Journal of Thermal Analysis and Calorimetry (2023) 148:5089–5095
https://doi.org/10.1007/s10973-022-11657-1

Development and optimization of image fire detection on deep learning algorithms

Yi Yang[1] · Mengyi Pan[1] · Pu Li[2] · Xuefeng Wang[1] · Yun-Ting Tsai[3]

Received: 16 May 2022 / Accepted: 19 September 2022 / Published online: 16 October 2022

Abstract
The primary function of fire detection is to detect fires and raise the alarm early. A detection algorithm is a key element of image fire detection (IFD) technology because it directly determines the IFD's performance. In this study, an IFD algorithm based on the YOLOv3 network was developed to detect smoke and flame simultaneously. Subsequently, six improvements were applied to promote the algorithm's ability to detect fire early. The results demonstrated that the modified YOLOv3 network achieved an average accuracy of 95%, which is 14.1% higher than that of the same model without modifications. The detection speed reached 22 Frames Per Second (FPS), which satisfies the requirements of real-time detection.

Keywords Image fire detection · YOLOv3 network · Detection ability · Average accuracy · Detection speed

Abbreviations

IFD	Image fire detection
FPS	Frames per second
DL	Deep learning
CNN	Convolutional neural network
NMS	Nonmaximum suppression
SGD	Stochastic gradient descent
IOU	Intersection-over-union
CBL	Conv + Bn + Leaky_relu
Avg IOU	Average intersection-over-union
AP	Average precision

Introduction

Fire prevention is becoming increasingly challenging because of accelerating urbanization and the continuous growth in building size. The quick and accurate detection of fire can effectively reduce fire losses. Most traditional detection algorithms use simple models, such as shallow convolutional neural network and Support Vector Machine; in addition, the complex environments in fire images could affect the performance of algorithm on detection. Therefore, it is difficult for traditional detection algorithms to detect small object proportions on images. However, the new image fire detection (IFD) technology can be used to automatically distinguish the characteristics of fire or smoke in a fire image by using a digital image processing method. IFD is not limited by space, height, air velocity, or dust, and it is a noncontact technology, thus avoiding some of the restrictions in traditional fire detection.

Currently, the development of detection algorithms is a research focus. As early as 1966, Foo [1] mentioned the application of brightness information in IFD. Subsequently, studies of and developments in detection algorithms have focused on fire image features [2–5]. However, the traditional detection algorithm artificially extracted fire image features, which exhibited weak generalization, a high false-positive rate, and low practicability. Therefore, deep learning (DL) algorithms, such as an advanced image classification convolutional neural network (CNN), were introduced to solve these problems. The common CNNs [6] included AlexNet [7], VGG [8], Inception [9], ResNet [10], a smoke detection algorithm, and a flame detection algorithm. Mao et al. [11] introduced time-series information to improve the accuracy of algorithms in IFD, and Namozov [12] used a modified VGG-Net to detect smoke and flame simultaneously. Dung

✉ Yi Yang
yangyi_610@163.com

Yun-Ting Tsai
tsaiyt@xjtu.edu.cn

1 School of College of Safety Science and Engineering, Xi'an University of Science and Technology, Xi'an 710054, China

2 Zhengzhou Airport Economy Zone Fire Brigade, Zhengzhou 450000, China

3 School of Chemical Engineering and Technology, Xi'an Jiaotong University, Xi'an 710049, China

et al. [13] used mixture-of-Gaussians background modeling to cluster and discriminate the background and foreground and then applied cascade classification to determine the candidate region. Zhong et al. [14] used a color model to determine the candidate region and then AlexNet to improve the flame detection in the candidate region, and Li et al. [15] studied object detection CNN in IFD. The results revealed that YOLOv3 provided the most suitable method for IFD.

