Multiple Categories Of Visual Smoke Detection Database

1st Yafei Gong
Faculty of Information Technology, Beijing University of Technology
Engineering Research Center of Intelligent Perception and Autonomous Control, Ministry of Education Beijing Laboratory of Smart Environmental Protection
Key Laboratory of Intelligent Control and Optimization for Industrial Equipment of Ministry of Education, Dalian University of Technology
Beijing, China
GongYaFei@emails.bjut.edu.cn

2nd Xuanchao Ma
Faculty of Information Technology, Beijing University of Technology
Engineering Research Center of Intelligent Perception and Autonomous Control, Ministry of Education Beijing Laboratory of Smart Environmental Protection
Key Laboratory of Intelligent Control and Optimization for Industrial Equipment of Ministry of Education, Dalian University of Technology
Beijing, China
m1614936559@163.com

3rd Hongyan Liu
Faculty of Information Technology, Beijing University of Technology
Engineering Research Center of Intelligent Perception and Autonomous Control, Ministry of Education Beijing Laboratory of Smart Environmental Protection
Key Laboratory of Intelligent Control and Optimization for Industrial Equipment of Ministry of Education, Dalian University of Technology
Beijing, China
liuh9221@gmail.com

4th Ke Gu
Faculty of Information Technology, Beijing University of Technology
Engineering Research Center of Intelligent Perception and Autonomous Control, Ministry of Education Beijing Laboratory of Smart Environmental Protection
Key Laboratory of Intelligent Control and Optimization for Industrial Equipment of Ministry of Education, Dalian University of Technology
Beijing, China
guke@bjut.edu.cn

Abstract—Smoke detection has become a significant task in associated industries due to the close relationship between the petrochemical industry’s smoke emission and its safety production and environmental damage. There are several production situations in the real industrial production environment, including complete combustion of exhaust gas, inadequate combustion of exhaust gas, direct emission of exhaust gas, etc. We discovered that the datasets used in previous research work could only determine whether smoke is present or not, not its type. That is, the dataset’s category does not map to real-world production situations, which are not conducive to the precise regulation of the production system. In order to reduce the gap between the algorithm and the actual application so that the new algorithm can more comprehensively cover and solve the actual situations, we created a multi-categories smoke detection database that includes a total of 70196 images. We further conduct the experiment by employing multiple models on the proposed database. The results demonstrate the effectiveness of the proposed database and show that the performance of the current algorithms needs to be improved.

Index Terms—database, smoke detection, Convolutional neural network

This work was supported in part by the Key Laboratory of Intelligent Control and Optimization for Industrial Equipment of Ministry of Education, Dalian University of Technology under Grant LICOCO222TB03; in part by the CAAI-Huawei Mindspore Open Fund under Grant CAAIXSJLJJ-2021-016A.

I. INTRODUCTION

Petrochemical companies are the primary energy producers in China, and smoke detection at petrochemical companies has long drawn considerable interest. Exhaust gas emitted by the venting flare must be entirely burnt to guarantee the long-term normal operation of the system and efficiently eliminate environmental pollution. The leading existing solution to reduce exhaust gas pollution emissions is to inject combustion-supporting steam to promote the complete combustion of exhaust gas [1]. With the proper amount of combustion-supporting steam, exhaust gases can be completely burned to reduce pollution. If the amount of combustion-supporting steam is inadequate, the smoke produced by incomplete combustion will pollute the air. However, when the amount of combustion-supporting steam is excessive, it not only results in a waste of resources but also obstructs the exhaust gas combustion, resulting in the direct release of the exhaust gas into the atmosphere. Therefore, the key to the highly effective combustion of exhaust gas is the suitable regulation of combustion-supporting steam. Current approaches rely on sensors and human adjustment. Still, due to their unique features, it is impossible to ensure exhaust gas’s effective...
and efficient combustion using these two methods [2]–[4]. With the advancement of artificial intelligence, image-based technologies may offer novel solutions to smoke detection.

