A Deep Learning Approach to Reduce False Alarms for Optical Smoke Detectors

Ming Liu, Hongbin Zhou, Yupeng Ren and Wei Lu
Zhejiang Dahua Technology Co., Ltd. 1199 Binan Road, Binjiang District, Hangzhou, Zhejiang Province, China
Email: zhbb@zju.edu.cn

Abstract. Optical smoke detectors (OSDs) are fire-fighting equipment used to detect fire by detecting smoke with scattering phenomenon. From the principle of OSDs we can see that they are vulnerable to false alarms caused by dust or water steam. To reduce false alarms and to make OSDs more reliable, we present a deep learning approach to train a classifier to distinguish fire event from non-fire ones based on time series data. The classifier is modelled with a 1-Dimension convolutional neural network, and generative adversarial network is used to augment and balance training data. Experiment shows that our classifier can reduce more than 50% false alarms caused by water steam while maintaining sensitivity for fire events.

Keywords. Optical smoke detector; false alarms; deep learning; GAN.

1. Introduction
The optical smoke detectors (OSDs) are fire-fighting equipment used to detect fire by detecting smoke produced by material burning. Commercial OSDs, which have dominated the market, are mainly composed of light source, light sensor, optical chamber, data processing circuit and shell. The optical chamber interrupts the light sensor from the light source so that the light sensor does not have significant reading. While smoke particles produced by fire get a way into the optical chamber, these particles scatter the light beam from light source to light sensor, which produces an electrical signal. The intensity of the signal indicates the consistency of smoke particles in the air. In other words, the OSDs identify fire event by detect the scattering phenomenon in optical chamber.

However, there are other particles in the air may cause scattering phenomenon besides smoke particles, like dust and water steam, which are quite common in our living environment. A naive threshold based classifier cannot distinguish signal caused by smoke from signal caused by dust or water steam, which makes OSDs extremely vulnerable to false alarms. It is of great importance to investigate a method to reduce false alarms and to make OSDs more reliable.

The core issue leads to false alarms lies in the lack of information gathered by OSDs. An intuitive solution is to increase the source of information by adding more sensors of variable kinds, with the help of multi-sensor fusion technology OSDs can be more robust to disturb. However integrating more sensors means higher manufacturing cost, which is fatal to the popularization of commercial OSDs. An alternative way is to expand the information in time dimension. Compare to a single data point of particle consistency sampled by OSDs, a curve recording how particle consistency changes in a period of time contains much more information. The problem becomes how to extract useful information from time series data.

In this paper we present a deep learning approach to reduce false alarms for OSDs. A 1-Dimension convolutional neural network (CNN) is designed and trained as a classifier to extract features from the...
sampled curve and decide whether or not the curve comes from a fire event. CNN training is notoriously data consuming, nevertheless it is expansive to sample and label data with OSDs, to make things work we apply generative adversarial network (GAN) as a data augmentation method. The experiment shows that the CNN classifier can reduce more than 50% false alarms caused by water steam while maintaining its sensitivity for fire events. The main contributions of this work can be summarized as follows.

- We propose a pipeline for OSDs false alarms reduction which filters more than 50% false alarms caused by water steam.
- We design a 1-Dimension CNN which can work on a small embedded system to extract information from OSDs data.
- We use a conditional Wasserstein GAN with gradient penalty to expand OSDs dataset with high quality artificial data.

2. Related Work

The subject of our work intersects with three domains of research, namely OSDs optimization, CNN designing & training, and GAN based data augmentation.