An IFD algorithm based on DL experiences two problems: first, image classification is the focus of the detection algorithm, and it therefore lacks the ability to extract candidate regions, limiting its early fire detection ability. Second the part of algorithms uses the objective detection methods for fire detection. This only applies transfer learning to develop an algorithm, and the fire detection performance of the algorithm is not optimized, which is lacking practicability [16–27].

This study aimed to develop an IFD with a stronger early fire detection capability. To achieve this, a modified YOLOv3 algorithm was developed with six improvements: (1) Addition of images containing small object proportions; (2) data enhancement; (3) addition of a backend object detection network feature map; (4) improvements to the backend object detection network structure; (5) improvements to the anchorpoint setup; (6) improvements to the nonmaximum suppression (NMS).

Algorithm development and optimization

Development of the algorithm

The computer used in the study was an Intel Core I7-7700 CPU @ 3.6 GHZ, 16 GB DDR4 RAM 2400 MHz with a NVIDIA Titan X Pascal GPU with 3840 CUDA. The operating system was Ubuntu 16.0.4. The data set consisted of 29,180 images (13,400 fire images and 15,780 nonfire images) with various scenarios obtained from Li et al. [15] The fire image data set has been divided into development subsets and test subsets by using the min-Hash approximate image replacement method.

YOLOv3 network

YOLOv3 was used to design and generate a fire detection network. YOLOv3 is a network of object detection CNNs that can develop the image fire detection algorithm. It was trained using Microsoft's COCO data set (a large-scale detection, segmentation, and captioning data set) and then retrained using transfer learning [28, 29]. The YOLOv3 network consisted of a frontend feature extraction network and a backend classification network. The frontend was frozen in transfer learning, and the backend was trained and optimized through the training and verification of the fire image data set obtained from Li et al.

Stochastic gradient descent (SGD) was used to update the parameters. The batch size was set to 64, the SGD momentum to 0.9, and the intersection-over-union (IOU) threshold to 0.6. The NMS method was used to determine the number of candidate boxes as 300, the initial learning rate as 0.001, and the total number of iterations as 200 K. The learning rate was reduced by 10 when the number of iterations reached 120 K and 160 K. The other parameters remained set to the original network.

Improvement of the algorithm

The test set in Li et al.'s fire image data set was used to evaluate the reliability of the trained YOLOv3 algorithm. According to the results, the average accuracy was 84.5% and the detection speed reached 28 Frames Per Second (FPS), which represents a high detection level. However, in the early stages of fire development, the proportion of flame and smoke in an image is less than 20%; in a large building, the proportion is less than 1%. Therefore, an object proportion of less than 20% was used to evaluate the average accuracy of YOLOv3, which was only 80.95%.

An error analysis was conducted to achieve an object proportion of less than 20%. When the confidence threshold reached 0.5, 100 missed detection samples from the test set were randomly selected to evaluate YOLOv3's missed detection rate. The missed detection samples with the different object proportions are shown in Table 1, which demonstrates that lower object proportions are associated with higher missed detection rates. Therefore, improved detection ability for lower object proportions is required.

Six improvements for YOLOv3

Addition of images containing small object proportions

The performance of algorithms depends on the design of the network architecture and the selection of suitable data sets. If the development data set differs considerably from the actual scenario in real-time detection, the network architecture is not able to achieve its desired effect. Therefore, the

Table 1 The missed detection samples with the different object proportions

Object proportions/%	Missed detection samples
(0, 5)	55
[5, 10]	23
[10, 15)	10
[15, 20)	12

development data set should be consistent with real-life scenarios. 1,231 fire images containing small object proportions were added to the original development set to achieve this. Here, small object means that the object is proportion small in an image and difficult to be detected by algorithm (Fig. 1).

Data enhancement

Data enhancement is an effective method to expand data samples and improve an algorithm's generalization ability and robustness. The shooting angle, pixel size, brightness, and other factors can cause differences between images of the same scene. Therefore, in this study, data enhancement was used to transform the original training data. Subsequently, the transformed data were used to train the neural network to improve the detection ability of the algorithm for different scenes (Fig. 2).