The smoke color can directly reflect the state of combustion of exhaust gas in the image. During the combustion of exhaust gas, the color of smoke will vary according to the degree of combustion and will also represent the industrial environment’s operating conditions. For instance, in petrochemical plants and refineries, the smokeless state indicates that the system is operating normally. When black smoke is detected, the system is in abnormal operating, meaning there is insufficient combustion of the exhaust gases. The higher the carbon content, the darker the hue, and the larger the environmental harm, the blacker the smoke. The white smoke indicates excessive combustion-supporting steam, and the exhaust gas is released into the atmosphere, seriously polluting the environment. Consequently, when black smoke or smokeless is discovered, it may cause significant safety dangers and environmental issues.

Particulate matter detection [5]–[10] and image-based smoke detection methods have been the subject of extensive research [11]–[16] to reduce pollution. As far as we know, the available methods with their data sets can only determine whether smoke is present or absent. They cannot determine the type of smoke [17]–[21]. Furthermore, existing industrial control systems can only make fundamental decisions based on limited smoke detection results. The new algorithms developed using these datasets cannot provide additional information to assist the control systems in producing more precise regulation. As a result, creating a database with multiple categories significantly impacts resolving the engineering issues mentioned above. Therefore, we proposed and developed a multi-classification database for smoke detection in this paper with a total of 70196 image patches.

The remaining structure of this paper is organized as follows. The process for creating the dataset is described in depth in Section 2. The effectiveness of the database and the performance of the algorithm are examined in Section 3. The main conclusion is offered in Section 4 at the end.

II. DATABASE

During the combustion of exhaust gases, three types of smoke are often produced: white smoke, black smoke, and smokeless, which is also the most common form of smoke. Similar findings have been seen in other industrial domains, such as thermal power plants. As a result, we designed a new three-categories smoke detection database (TCSSD) for smoke detection based on [16] to continue the research on the engineering difficulties mentioned above. The dataset in [16] was used to determine the presence or absence of smoke in an image. It contains 24217 images, including 5695 smoke images and 18522 smokeless images. The dataset we proposed can be constructed by processing the original dataset. The specific process is described in the remainder of this section.

The proposed dataset includes a training set, a validation set, and a testing set following standard dataset design practices. In the process of designing the database, it is found that the color of smoke in an image does not precisely match pure black or white. The subjective scoring technique is utilized to identify the categories of an image according to the color of the smoke. This study uses 20 data from 20 different sources to determine an image’s label to minimize the learning error brought on by inaccurate labels. All images are displayed on an HP monitor with resolution of 2560 x 1440. Each participant needs to understand the content of an image and then determine the specific categories of the image, including smokeless, black smoke, and white smoke. A short pre-exercise is conducted to help the participant understand the calibration criteria: the image content and color. The labeling process runs in the environment under moderate illumination and low noise conditions. Each participant would take five minutes break every thirty minutes to avoid unnecessary errors due to visual fatigue. Once the labels are acquired, choosing a final label from a range of possible results can be equivalent to the hard voting process in ensemble learning. To be specific, let $X = [x_1, x_2, \ldots, x_n]$, where $n$ is the number of categories and $x_i$ is the number of votes an image has received to be classified in a specific category ($n$ equals three in this paper). The sum of $x_i$ should also equal $M$, where $M$ is the total number of votes cast for an image, which is 20. The following is the image’s final category:

$$x = \text{argmax}(X)$$ (1)

argmax denotes that the final category of the image is obtained from the subscript of the value that causes $X$ to obtain the maximum value.