2.1. OSDs Optimization

Kitsak et al. [1] use laser as light rather than LED and the light is split into two coherent beams, under the effect of scattering and interference, a time-varying signal can be received by the light sensor, which contains enough information to distinguish false alarms from fire events. To et al. [2] use a hand designed feature extraction method to convert the time series into a gradient histogram, and dissimilarity measure like Euclidean Distance is used to classifier these histograms. Chen et al. [3] add gas sensor to the smoke sensor system to identify false alarms. Cheon et al. [4] add temperature sensor to the smoke sensor system, which not only sense heat from a fire but also compensate the smoke detector. Aspey et al. [5] use a polychromatic LED to produce light with variable wavelength, and spectral analysis is applied to uniquely identify the combustion material. Cashdollar et al. [6] use a white light source and three light sensors sensitive to different wavelength to get more information of the scattering particles, from which the size and mass of particles can be measured. Wang et al. [7] combine temperature sensor, infrared sensor and OSDs together to get more information, and use BP neural network to identify false alarm.

2.2. CNN Designing & Training

CNN is a successful architecture in a variety of domains like computer vision task, natural language processing task, speech recognition task, and automatic translation task. LeCun et al. [8] design a network to recognize handwrite digits, this network scanning the input image with a single neuron that has a local receptive, such a network is known as CNN and is much more expressive and efficient than conventional fully connected neural network. Nair et al. [9] introduce ReLU activation function into deep learning, which is faster and less vulnerable to gradient vanishing than conventional Sigmoid activation function. Sergey et al. [10] propose Batch Normalization to deal with the covariate shift problem, which may be responsible for network losing learning capability, Batch Normalization makes deep CNNs more trainable. Diederik et al. [11] introduce Adam optimizer, which has become one of the most popular optimization algorithms in deep learning community.

2.3. GAN Based Data Augmentation

Goodfellow et al. [12] propose GAN which is an elegant architecture that contains two parts, a Generator and a Discriminator, the Generator tries to generate something ‘real’ but not exist while the Discriminator tries to distinguish real data from generated ones. These two networks progress in an adversarial way, until the Generator can generate real enough data. Ever since the original GAN is proposed, it becomes a hot point in deep learning research, a bunch of works have been done in this domain. Arjovskiy et al. [13] propose WGAN by introducing Wasserstein Distance to replace the JS
divergence objective function in original GAN. The advantage of WGAN is that it’s free from mode collapse and more stable in training, however WGAN uses Weight Clipping to guarantee that Lipschitz constraint is met, which is ‘a clearly terrible way’ according to the authors and they encourage further investigation on the Lipschitz constraint. Gulrajani et al. [14] introduce a soft way to deal with Lipschitz constraint by penalizing the norm of gradient, the so called WGAN-GP algorithm is more robust and performs better than WGAN. Mirza et al. [15] propose cGAN which can generate data belongs to a given label. Frid-Adar et al. [16] use GANs as a data augmentation method to improve the performance of liver lesion classifier.

3. Method

3.1. Data Collection
We collect OSDs data according to the following rules. The sampling frequency is adaptively adjusted according to the current sample value, higher sample value leads to higher sampling frequency. We store the sample value it in a first-in-first-out queue, whose capacity is set to be 24, large enough to cover the intensity ascending process. An event is defined by sample value larger than the pre-set threshold 3 times in a row. When an event takes place, 4 extra data points will be sampled and then the whole queue, as illustrated in figure 1, will be saved along with whether it’s a fire event or non-fire event as a piece of complete data. We sample data in 7 scenes, namely beech stick smouldering, cotton rope smouldering, polyurethane foam (without flame retardant) burning, n-heptane/toluene mixture burning, joss stick smouldering, smoke chamber testing, and water steam blast. The first 6 scenes belong to fire event, and the last scene belongs to non-fire event, more detail to be found in table 1.

![Figure 1. A curve of particle consistency.](image)

Table 1. Number of data collected in different scenes.

| Scene                   | Data number | Label |
|-------------------------|-------------|-------|
| beech stick             | 72          | fire  |
| cotton rope             | 72          | fire  |
| polyurethane foam       | 72          | fire  |
| n-heptane/toluene       | 72          | fire  |
| joss stick              | 432         | fire  |
| smoke chamber           | 288         | fire  |
| water steam             | 316         | non-fire |
3.2. Data Augmentation
Data augmentation is of great importance in deep learning. Popular data augmentation methods in computer vision, like rotation [17] or flipping [18], don’t make sense on OSDs data. Meanwhile OSDs data is sampled at a low frequency so that the augmentation cannot be done in frequent domain like [19] do in speech recognition. To make things work, we use two kinds of augmentation, namely Gaussian noise and GAN.