Addition of the backend object detection network feature map

The YOLOv3 through 8 times downsampled feature map to detect small objectives. But, when the size of objectives is less than 8×8 pixels, the algorithm has difficult to detect it. Therefore, addition the object detection network feature map of backend to improve the ability on detecting small objectives. The improvements in the YOLOv3 network structure achieved by the $4 \times$ downsampled feature map (Fig. 3).

Improvements to the backend object detection network structure

A previous study [30] used a residual unit to improve feature learning efficiency and reduce gradient dispersion. In the present study, a similar improvement was achieved by changing the original five CBL units in the convolutional block unit of the YOLOv3 network structure to two residual block units and one CBL unit (Fig. 4).

Improvements to the anchorpoint setup

The K-means clustering method was used to obtain a new proposal for the region size of the box for the fire image data set, thereby reducing the complexity of box regression in the next step. The average intersection-over-union (Avg IOU) Eq. (1) serves as the cluster analysis metrics for determining the optimal value of K:

$$I = avg \max \frac{\sum_{i=1}^{k} \sum_{j=1}^{nk} IOU(B, C)}{n} \tag{1}$$

where B denotes the cluster sample box, C denotes the center of the cluster, k denotes the number of cluster centers, n_k denotes the number of samples in the kth cluster center, n is the total number of samples, and IOU (B, C) denotes the intersection ratio of the central box and the sample box in the cluster (Fig. 5).

Fig. 1 The 1,231 fire images containing small object proportions were added to the original development set

(a) Original image | (b) Contrast enhancement | (c) Gaussian noise | (d) Brightness transform | (e) Motion blur

Fig. 2 The examples of images transformed using different methods of data enhancement. These images were produced with the addition of object proportions of less than 20% from the original development set

