Smoke100k: A Database for Smoke Detection

Hsiang-Yin Cheng*, Jia-Li Yin*, Bo-Hao Chen*, and Zhi-Min Yu†
*Department of Computer Science and Engineering
Yuan Ze University, Taoyuan 320, Taiwan
Email: bhchen@saturn.yzu.edu.tw
† College of Mathmatics and Computer Science
Fuzhou University, Fuzhou 350116, China
Email: 228815616@qq.com

***Abstract*—Due to the complex scenarios and the limited feature information in a single image, a precise smoke detection is much more challenging in practice. Most of previous smoke detection methods either extract textural and spatiotemporal characteristics of smoke or separate the smoke and background components of the image. However, those methods often fail in detecting smoke positions because of the limited feature information within a single image. Moreover, the task of smoke detection can be better achieved if the extra information from collected training dataset is available. One key issue is how to build a training dataset of paired smoke images and ground-truth bounding box positions for end-to-end learning. This paper proposes a large-scale benchmark image dataset to train a smoke detector. With the built dataset, experimental results demonstrate that the discriminative models can be effectively trained as the smoke detector to detect the smoldering fires precisely.**

***Keywords*—smoke detection, object detection, deep learning.**

I. Introduction

Detecting the smoldering fires with good protection from flaming fires is an essential goal of fire alarm system in most commercial and industrial facilities. The smoke alarms, however, are sometimes failed to activate in manufactured housing because of outdoor working conditions and the limited sensing range of alarm device. The resulting chances of a fire occurring will not only damage the property, but also cause harm to people from fire [1]. With recent advances in computer vision driven by deep learning, the smoke detection from ubiquitous surveillance cameras has become more viable to improve people's safety relative to fire incidents and make the fire alarm system more reliable. To detect the exact position of smoke in an image, traditional smoke detection techniques include those feature extraction-based algorithms [1], which extract textural and spatiotemporal characteristics of smoke from an image by using image descriptors, and image decomposition-based algorithms [2], which separate the smoke and background components of the image based on the morphological component analysis (MCA). These approaches, however, are difficult to achieve a precise smoke detection due to the complex scenarios and the limited feature information in a single image.

Inspired by the recent success of data-driven methods, especially the deep convolutional neural networks (CNNs) [3] in object detection, this work makes the first attempt to deal with this challenging problem. Compared with traditional

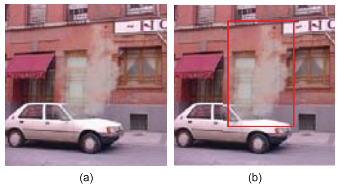

Fig. 1. Example of (a) smoke image, and (b) the corresponding image with bounding box around the smoke.

smoke detection techniques which use several images to learn image priors for smoke localization, the data-driven methods utilize a set of smoke images and the corresponding ground-truth bounding box positions to learn a discriminative model to localize the smoke from the given image, as an example shown in Fig. 1. In the case of smoke detection, unfortunately, it is very difficult to generate such paired data since the lack of a simple model to simulate the smoke image generation process. To the best of our knowledge, by far there is no dataset of paired smoke images and ground-truth bounding box positions available for training a discriminative model for smoke detection. Therefore, one significant contribution of this study is that we build such a dataset "Smoke100k" of paired smoke images and ground-truth bounding box positions, which makes the discriminative learning of smoke detectors possible.

II. Proposed Smoke Image Dataset

The Smoke100k dataset consists of 100k synthesized smoke images $\mathbf{x}_i$, corresponding smoke-free images $\mathbf{y}_i$, smoke masks $\mathbf{z}_i$, and bounding box positions $\mathbf{p}_i$. Additionally, the smoke-free images $\mathbf{y}_i$ were collected from LabelMe dataset [4], NYU dataset [5] and normalized into the size of $224 \times 224 \times 3$.

To build this dataset, we manually make 920 base masks via PhotoShop to simulate the gaseous behavior of rising smoke for synthesizing the smoke masks, where each base mask is composed of different densities of smoke. As can be seen in Fig. 2 (a)-(c), the densities of smoke in the smoke masks are divided into three levels: low, middle, and high densities. The amount of base masks in each level is indicated in Table I. Additionally, the base masks are added in smoke-free images by different angles and shapes, by which to enlarge the variation. This can be expressed as follows:

$$\mathbf{x}_{i,c} = \mathbf{y}_{i,c}\mathbf{z}_{i,\alpha} + (1 - \mathbf{z}_{i,\alpha})\mathbf{z}_{i,c}, \text{ s.t. } c \in \{R, G, B\} \quad (1)$$

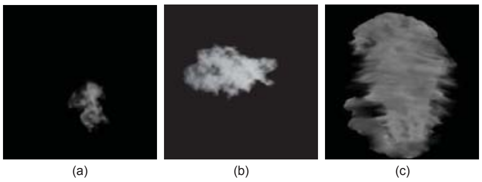

Fig. 2. Illustration of smoke masks with different density levels of smoke. (a) Smoke mask of low density level. (b) Smoke mask of middle density level. (c) Smoke mask of high density level.

