown in Figure 1.

Deep learning has requirements for the amount of data. In the field of fire alarm, the source of image annotation data is very limited, resulting in easy overfitting of the model and reducing the recognition effect, while transfer learning can

effectively avoid the problem of overfitting and improve the recognition accuracy when there are few datasets [20, 21]. Traditional machine learning needs a large amount of labeled training data. Without these labeled data, the trained model will perform poorly, and labeling a large amount of data requires a lot of time and manpower. The transfer learning does not need too much annotation data. It can use the knowledge or model learned in the source domain to be applied to the target domain to complete specific tasks.

The construction of a new discrimination model using transfer learning mainly includes the following aspects. Firstly, we cut the original image, enhance the data, and normalize it as the initial dataset. Secondly, divide the dataset into training set, verification set, and test set, then load the pretrained model on the ImageNet dataset, reset the full connection layer, and then put the training set in the pretrained model for training, respectively. In the training process, the model is finetuned and a new model is trained. Then, the model is evaluated and tested through the verification set and test set, and the optimal model with migration ability is selected for smoke image detection at the fire scene.

Figure 2 indicates the flowchart of transfer learning, which mainly includes four parts. First, the data are preprocessed. In the preprocessing stage, all image data are resized according to corresponding categories (unified into three channels). The size of the image is $3 \times 150 \times 150$ after random transformation such as random rotation, cutting, flipping, and normalization. Secondly, a network based on deep transfer learning is constructed. In this process, a fully connected network was pretrained using VGG-16 network within smoke dataset. The specific procession is to take the smoke dataset as the input to obtain the convolution part output in the VGG-16 network trained in ImageNet. This output is then used to train a fully connected network. Then, the reserved parameters of convolution part in the VGG network trained in ImageNet are transferred. The fully connected network of the previous pretraining is connected to obtain the learning model based on deep migration. Then, the model is trained and the parameters are adjusted, and finally, the model is predicted.

3.2. The Core Process of Fast R-CNN. In the new fast R-CNN structure, feature extraction, candidate box extraction, bounding box regression, and classification are all integrated into one network, which improves the model performance, especially the detection speed to a great extent. Fast R-CNN is mainly composed of Conv layers, Region Proposal Networks (RPN), Roi Pooling, and Classification [21].

The input objects of Fast R-CNN mainly include image samples and calibration parameters [22]. By calculating the coverage of the calibration box and object proposals for each sample image, a set of regions of interest (RoI) relative to each sample image can be known. Table 1 lists the specific composition of ROI for each image. Convolution neural network mainly uses partial convolution layer and maximum pool layer to obtain the convolution attribute of samples, which is shown in Figure 2. Then, the ROI pool