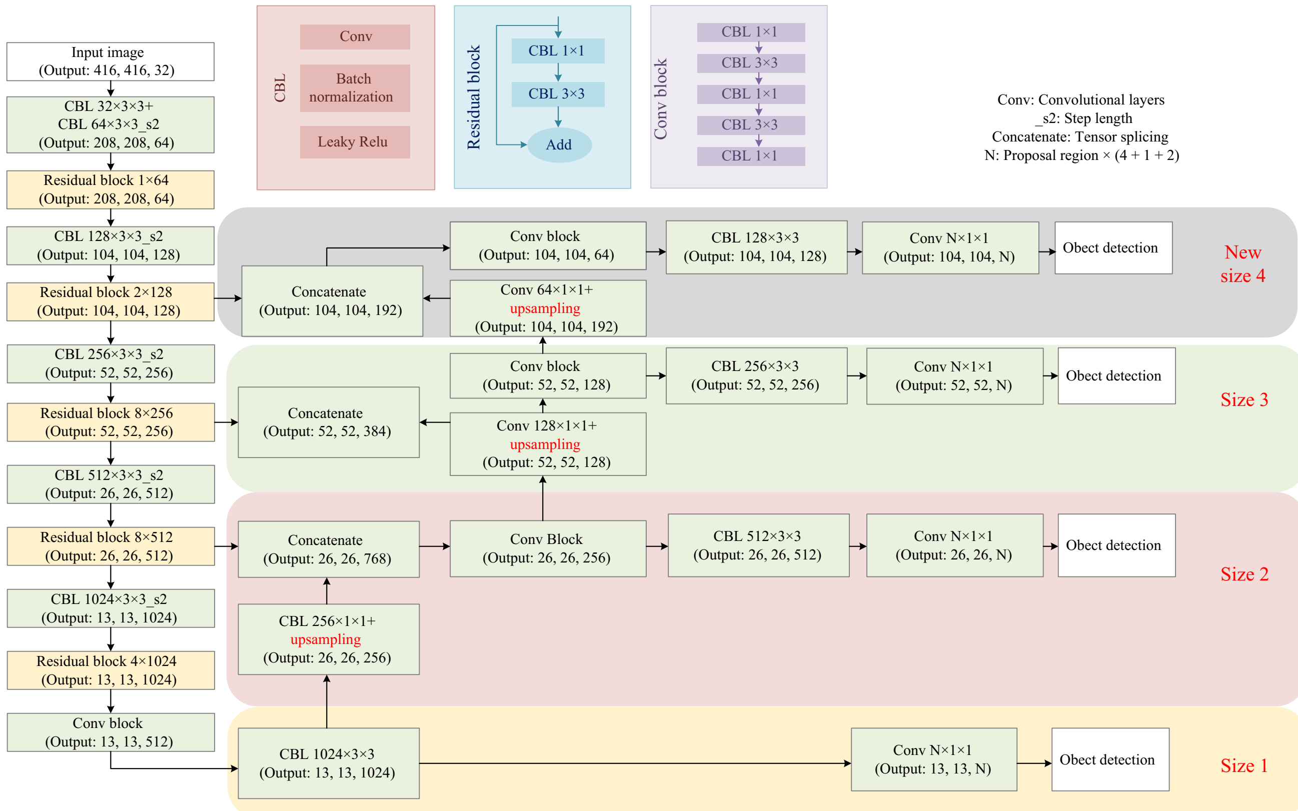

Fig. 3 An 8 downsampled feature map was produced through the YOLOv3 network. Subsequently, 2 upsampling was conducted combined with the second group of residual blocks in the frontend Darknet-53 feature extraction network to obtain a 4 downsampled feature map

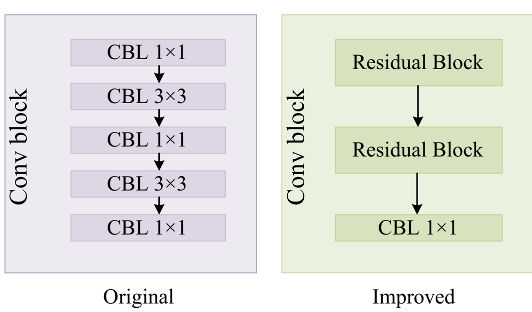

Fig. 4 The improvements of Conv Block unit. Conv Block unit consists of 5 Conv + Bn + Leaky_relu (CBL) unites that original 5 CBL units in the Conv Block unit were changed to 2 Residual Block units and 1 CBL unit

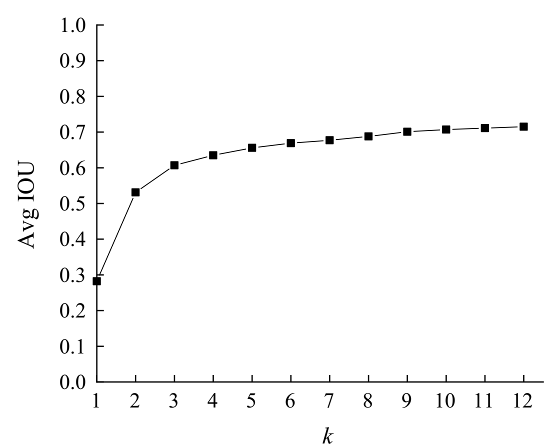

Fig. 5 The Avg IOU against different K values using a cluster analysis. When $k \geq 10$, the Avg IOU was stable. The central box could therefore be generated using a cluster analysis at $k = 10$, which indicated an improvement in the region proposal scheme

A cluster analysis of the object box in the fire image data set was conducted using $K = 1$–12. The Avg IOU before and after improvement is displayed in Table 2, and the improved region proposal is listed in Table 3. The region proposal IOU increased to 9.2%, indicating that, as a result of these improvements, the detection of small object proportions in large-scale feature maps can acquire more region proposals.