Another issue discovered while building the database is the stark variation in the number of images in each category. The network’s performance depends on the distribution of samples among the categories [21]. Specifically, there are 8363 smokeless images, 1423 black smoke images, and 778 white smoke images in the validation set that was initially partitioned, compared to 8511 smokeless images, 1423 black smoke images, and 778 white smoke images in the original training set. As a result, data augmentation is needed for images of white and black smoke. The color change enhancement has significant interference with the smoke image, which should be considered when developing the specific algorithm. Therefore, the rotation operation is mainly employed to augment data. The rotation angle that is applied to the image can be calculated by $360/N \times I$, where $N$ is the multiple of the data increase and $I = [1, 2, \ldots, N]$. The rotation operation with $n$ equals six for the black smoke images is utilized. For the white smoke images, the rotation operation with $n$ equals 12 is used, which closely balances the number of images in the three categories. By utilizing rotation, the total number of images in the three categories is almost equal.

The finished dataset is as follows: The training set consists of 26538 images, comprising 9336 image patches of white smoke, 8511 image patches of smokeless, and 8538 image patches of black smoke. A total of 26483 image patches, 8363 smokeless image patches, 8928 black smoke image patches,
Fig. 1. Examples of data, the patches from left to right are the original image and the image enhanced by rotation.

and 9192 white smoke image patches make up the validation set. There are a total of 17328 image patches in the testing set, including 9888 image patches of smokeless patches, 4644 image patches of black smoke, and 2796 image patches of white smoke. The images in the dataset are displayed in Fig. 1, and it is essential to note that only the images from the training and validation sets are rotated.

III. EXPERIMENT

In this section, we will further examine the performance of existing state-of-the-art general classification networks and networks designed specifically for smoke detection and assess this database’s usefulness.

A. Experimental Protocol

**Operating Environment.** All models used in this paper have the same configuration, allowing for a fair comparison of model performance. Our experimental framework is PyTorch. An Ubuntu computer running the Inter(R) Gold 6248R CPU at 3.00GHz and an NVIDIA GeForce RTX 3090 graphics card power the experimental environment. All models used CrossEntropyLoss and SGD optimizer with epoch equal to 200 and batch size equal to 128.

**Competing Models.** The comparative experiments are conducted on our database with a total of seven start-of-the-art models, and those models are Alex-Net [22], VGG-Net [23], Res-Net [24], Google-Net [25], Mobile-Net [26], Shuffle-Net [27], DCNN [16]. The first four networks have seen significant advancements in public image classification datasets in recent years and have been widely employed. MobileNet and ShuffleNet are created for real-world industrial uses. They are simpler to apply in the industry, having great precision and quick calculating speeds. DCNN is developed by Gu et al. specifically for smoke detection.

**Evaluation Criteria.** For quantifying the performance of each model, The loss, precision, recall, and accuracy of the validation set and testing set, as well as the recall and precision of each category, were recorded. We could analyze the model’s behavior and assess the database’s efficacy from these indicators. The following are the definitions of these indicators:

\[
\text{accuracy} = \frac{TP + FP}{TP + TN + FP + FN} \\
\text{precision} = \frac{TP}{TP + FP} \\
\text{recall} = \frac{TP}{TP + FN} \\
\text{f1score} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}
\]

TP, TN, FP, and FN represent True positive, True negative, False Positive, and False negative, respectively. In the binary classification task, recall and precision are typically calculated for the positive and negative samples. We determine the recall and precision for each category, then average the results to get the overall recall and precision. A good model is expected to achieve greater values on these indicators as far as possible.

B. Performance Comparison

On the TCSDD database, we record and compare the performance of those mentioned seven start-of-the-art models using accuracy, recall, precision, and f1-score. Loss is also employed as a measure of the robustness and generalization of networks. In this study, the loss of a model is determined by averaging the total loss that results from adding the losses of all the samples in the testing set. The performance results of the network are shown in table I.