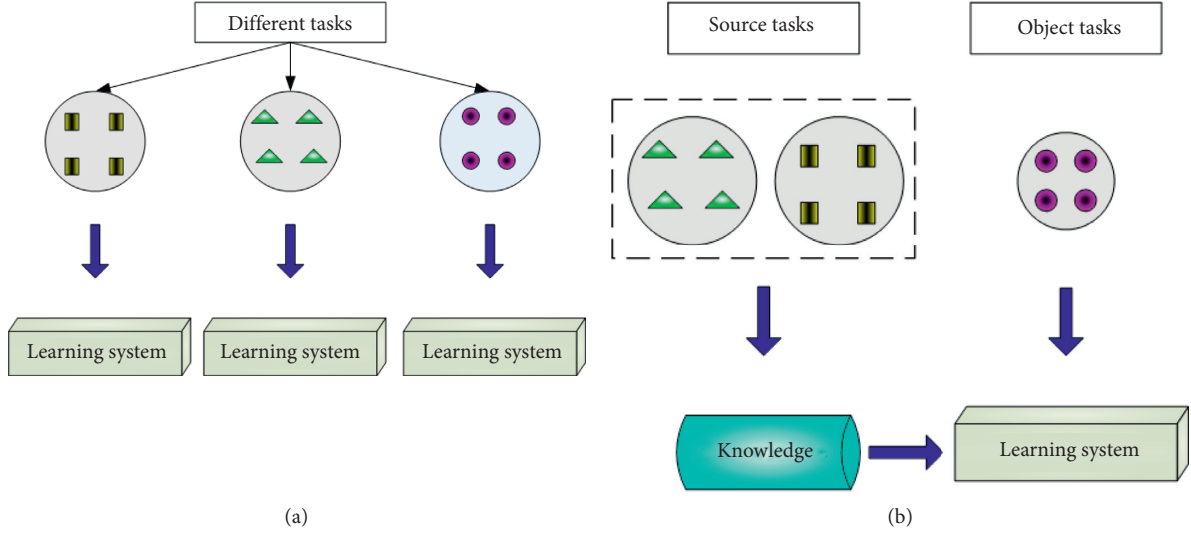


FIGURE 1: Comparison between traditional machine learning and transfer learning. (a) Learning process of traditional machine learning and (b) learning process of transfer learning.

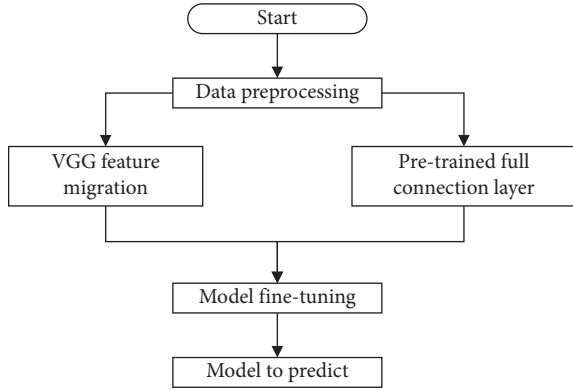


FIGURE 2: The model flowchart of transfer learning.

layer obtains the normalized feature vector through the convolution features in the region of interest.

When these eigenvectors are processed in fully connected layers, the results can be shared and enter different layers, respectively. For softmax regression calculation, the probability estimation value of class objects is calculated by one layer, and the other layer completes the data output to obtain the detection frame coordinate value of class objects on the image.

3.3. RoI Pool Layer Calculation. ROI pool layer mainly converts the ROI characteristic matrix into usable normalized characteristic matrix through the largest pool layer in fixed spatial amplitude $H \times W$. The RoI coordinates are represented by quaternions (r, c, h, w) representing the corner (r, c) and (h, w) of RoI, respectively. The RoI pool layer segments the RoI window of $h \times w$ with the subwindow of $H \times W$ size and gets the subwindow of $h/H \times w/W$ approximately. The corresponding grid output is obtained by calculating each subwindow by using pooling. With this layer, the input images that will no longer constrain the

training process must be consistent in specifications. The RoI pool layer integrates and manages RoI characteristic matrices of different sizes.

3.4. Proposed Algorithm. As shown in Figure 3, two output layers are finally obtained, and the classification results and the coordinate values of the detection box can be calculated, respectively.

The first part of the output layer is to solve the probability P of each RoI in class K , which is mainly calculated by the softmax regression method. The other part of the output layer calculates the coordinate value of the K -type detection frame. Finally, the multitask loss function is used for regression calculation of each calibrated RoI type and detection frame coordinate value as follows:

$$L(p, u, t^u, v) = L_{\text{cls}}(p, u) + \lambda [u \geq 1] L_{\text{loc}}(t^u, v), \quad (1)$$

$$L_{\text{cls}}(p, u) = -\log p_u,$$

where $[u \geq 1]$ means that when u is greater than or equal to 1, the value of the brackets is 1, and the other values are 0. L_{loc} indicates the logarithmic loss value of a certain type of probability. The other L_{loc} is detection frame coordinate loss. They are obtained according to the actual detection frame result $v = (v_x, v_y, v_w, v_h)$ of class u and the predicted coordinate value $t^u = (t_x^u, t_y^u, t_w^u, t_h^u)$ of class u , which can be expressed as follows:

$$L_{\text{loc}}(t^u, v) = \sum_{i \in \{x, y, w, h\}} s_{L1}(t_i^u - v_i), \quad (2)$$

$$S_{L1}(X) = \begin{cases} 0.5x^2 & |x| < 1 \\ |x| - 0.5 & |x| \geq 1. \end{cases}$$

In the training process, a random descent small batch is defined, which consists of N sample images as well as R RoIs. In the paper, we assume that $R = 128$ and $N = 2$. There are 2 images per batch and 64 RoIs per image.