Improvements to the NMS

The early stages of intensive image sampling may generate multiple region proposals for each image position, and therefore, the same objects can be predicted by multiple overlapping boxes. Equation (2) uses the NMS method to solve this problem, enabling the YOLOv3 algorithm to filtrate the output of prediction boxes:

$$s_{\text{confidence}} = \begin{cases} s_{\text{confidence}}, & IOU(M,b) < I \\ 0, & IOU(M,b) > I \end{cases} \tag{2}$$

where $S_{\text{confidence}}$ refers to the confidence of the prediction box, M refers to the prediction box with the maximum confidence in the box list, b refers to the prediction box that compares with M, IOU (M, b) refers to the intersection ratio of two boxes, and I refers to the IOU threshold.

However, if the IOU value of a box is greater than the threshold value, this box might be deleted. Therefore, when flame and smoke overlap in the images, the NMS can decrease the average detection accuracy. Bodla et al. [30] noted an improvement in the YOLOv3 algorithm using a soft-NMS method, as displayed in Eq. (3):

$$s_{\text{confidence}} = \begin{cases} s_{\text{confidence}}, & IOU(M,b) < I \\ s_{\text{confidence}}(1-IOU(M,b)), & IOU(M,b) > I \end{cases} \tag{3}$$

Thus, when the IOU is larger, the confidence is lower. This method also decreased the probability of a box being removed, thereby improving the detection ability when flame and smoke overlap.

Evaluation of the algorithm's performance

To evaluate the algorithm's performance, the six improvements were used to develop different models separately, and a combination of all the improvements was used to obtain the modified YOLOv3. These seven models were comparable in terms of their algorithm performance. Table 4 presents the various models with their corresponding improvements.

Average precision (AP) was used to evaluate the detection ability of the different models. Table 5 lists the AP (fire and smoke individually), mAP (average AP of fire and smoke together), and detection speed calculated using the different models. According to the results, the AP increased in all the improved models and the detection speed decreased in the modified model, which satisfied the requirement of a detection speed of ≥ 20 FPS. Because the flame features

Table 2 Avg IOU before and after improvement

Region proposal schemes	Original schemes	Improvement schemes
Size	373×326, 156×198, 116×90, 59×119, 62×45, 30×61, 33×23, 6×30, 10×13	316×195, 206×257, 126×92, 106×122, 63×58, 34×78, 37×55, 45×23, 22×25, 10×16
Number	9	10

Table 3 The improved region proposal

Feature map	13×13	26×26	52×52	104×104
Receptive field	Big	Medium	Smaller	Small
Detection the object	Big	Medium	Smaller	Small
Region proposal	316×195, 206×257	126×92, 106×122	63×58, 34×78, 37×55	45×23, 22×25, 10×16

Table 4 Model design scheme

Optimization strategy		Model
Original	No improvement	YOLOv3
Improvement of development data	Addition of images containing small object proportions	YOLOv3_a
	Data enhancement	YOLOv3_b
Improvement of algorithm design	Addition of a backend object detection network feature map	YOLOv3_c
	Improvements to the backend object detection network structure	YOLOv3_d
	Improvements to the anchorpoint setup	YOLOv3_e
	Improvements to the NMS	YOLOv3_f
Combination of all improvement methods		modified YOLOv3

Table 5 Evaluation results of model

Model	AP/% Smoke	AP/% Fire	mAP/%	Detection speed (FPS)
YOLOv3	78.3	83.5	80.9	28
YOLOv3_a	83.6	87.9	85.8	28
YOLOv3_b	87.2	91.5	89.4	28
YOLOv3_c	91.7	93.8	92.8	24
YOLOv3_d	86.9	90.4	88.7	27
YOLOv3_e	90.5	92.7	91.6	28
YOLOv3_f	82.6	86.1	84.4	26
YOLOv3_modified	93.8	96.1	95.0	22

were clearer than the smoke features, the AP for detecting smoke was lower than that for detecting flame.

Furthermore, although the YOLOv3_b model was only optimized using data, its AP was still higher than that of original YOLOv3, indicating that improvements in development data are important for promoting algorithm performance.

