HCHODetector: Formaldehyde concentration detection based on deep learning

Zhihao Cao1, Mingfeng Shao1, Aiju Shi2, and Hongchun Qu1

1College of Information Science and Engineering, Zaozhuang University, Zaozhuang, Shandong Province, 277160, China
2College of chemistry and Materials Science, Shandong Agricultural University, Taian, Shandong Province, 271018, China
*Corresponding author’s e-mail: czh@sdau.edu.cn

Abstract. Currently, deep learning technology is developing rapidly. Deep learning is mainly used in the fields of vision and hearing for human beings, but less in the field of olfactory. Formaldehyde is a common gas harmful to human health. However, the traditional methods of Formaldehyde concentration detection are inefficient in some cases. As for this problem, this paper proposes a novel formaldehyde detector namely HCHODetector. Specifically, this detector is based on deep learning and HSV colour space augmentation. Moreover, we propose a novel Mask-guided module and a novel pre-training network to enhance the colour discrimination ability of HCHODetector. As a consequence, the experimental results show that the detection error is within 0.08 mg/m³ in the actual environment, which provides a new idea for Formaldehyde concentration detection.

1. Introduction

Deep learning is an emerging area of machine learning research. Currently, deep learning technology is developing rapidly in artificial intelligence [1]. It comprises multiple hidden layers of artificial neural networks. Deep learning is mainly used in the fields of vision and hearing for human beings, such as image processing [2] and speech processing [3], but less in the field of olfactory, such as gas detection.

Formaldehyde is a kind of irritating gas, which has many uses in industry. The commonly used boards, paints, carpets, wallpaper and so on in interior decoration contain and release formaldehyde. High indoor concentration of formaldehyde will lead to formaldehyde poisoning. Formaldehyde poisoning can cause eye tears, skin allergy, nasopharynx discomfort, cough, acute and chronic bronchitis and other respiratory diseases, as well as nausea, vomiting, gastrointestinal dysfunction. Therefore, timely detection of formaldehyde is conducive to health.

For human beings, gas recognition is more complex than visual recognition. Nobel laureates Richard Axel and Linda Buck have discovered in the study of olfactory mechanism that although human beings have only about 1000 olfactory genes, they can sense and recognize more than 10000 kinds of odorant chemicals [4]. The mechanism of gas recognition is very complex.

In this paper, a novel Formaldehyde detector namely HCHODetector is proposed. HCHODetector can detect formaldehyde concentration by convolutional neural networks (CNN) instead of traditional detection instrument. As a consequence, the experimental results show that the detection error is within 0.08 mg/m³, which provides a new idea for Formaldehyde concentration detection.
2. Related works

2.1. Formaldehyde
Formaldehyde is a common gas harmful to human health, the chemical formula is HCHO. Human exposure to formaldehyde is mainly through respiratory tract inhalation, oral intake and skin contact. At present, Formaldehyde detection usually uses special equipment or detection kit. Special equipment is very expensive and not suitable for daily testing.

Some common Formaldehyde detection kits on the market produce different colours through complex chemical reactions with phenol reagent, RO pure water, Ammonium iron sulfate, hydrochloric acid and other chemical reagents. Then people roughly judge the indoor Formaldehyde concentration by comparing the colour produced with the standard colorimetric card. Due to the strong subjectivity of naked eye observation and the influence of light, this detection method is inefficient. The common Formaldehyde detection colorimetric card on the market is shown in Figure 1.

![Figure 1. The common Formaldehyde detection colorimetric card on the market](image)

2.2. Convolutional neural networks (CNN)
CNN is an important model in the field of deep learning, which is widely used in object detection, image classification and image segmentation. Original images can be input into CNN without complex pre-processing, which can avoid low-level feature extraction manually and data reconstruction. Therefore, CNN has been widely used in the field of image pattern recognition. The parameters of CNN will be optimized by back-propagation algorithm. The CNN is generally composed of convolutional layers, pooling layers, fully connected layers and one classifier. The convolutional layer uses weight sharing, so the network can extract image features with fewer parameters, which reduces the network complexity [5]. It should be noted that CNN are not sensitive to rotation invariance of image [6].

2.3. HSV Colour Space
Colour space is a kind of mathematical method to visualize colour. The common colour space includes RGB and HSV. Traditional CNN only deals with images in HSV colour space.

HSV is a kind of colour space whose channel scale conforms to the human perception law of colour, including H, S and V channels. Among them, H channel represents hue, representing different colours, i.e. the position of spectral colour (red, green, etc.), with the range of 0° to 360°. S channel represents the saturation, and the value is from 0 to 1. It is the pure spectral colour if S = 1. Adding white can reduce the saturation. V channel represents the value, which indicates the brightness of the colour. The value of V is also from 0 to 1. Adding black can reduce the V value. In HSV colour space, H channel and S channel contain colour information and are closely related to the way the human eye perceives colour, while V channel has nothing to do with the colour information. The HSV colour space model is shown in Figure 2.

When the image is taken and detected in the outdoor environment, the illumination has a great influence on the image. HSV space is sensitive to colour, but not sensitive to external factors such as illumination. It is not only a data augmentation method, but also can effectively reduce the influence of light factors in the detection. Ref. [7] proved that using HSV colour space for data augmentation can effectively improve the feature extraction effect of CNN.
3. **HCHODetector**

In this section, the experimental data, network pre-training and network training of HCHODetector are introduced.