**TABLE I**

| Networks  | AlexNet | ResNet | Vgg | GoogleNet | ShuffleNet | MobileNet | DCNN |
|-----------|---------|--------|-----|-----------|------------|-----------|------|
| ACC       | 0.928   | 0.923  | 0.941 | 0.954     | 0.922      | 0.942     | 0.935 |
| R         | 0.928   | 0.922  | 0.934 | 0.949     | 0.928      | 0.9421    | 0.926 |
| P         | 0.907   | 0.908  | 0.929 | 0.943     | 0.904      | 0.926     | 0.926 |
| F1        | 0.918   | 0.915  | 0.932 | 0.946     | 0.916      | 0.933     | 0.926 |
| loss      | 0.328   | 0.380  | 0.895 | 0.191     | 0.343      | 0.276     | 0.358 |

Accuracy, precision, recall, and the f1-score are each denoted by the letters Acc, R, P, and F1 in table I. It can be shown from table I that GoogleNet performs best, with an accuracy of 0.954, a recall of 0.949, and a precision of 0.943. Additionally, we discover that all models’ precision is lower than recall, which indicates that the model’s precision needs to be increased. In real-world applications, we not only need to be able to identify particular scenarios accurately, but we also don’t want the model to become perplexed by various scenes. Therefore, it is necessary to work on increasing precision while keeping recall. For the indicators
of each category, the experimental results demonstrate that the smokeless category has a higher recall and precision than the white smoke and black smoke categories, which is consistent with prior research’s findings and suggests that the model does a good job of differentiating between smoke- and non-smoke-filled environments. In other words, it can detect whether smoke is present in the image in the binary classification test. When employing TCSDD, the main errors occurred in black and white smoke categories. A lot of work has been put into building the database to prevent importing inaccurate label information as much as possible. The results in table I demonstrate that there are additional causes for the uncertainty, like backdoor confusion in the image or color confusion brought on by changes in illumination. The models should further mine features unaffected by environmental noise and brightness variations.

According to the above content, a basic ranking of the models may be determined based on the metrics, and the orders are as follows: GoogleNet > MobileNet > VGG ≈ DCNN > AlexNet ≈ ResNet18 ≈ ShuffleNet. GoogleNet achieves the best performance, which may be related to its multi-scale information extraction mechanism. VGG lags behind Mobilenet because its loss on the test set is substantially higher than Mobilenet’s, suggesting that its generalization performance might be slightly worse. Mobilenet also shows excellent capabilities, increasing the potential for practical industrial applications. In addition, DCNN also offers great detection capabilities.

IV. CONCLUSION

Smoke detection is essential to reducing air pollution and ensuring safe production. In this paper, we create a three-categories smoke database dedicated to petrochemical smoke detection. First, the image labels are obtained by subjective scoring, and then data augmentation is performed by rotation to solve the problem of category imbalance. The final generated dataset contains 70,196 image patches. We have run a performance competition on the proposed database with some state-of-the-art models, and the results show that performance still needs to be improved. In our future work, we will strive to leverage self-supervision and other techniques to mine smoke features to enhance the performance of fully-supervised tasks for the challenge of distinguishing different forms of smoke.

REFERENCES

[1] E. A. Emam, "Environmental pollution and measurement of gas flaring," Int. j. innov. res. sci. eng. technol, vol. 2, pp. 252-262, 2016.
[2] G. Kokkullnik, G. Gonca, V. Ayhan, I. Cesar, A. Parlak. "Theoretical and experimental investigation of diesel engine with steam injection system on performance and emission parameters," Applied Thermal Engineering, vol. 54(1), pp. 161-170, 2013.
[3] R. Srinivasarao, K. M. Krishna. "Automatic control of soot and unburnt hydro carbons from flares in oil and gas industry," In 2014 International Conference and Utility Exhibition on Green Energy for Sustainable Development (ICUE), pp. 1-5, 2014.
[4] M. R. Rahimpour, Z. Jamshidnejad, S. M. Jokar, G. Karimi, A. Ghorbani, A. H. Mohammadi. "A comparative study of three different methods for flare gas recovery of Asalouye Gas Refinery," Journal of Natural Gas Science and Engineering, vol. 4, pp. 17-28, 2012.