An evaluation of the model's performance was conducted after the improvements to the algorithm design had been completed. The results revealed that the addition of the backend object detection network feature map and improvements to the anchorpoint setup increased the AP to 11.9% and 10.7%, respectively. The AP of the modified YOLOv3 model reached 95%, which was 14.1% higher than that of the original model. The detection speed of the modified YOLOv3 reached 22 FPS, which satisfied the real-time detection requirements.

Conclusions

This study provides an effective and reliable method for detecting smoke and flame in the early stages of a fire using images. The procedure for developing the model has been clearly described as well as the six improvements for promoting YOLOv3's detection ability and speed and decreasing the missed detection rate. The AP of the modified YOLOv3 reached 95%, which was 14.1% higher than that of original model, and the detection speed satisfied the requirements of real-time detection. This model can be used to develop IFD technology for real-life situations and decrease the risk of fire losses.

The purpose of this study was to develop and optimize an image fire detection algorithm on deep learning. Six improvements were applied to promote the algorithm's ability to detect fire early. These results were confirmed through the performance of algorithm. In the future study, we can consider the complex situations on real environment, thereby enhancing the detection ability of IFD.

Acknowledgements This project was supported by National Natural Science Foundation of China (Grant no. 51904229), the Natural Science Basic Research Program of Shaanxi (Grant no. 2020JQ-753). We would like to extend our sincere appreciation to the organizations.

Author contributions All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by YY, MY P, and P L. The first draft of the manuscript was written by MY P. The writing-review and editing was worked by X W and YT T. All authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

References

1. Horng WB, Peng JW, Chen CY. A new image–based real–time flame detection method using color analysis. 2005.
2. Ko BC. Wildfire smoke detection using temporospatial features and random forest classifiers. Opt Eng. 2012;51:7208. https://doi.org/10.1117/1.OE.51.1.017208.
3. Tung TX, Kim JM. An effective four-stage smoke-detection algorithm using video images for early fire-alarm systems. Fire Saf J. 2011;46:276–82. https://doi.org/10.1016/j.firesaf.2011.03.003.
4. Wei Z, Wang X, An W, Che J. Target-tracking based early fire smoke detection in video. 2009 Fifth International Conference on Image and Graphics. 2009:172–6. https://doi.org/10.1109/ICIG.2009.173.
5. Hu Y, Lu X. Real-time video fire smoke detection by utilizing spatial-temporal ConvNet features. Multimedia Tools and Applications. 2018;77:29283–301. https://doi.org/10.1007/s11042-018-5978-5.
6. Muhammad K, Ahmad J, Mehmood I, Rho S, Baik SW. Convolutional neural networks based fire detection in surveillance videos. IEEE Access. 2018;6:18174–83.
7. Krizhevsky A, Sutskever I, Hinton G. Image net classification with deep convolutional neural networks. Adv Neural Inform Process Syst. 2012;60(6):84–90.
8. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. Comput Sci 2014. https://doi.org/10.48550/arXiv.1409.1556.
9. Filonenko A, Kurnianggoro L, Jo KH. Comparative study of modern convolutional neural networks for smoke detection on image data// International conference on human system interactions. IEEE. 2017:64–8. https://doi.org/10.1109/HSI.2017.8004998.
10. C Tao, Jian Z, Pan W. Smoke detection based on deep convolutional neural networks// 2016 International Conference on Industrial Informatics - Computing Technology, Intelligent Technology, Industrial Information Integration (ICIICII). IEEE. 2016. https://doi.org/10.1109/ICIICII.2016.0045.
11. Mao W, Wang W, Dou Z, Li Y. Correction to: fire recognition based on multi-channel convolutional neural network. Fire Technol. 2018;54:809. https://doi.org/10.1007/s10694-018-0705-3.
12. Namozov A, Cho YI. An efficient deep learning algorithm for fire and smoke detection with limited data. Adv Electr Comput Eng. 2018;18:121–8.
13. Dung NM, Kim D, Ro S. A video smoke detection algorithm based on cascade classification and deep learning. KSII Trans Internet Inf Syst. 2018;12:6018–33.
14. Zhong Z, Wang M, Shi Y, Gao W. A convolutional neural network-based flame detection method in video sequence. Signal Image Video Process. 2018;8:1619–27.