TABLE 1: RoIs sources.

Category	Proportion (%)	Mode
1	25	IoU in the [0.5, 1] area, as a background
2	75	IoU is in the larger area of [0.1, 0.5), as the foreground

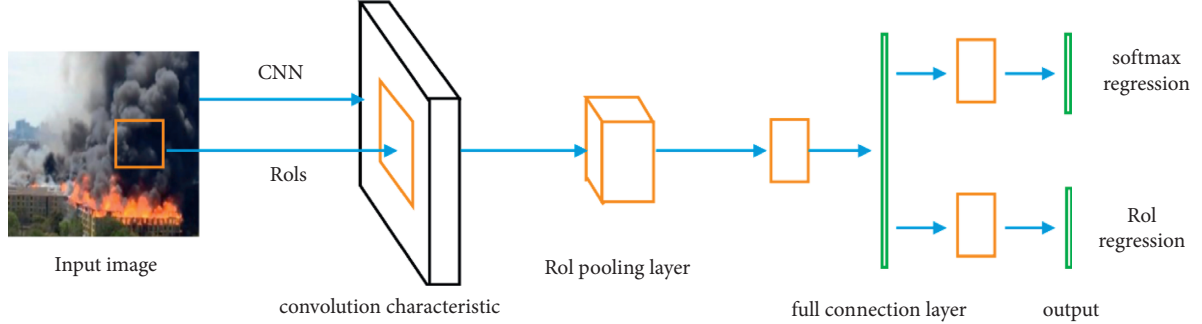


FIGURE 3: The constitution of fast R-CNN.

When passing through the RoI pool layer during the process of reverse propagation, the calculation of reverse propagation is carried out by the following formula:

$$\frac{\partial L}{\partial x_i} = \sum_r \sum_j [i = i^*(r, j)] \frac{\partial L}{\partial y_{rj}}, \quad (3)$$

where $[\cdot]$ indicates that the residual is propagated backwards in the RoI layer. It is necessary to judge whether the residual node i is connected to the maximum value of the input value of the RoI. If it is true, it accumulates the residual and the bracket value of is 1. Otherwise, the bracket value is 0.

4. Experiment and Analysis

In the paper, the common dataset and network acquisition from flame and flue gas are used as the dataset for the experiment. The experimental data mainly include the training set for model construction and the test set for model testing. The specific experimental data are listed in Table 2, while Figure 4 shows the experimental data example.

From Table 2 and Figure 4, it can be seen that the dataset used in this paper belongs to the small dataset. The ImageNet classification dataset contains 1000 categories of image data. The large amount and variety of data provide strong support for the model construction based on deep migration learning.

The methods of Tensorflow [9] and Keras [23] are used to train the fast R-CNN proposed in the paper. For comparison, we use Tensorflow and Keras to realize deep CNN. Training data are mainly completed on Intel Xeon computer and NVIDIA gtx1060 GPU.

The evaluation methods of detection rate (DR), false alarm rate (FAR), and accuracy rate (AR) are used to compare the proposed algorithm with others, which can be expressed as follows:

$$DR = \frac{P_p}{Q_p} \times 100\%,$$

$$FAR = \frac{N_p}{Q_p} \times 100\%, \quad (4)$$

$$AR = \frac{P_p + N_n}{Q_p + Q_p} \times 100\%,$$

where Q_p denotes the positive samples, P_p indicates true positive samples detected correctly, and Q_n expresses negative samples, whereas N_p is negative samples misclassified as positive. When the recognize algorithm performs good, it will get larger AR and DR and smaller FAR.

The proposed algorithm is compared with several traditional smoke detection methods, such as HLTPMC [24] and MCLBP [25]. At the same time, it is also compared with some typical deep CNNs, including AlexNet [26] and ZF-Net [27]. The comparison results obtained through the experiment are shown in Table 3.