TABLE I
AMOUNT OF BASE MASKS IN SMOKE100K DATASET

Density	Low	Middle	High
# base masks	220	400	300

TABLE II
AMOUNT OF SYNTHESIZED SMOKE IMAGES IN EACH SUBSET OF SMOKE100K DATASET

Subset	Smoke100k-L	Smoke100k-M	Smoke100k-H
# Training images	22k	24k	21k
# Testing images	11k	12k	12k

where $\mathbf{x}_{i,c}$, $\mathbf{y}_{i,c}$, and $\mathbf{z}_{i,c}$ denote the ith synthesized smoke image, smoke-free image, and smoke mask, respectively; and $\mathbf{z}_{i,\alpha}$ denotes the alpha channel for transparency of smoke mask. By doing so, the ground-truth bounding box positions $\mathbf{p}_i$ can be also acquired in the generation process.

Hence, there are three subsets of synthesized smoke images $\mathbf{x}_i$ for simulation of different smoldering fires. Three subsets are detailed as follows:

1) Smoke100k-L: samples are synthesized by smoke masks selected from the Low level with twenty kinds of angles.
2) Smoke100k-M: samples are synthesized by smoke masks selected from the Middle level with eight kinds of angles.
3) Smoke100k-H: samples are synthesized by smoke masks selected from the High level with fifteen kinds of angles.

The amount of synthesized smoky samples in each level are indicated in Table II. Note that we also append an additional random factor of shape to further increase the variation for generation process in each level.

III. EXPERIMENTS

We evaluate the state-of-the-art CNN detectors [3], i.e., MobileNetV2 and ResNet50, using transfer learning on the built Smoke100k dataset. Adam optimizer with a batch size of 25 patches is employed in training process, where the learning rate, β_1, and β_2 are set to 0.001, 0.9, and 0.999, respectively. Figures 3-5 show the detection results on three smoke images. The ground-truth bounding box positions in Smoke100k dataset are provided in Figs. 3-5 (a). These data-driven methods are able to extract enough information from Smoke100k dataset, and the detection results by MobileNetV2 and ResNet50 detectors produce adequate accuracy indicated by Intersection over Union (IoU) [3] higher than 79%, as shown in Figs. 3-5 (b) and (c).

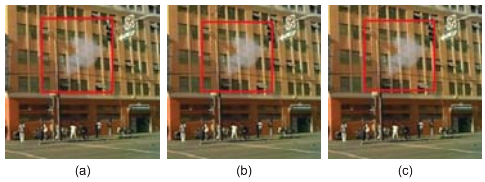

Fig. 3. Smoke detection results on Smoke100k-L dataset. (a) Original. Predicted results detected by using (b) MobileNetV2 (IoU = 84%), and (c) ResNet50 (IoU = 85%).

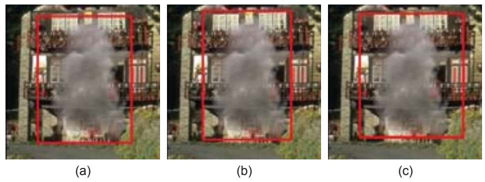

Fig. 4. Smoke detection results on Smoke100k-M dataset. (a) Original. Predicted results detected by using (b) MobileNetV2 (IoU = 79%), and (c) ResNet50 (IoU = 81%).

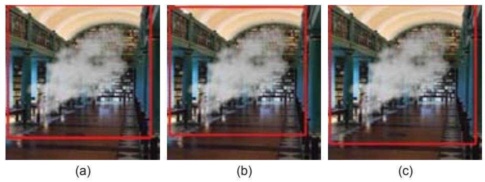

Fig. 5. Smoke detection results on Smoke100k-H dataset. (a) Original. Predicted results detected by using (b) MobileNetV2 (IoU = 88%), and (c) ResNet50 (IoU = 92%).

IV. CONCLUSIONS

This paper proposes a large-scale benchmark image dataset called Smoke100k for training a smoke detector. The Smoke100k dataset contains 100k synthesized smoke images, corresponding smoke-free images, smoke masks, and bounding box positions. Experiment results show that the proposed Smoke100k dataset is able to be exploited to train a discriminative model for more accurate smoke detection.

ACKNOWLEDGMENT

This work was supported by the Ministry of Science and Technology, Taiwan, under Grant Nos. MOST 108-2221-E-155-034-MY3, and MOST 107-2221-E-155-052-MY2.

REFERENCES

[1] K. Dimitropoulos, P. Barmpoutis, and N. Grammalidis, "Higher order linear dynamical systems for smoke detection in video surveillance applications," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 27, no. 5, pp. 1143–1154, May 2017.

[2] H. Tian, W. Li, P. O. Ogunbona, and L. Wang, "Detection and separation of smoke from single image frames," *IEEE Transactions on Image Processing*, vol. 27, no. 3, pp. 1164–1177, March 2018.

[3] Z. Zhao, P. Zheng, S. Xu, and X. Wu, "Object detection with deep learning: A review," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–21, 2019.

[4] B. C. Russell, A. Torralba, K. P. Murphy, and W. T. Freeman, "Labelme: A database and web-based tool for image annotation," *International Journal of Computer Vision*, vol. 77, no. 1, pp. 157–173, May 2008. [Online]. Available: https://doi.org/10.1007/s11263-007-0090-8

[5] P. K. Nathan Silberman, Derek Hoiem and R. Fergus, "Indoor segmentation and support inference from rgbd images," in *ECCV*, 2012.