As presented in figure 2, our GAN combines the features of cGAN [15] and WGAN-GP [14]. The Generator takes a piece of noise which sampled from a Gaussian distribution and a label as input, and output a piece of data that matches the given label. While the Discriminator take a piece of data as input, and output the prediction of label and Wasserstein distance. Loss of Discriminator consists cross entropy loss, adversarial loss, and gradient penalty. Loss of Generator consists cross entropy loss and adversarial loss. It is worth noting that in this architecture Generator and Discriminator cooperate together to avoid ambiguity in classification.

We use Adam algorithm [11] for both Generator and Discriminator with learning rate of 2e-4. To balance the process of Generator and Discriminator, we update these two networks with a ratio of 1:5. A weight of 10 is assigned to gradient penalty and 1 for other losses. We train the GAN with all of our real data for 1e4 epochs. Figure 3 shows some data generated by the Generator, and compared with real data we collected in variant scenes.

3.3. CNN Classifier
To deploy our classifier on commercial OSDs which have strictly limited storage and computing resources, we have to restrict the depth and width of our network. We build our classifier with 4 convolution blocks, which contains 1-Dimension convolutional layer, ReLU layer, and batch normalization layer. The max width of the network is set to be 2 as show in figure 4. The classifier takes a vector of length 24 as input, and output a scalar which indicates the probability of the input data belongs to a fire event.

We use Adam algorithm [11] with learning rate of 3e-4 to train the classifier. Real data is split into two subsets, namely train set and validate set, in a ratio of 4:1. In mini batch, the ratio of real data versus artificial data is 1:1. For artificial data generation, we sample random labels from uniform distribution to balance positive and negative samples, and sample random noise from Gaussian distribution. Binary cross entropy loss is used to guide the training.

4. Experiment
To verify the performance of our classifier, we make controlled experiments in same scenes as we collect training data. The first four experiments are made in a fire laboratory following Chinese national standard GB 4715-2005. Joss stick is lit in a pipe to assemble the smoke together and lead it to testing OSDs. Smoke chamber is a standard equipment to test OSDs as stipulated in GB 4715-2005. Water steam experiment is made in a box with an electric stove to heat water, when the lid is removed, a burst of water steam blast pounds the testing OSDs mounted on the box ceiling, as showed in figure 5.

![Figure 2. The architecture of our GAN in augmentation](image-url)
Figure 3. Real data versus artificial data: row a is real data from fire event, row b is artificial data from fire event, row c is real data from non-fire event, and row d is artificial data from non-fire event.

Figure 4. The architecture of our CNN classifier.

The experiments result is recorded in table 2, the recording 6/15 means we repeat that experiment 15 times and in 6 times the testing OSD gives an alarm. The first 6 scenes are fire event, and we expect the OSDs to give alarms, both OSDs pass these tests with 100% alert rate. In the scene of water steam, which we expect the OSDs to keep silence, the OSD without our classifier experiences 13 false alarms while the one with our classifier has 6 false alarms, which means more than 50% false alarms are filtered by our classifier.

From the false alarm records of our classifier, we find that in some cases, the time series curve of water steam blast appears to be a step signal, which is quite similar to the curve of joss stick. Under the principle of “Zero Leakage Alarm”, we prefer to give alarm for all the step signals. Such a step signal phenomenon is due to a large amount of particles get into the optical chamber in a short time,
say about a few sample periods, and light sensor reaches saturation. A higher sample frequency may relieve the problem, however we should also balance performance and battery life time. More investigation will be made in the aspect.

| Scene              | w    | w/o  |
|--------------------|------|------|
| beech stick        | 2/2  | 2/2  |
| cotton rope        | 2/2  | 2/2  |
| polyurethane foam  | 2/2  | 2/2  |
| n-heptane/toluene  | 2/2  | 2/2  |
| joss stick         | 12/12| 9/9  |
| smoke chamber      | 2/2  | 2/2  |
| water steam        | 6/15 | 13/15|

Figure 5. The steam source we used is not large enough to trigger OSDs in an empty room, so we use a box to concentrate steam. An OSD with our classifier and a common OSD are mounted on the box ceiling side by side. The OSDs is about 50cm above the electric stove we used to produce steam. To exclude the influence of mounting location, we exchange the place of these two OSDs during experiment frequently. After the experiments condensation is found on the inner surface of optical chamber.