Compared to traditional algorithms and several classical deep CNNs, the algorithm based on deep learning (fast R-CNN) is improved in detection rate and false positive rate. According to the comparison results of various indicators in Table 3, under the same test set, the algorithm in this paper is improved compared with HLTPMC, MCLBP, AlexNet, and ZF-Net. As a typical fire detection algorithm based on candidate areas, fast R-CNN will consume a lot of time when recommending candidate areas by RPN (region proposal networks). Although the overall performance of the algorithm is good, it takes a long time to detect a fire image. Although the detection time of Fast R-CNN is much shorter than that of other algorithms, it is still insufficient in the real-time requirements of fire detection.

TABLE 2: Experimental data.

Datasets	No. of smoke image	No. of no-smoke image	Total number	Purpose
Set1	350	350	700	Training
Set2	150	150	300	Test



FIGURE 4: The experimental data examples. (a) Smoke images. (b) Nonsmoke images.

TABLE 3: The experimental results are compared with other methods.

Algorithm	AR	DR	FAR
HLTPMC [24]	95.35	96.21	4.68
MCLBP [25]	96.24	95.38	1.93
AlexNet [26]	97.37	93.91	0.52
ZF-Net [27]	96.98	92.85	0.48
This paper	97.83	96.78	0.23

AR, accuracy rate; DR, detection rate; FAR, false alarm rate. Compared with other smoke detection methods, the AR, DR, and FAR obtained by this method are 97.83, 96.78, and 0.23, respectively.

The experimental results show that this algorithm not only has good detection results in the images of large fire areas but also can effectively detect small fire areas. When detecting fire-like objects, it can effectively distinguish street lamps, firefighters, and sanitation workers. The missed detection rate and false detection rate are low, and it has a good performance in fire detection.

5. Conclusion

In view of the shortcomings of existing methods in the application of fire detection, for improving the low smoke recognition rate, we proposed a Fast R-CNN smoke detection method. In this paper, a deep convolution neural network is used to extract the convolution features of visual task example images. Fast R-CNN normalization and parallel regression methods are used for calculation. Finally, the visual task-related smoke detection model was obtained. The experimental results show that the fast R-CNN smoke detection method proposed in this paper can effectively improve the detection rate and reduce the false alarm rate. Compared with similar fire detection algorithms, the improved method in this paper has better robustness to fire

detection and has good performance in both accuracy and speed. In the following work, the network will be optimized and improved to further improve the detection effect and algorithm speed. At the same time, the existing fire dataset will be expanded to increase the diversity of samples in the dataset and improve the sample quality of the training set [18–20].

Data Availability

The labeled dataset used to support the findings of this study is available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] H. Yim, S. Oh, and W. Kim, “A study on the verification scheme for electrical circuit analysis of fire hazard analysis in nuclear power plant,” *Journal of the Korean Surgical Society*, vol. 30, no. 3, pp. 114–122, 2015.
- [2] W. Wang, W. Mao, and H. E. Jianliang, “Smoke recognition based on deep transfer learning,” *Journal of Computer Applications*, 2017.
- [3] S. Miao, K. Hu, and H. Gao, “Small fire smoke region location and recognition in satellite image,” in *Proceedings of the 9th International Congress on Image and Signal Processing, Bio-medical Engineering and Informatics*, pp. 714–718, Datong, China, October 2017.
- [4] T. Tang, L. Dai, and Z. Yin, “Smoke image recognition based on local binary pattern,” in *Proceedings of the 5th International Conference on Mechatronics, Materials, Chemistry and Computer Engineering*, Chongqing, China, July 2017.
- [5] M. Yamada and S. Shiratori, “Smoke sensor using mass controlled layer-by-layer self-assembly of polyelectrolytes