15. Li P, Zhao W. Image fire detection algorithms based on convolutional neural networks. Case Stud Therm Eng. 2020;19:100625.
16. Xiao Y, Liu J, Zeng J, Lu X, Tian Y, Shu CM. Coupling effect of operational factors on heat extraction from a coal pile using a two-phase closed thermosyphon. Energy. 2022;239:122371.
17. Wang W, Huang Y, Hu S, Su W, Pan Y, Shu CM. Thermal hazards analysis for benzoyl peroxide in the presence of hexanoic acid. Process Saf Environ Prot. 2022;157:208–17.
18. Liu K, Xiao Y, Zhang H, Pang P, Shu CM. Inhibiting effects of carbonised and oxidised powders treated with ionic liquids on spontaneous combustion. Process Saf Environ Prot. 2022;157:237–45.
19. Song J, Deng J, Zhao J, Zhang Y, Shu CM. Comparative analysis of exothermic behaviour of fresh and weathered coal during low-temperature oxidation. Fuel. 2021;289:119942.
20. Li H, Zhai F, Li S, Lou R, Wang F, Chen X, Shu CM, Yu M. Macromorphological features and formation mechanism of particulate residues from methane/air/coal dust gas–solid two-phase hybrid explosions: An approach for material evidence analysis in accident investigation. Fuel. 2022;315:123209.
21. Tsai YT, Huang GT, Zhao JQ, Shu CM. Dust cloud explosion characteristics and mechanisms in MgH2-based hydrogen storage materials. AIChE J. 2021;67:e17302.
22. Tsai YT, Yang Y, Huang HC, Shu CM. Inhibitory effects of three chemical dust suppressants on nitrocellulose dust cloud explosion. AIChE J. 2020;66:e16888.
23. Yang Y, Ding L, Xu L, Tsai YT. Effect of metal chloride on thermal decomposition of nitrocellulose. Case Stud Therm Eng. 2021;28:101667.
24. Tsai YT, Tao F, Zhou Q. Explosion characteristics and suppression of hybrid Mg/H2 mixtures. Int J Hydrog Energy. 2021;46:38934–43.
25. Yang N, Jiang J, Huang AC, Tang Y, Li Z, Cui J, Shu CM, Xing Z. Thermokinetic model-based experimental and numerical investigation of the thermal hazards of nitrification waste. J Loss Prev Process Ind. 2022;79:104836.
26. Huang AC, Liao F, Huang C, Zhang TY, Shu YCM, Xing Z, Jiang J, Hsieh W. Calorimetric approach to establishing thermokinetics for cosmeceutical benzoyl peroxides containing metal ions. J Therm Anal Calorim. 2021;144:373–82.
27. Chen W, Chen W, You M, Tsai YT, Shu CM. Evaluation of thermal decomposition phenomenon for 1,1-bis(tert-butylperoxy)-3,3,5-trimethylcyclohexane by DSC and VSP2. J Therm Anal Calorim. 2015;122:1112–33.
28. Redmon J, Farhadi A. YOLOv3: An Incremental Improvement. arXiv e-prints. 2018;1804:02767. https://doi.org/10.48550/arXiv.1804.02767.
29. Ju L. Wang, Hui, Chang, The Application of Improved YOLO V3 in Multi-Scale Target Detection. Appl Sci. 2019;9:3775.
30. Bodla N, Singh B, Chellappa R, Davis LS. Soft-NMS -- Improving object detection with one line of code. 2017:5562–70. https://doi.org/10.1109/ICCV.2017.593.