5. Conclusions
In this work we present a deep learning approach to reduce false alarms for OSDs. We use a conditional Wasserstein GAN with gradient penalty to augment and balance training data. A compact and efficient 1-Dimension CNN classifier is designed and trained to distinguish fire event from non-fire event. Our experiments demonstrate that our classifier can filter more than 50% false alarms caused by water steam.

References
[1] Kitsak A, Esipovich D and Poliakov V 2011 Increase of noise immunity of laser smoke fire detectors Szkola Głowna Służby Pożarniczej.
[2] To H and Fong N 2013 Investigation of the performance and improvement of optical smoke detectors Procedia Engineering 62 985-993.
[3] Chen S J, Hovde D C, Peterson K A and Marshall A 2007 Fire detection using smoke and gas sensors Fire Safety Journal 42 (8) 507-515.
[4] Cheon J, Lee J, Lee I, Chae Y, Yoo Y and Han G 2009 A single-chip CMOS smoke and temperature sensor for an intelligent fire detector IEEE Sensors Journal 9 (8) 914-921.
[5] Aspey R A, Brazier K J and Spencer J W 2005 Multiwavelength sensing of smoke using a polychromatic LED: Mie extinction characterization using HLS analysis IEEE Sensors Journal 5 (5) 1050-1056.
[6] Cashdollar K L, Lee C K and Singer J M 1979 Three-wavelength light transmission technique to measure smoke particle size and concentration Applied Optics 18 (11) 1763-1769.
[7] Wang W, Jiang B and Yang L 2014 Design on fire alarm system based on bp neural network and multiple source information fusion International Conference on Mechatronics.
[8] LeCun Y, Boser B, Denker J S, Henderson D and Jackel L D 1990 Handwritten digit recognition with a back-propagation network Advances in Neural Information Processing Systems pp 396-404.

[9] Nair V and Hinton C E 2010 Rectified linear units improve restricted Boltzmann machines Proc. of the 27th International Conference on Machine Learning pp 807-814.

[10] Ioffe S and Szegedy C 2015 Batch normalization: Accelerating deep network training by reducing internal covariate shift.

[11] Kingma D and Ba J 2014 Adam: A Method for Stochastic Optimization Computer science.

[12] Goodfellow I J, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A and Bengio Y 2014 Generative adversarial nets Advances in Neural Information Processing Systems 3 2672-2680.

[13] Arjovsky M, Chintala S and Bottou L 2017 Wasserstein GAN Neural Information Processing Systems Foundation 12 5768-5778.

[14] Gulrajani I, Ahmed F, Arjovsky M, Dumoulin V and Courville A 2017 Improved training of Wasserstein GANs.

[15] Mirza M and Osindero S 2014 Conditional generative adversarial nets arXiv preprint arXiv:1411.1784.

[16] Frid-Adar M, Klang E, Amitai M, Goldberger J and Greenspan H 2018 Synthetic data augmentation using GAN for improved liver lesion classification 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018).

[17] Xie S and Tu Z 2017 Holistically-Nested Edge Detection International Journal of Computer Vision 125 (5) 3-18.

[18] Yang H and Patras I 2015 Mirror, mirror on the wall, tell me, is the error small? 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

[19] Park D S, Chan W, Zhang Y, Chiu C C, Zoph B, Cubuk E D and Le Q V 2019 SpecAugment: A simple data augmentation method for automatic speech recognition Interspeech.