- films,” *Sensors and Actuators B: Chemical*, vol. 64, no. 1, pp. 124–127, 2000.
- [6] A. Keller, M. Rüegg, and M. Forster, “Open photoacoustic sensor as smoke detector,” *Sensors and Actuators B: Chemical*, vol. 104, no. 1, pp. 1–7, 2005.
 - [7] J. Cheon, J. Lee, I. Lee, Y. Chae, Y. Yoo, and G. Han, “A single-chip CMOS smoke and temperature sensor for an intelligent fire detector,” *IEEE Sensors Journal*, vol. 9, no. 8, pp. 914–921, 2009.
 - [8] C. Yu, Y. Zhang, and J. Fang, “Video smoke recognition based on optical flow,” in *Proceedings of the 2nd International Conference on Advanced Computer Control*, vol. 10, no. 7, pp. 16–21, Shenyang, China, March 2010.
 - [9] M. Abadi, “TensorFlow: large-scale machine learning on heterogeneous distributed systems,” 2016, <https://arxiv.org/abs/1603.04467>.
 - [10] P. Guillemant and J. Vicente, “Real-time identification of smoke images by clustering motions on a fractal curve with a temporal embedding method,” *Optical Engineering*, vol. 40, no. 4, pp. 554–563, 2001.
 - [11] I. Kopilovic, B. Vagvolgyi, and T. Sziranyi, “Application of panoramic annular lens for motion analysis tasks: surveillance and smoke detection,” in *Proceedings of the 15th International Conference on Pattern Recognition. ICPR-2000*, pp. 714–717, Barcelona, Spain, September 2000.
 - [12] T. X. Tung and J.-M. Kim, “An effective four-stage smoke-detection algorithm using video images for early fire-alarm systems,” *Fire Safety Journal*, vol. 46, no. 5, pp. 276–282, 2011.
 - [13] F. N. Yuan, “Motion accumulation and translucence based model for video smoke detection,” *Journal of Data Acquisition & Processing*, vol. 22, no. 4, pp. 396–400, 2007.
 - [14] D. Krstinić, D. Stipančević, and T. Jakovčević, “Histogram-based smoke segmentation in forest fire detection system,” *Information Technology and Control*, vol. 38, no. 3, pp. 237–244, 2009.
 - [15] J. Gubbi, S. Marusic, and M. Palaniswami, “Smoke detection in video using wavelets and support vector machines,” *Fire Safety Journal*, vol. 44, no. 8, pp. 1110–1115, 2009.
 - [16] F. Yuan, “Video-based smoke detection with histogram sequence of LBP and LBPV pyramids,” *Fire Safety Journal*, vol. 46, no. 3, pp. 132–139, 2011.
 - [17] A. J. Dunnings and T. P. Breckon, “Experimentally defined convolutional neural network architecture variants for non-temporal real-time fire detection,” in *Proceedings of the 2018 25th IEEE International Conference on Image Processing (ICIP)*, pp. 1558–1562, Athens, Greece, October 2018.
 - [18] P. Barmpoutis, K. Dimitropoulos, K. Kaza, and N. Grammalidis, “Fire detection from images using faster R-CNN and multidimensional texture analysis,” in *Proceedings of the 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 8301–8305, Brighton, UK, May 2019.
 - [19] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, “How transferable are features in deep neural networks,” *Advances in Neural Information Processing Systems*, pp. 3320–3328, 2014.
 - [20] S. J. Pan and Q. Yang, “A survey on transfer learning,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, 2010.
 - [21] H.-C. Shin, H. R. Roth, M. Gao et al., “Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning,” *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1285–1298, 2016.
 - [22] S. C. Hsu, Y. W. Wang, and C. L. Huang, “Human object identification for human-robot interaction by using fast R-CNN,” in *Proceedings of the Second IEEE International Conference on Robotic Computing*, pp. 201–204, Laguna Hills, CA, USA, January 2018.
 - [23] 2015 <https://keras.io/F.Chollet.Keras>. [Online].
 - [24] F. Yuan, J. Shi, X. Xia, Y. Fang, Z. Fang, and T. Mei, “High-order local ternary patterns with locality preserving projection for smoke detection and image classification,” *Information Sciences*, vol. 372, pp. 225–240, 2016.
 - [25] S. R. Dubey, S. K. Singh, and R. K. Singh, “Multichannel decoded local binary patterns for content-based image retrieval,” *IEEE Transactions on Image Processing*, vol. 25, no. 9, pp. 4018–4032, 2016.
 - [26] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” *Proceedings of the Advances in Neural Information Processing Systems*, vol. 25, pp. 1097–1105, 2012.
 - [27] M. D. Zeiler and R. Fergus, “Visualizing and understanding convolutional networks,” in *Proceedings of the European Conference on Computer Vision*, pp. 818–833, Zurich, Switzerland, September 2014.