### 3.1. Experimental data

Using deep learning technology for image detection needs a lot of original image data. We purchased 9 kinds of popular Formaldehyde detection kits on the market. These kits mainly use the complex chemical reaction to change colour. Moreover, the chemicals used by each product manufacturer are not exactly the same, so there are different discoloration reactions. Then we use the industrial camera to take images in the professional gas distribution system, and set the Formaldehyde concentration range from 0.0 to 1.5mg/m³, which is the common indoor Formaldehyde concentration range. The sampling interval is 0.1mg/m³. A total of 16 cases are collected, and 50 images are taken in each case without data augmentation. Some images collected are shown in Figure 3.

![Original image in RGB colour space](image1.png) ![The image in HSV colour space](image2.png)

**Figure 3. The images collected by industrial camera**

It should be noted that although the collected image is a standard square, there are strict direction requirements for the image due to the sequencing of different detection kits from various manufacturers. When using CNN to extract features, the image direction plays a very important role. As a result, when the image collected is input to CNN, it is needed to maintain the consistency of the image direction.

The image captured by conventional industrial camera is in RGB colour space, and we convert it to HSV colour space for data enhancement. As mentioned above, the H and S channels represent hue and saturation respectively. The V channel represents the brightness, and the V is independent of the colour information of the image. Finally, the original RGB image and HSV image are concatenated as the input of CNN.

In order to increase the number of images, we use the image data augmentation methods include parallel shift, and noise addition. Because the image direction is important, image rotation which is a common data augmentation method is not used in HCHODetector.
3.2. Network pre-training

For conventional image classification tasks based on CNN, the networks are usually pre-trained on ImageNet dataset. Because the Formaldehyde detection does not involve the task of conventional image classification, there is no need to pre-train the CNN on ImageNet. The networks are based on MobileNet V1 as the backbone, in order to facilitate the HCHODetector to run on the smartphone. In this way, the Formaldehyde detection can be more convenient.

The Deep Auto Encoder (DAE) structure is used for pre-training. The whole pre-training network consists of Encoder and Decoder, as shown in Figure 4. The basic structure of Encoder is MobileNet V1. Through pre-training, MobileNet V1 can initially extract features of the image, such as location information of detection kits. After the pre-training, the orange part in Figure 4 namely Decoder will be dropped. Encoder input is the concatenated image in RGB+HSV colour space. The input image size is 224×224×6. Because the number of images used in the experiment is small, the use of this pre-training can improve the detection accuracy.

3.3. Network training

In this section, we introduced the network training process of HCHODetector. MobileNet V1 is used as the backbone of HCHODetector, and the output result is the Formaldehyde concentration in the environment.

For data normalization, we adjusted the concentration range to 0~1. Assuming that the current concentration is \(N\), the normalized \(N_{\text{norm}}\) can be calculated as follows:

\[
N_{\text{norm}} = \frac{N}{1.5 - 0.0}
\]

HCHODetector uses Sigmoid classifier in the end and outputs \(N_{\text{norm}}\). The range of output result is 0-1. Then \(N_{\text{norm}}\) can be revised to the true value \(N\) by the above formula. The revised \(N\) is predicted Formaldehyde concentration in the environment. The overall network structure of HCHODetector is shown in Figure 5.
In addition to the backbone namely MobileNet V1, we use the Mask-guided module in Ref. [8] to improve the detection accuracy. As described in Section 2.3, the H and S channels contain rich color information. We use the information contained in these two channels as additional supervision. Because the V channel can be considered to represent the light intensity, it can be regarded as an interference term, so it is not considered.

Mask-guided module leverages the H and S channel information as extra supervision to enhance the feature representation learning of the backbone network. This structure is analogous to Mask R-CNN [9]. The module utilizes some convolutional layers, one deconvolutional layer and a per-pixel sigmoid function. The average binary cross-entropy loss is applied to train this module. Note that the Mask-guided module only works in training stage and brings no extra computations in the inference stage. The Mask-guided module can be considered as another form of attention mechanism to assist the feature learning.

The whole network is optimized by \( L = L_{\text{Huber}} + L_H + L_S \). Note that \( L_{\text{Huber}} \) presents classification Huber loss between the true value and predicted value. \( L_H \) and \( L_S \) present the average binary cross-entropy loss of H channel and S channel respectively.

4. Experiments
We use RTX 2070 as GPU and CUDA 11.0 to accelerate computing. Firstly, the HCHODetector need to be pre-trained on the pre-training networks through the DAE structure. Secondly, the last part of DAE namely Decoder is removed. The Encoder is connected to the classifier using Huber Loss. The HCHODetector will be trained in the image dataset with data augmentation until the loss function converges.

Then HCHODetector is applied on the smartphone to detect Formaldehyde concentration in the environment. Firstly, the collected images are tested for the ideal environment detection. We select 16 images from the experimental dataset, and the experimental results are shown in Figure 6. It can be seen from the figure that the detection error is low relatively, and the maximum error is 0.04mg/m³, which has achieved good detection performance in the ideal environment.

![Figure 6. The experimental results in the ideal environment](image)

In order to further test the practicability of HCHODetector, we conduct the experiments in random actual environments, such as playground, classroom newly decorated and old warehouse. We get the predicted concentration value through HCHODetector running on the smartphone, and then get accurate concentration value through expensive special equipment. The detection results are shown in Figure 7. It can be seen from the figure that HCHODetector has achieved good detection performance in the actual environment, and the maximum error is 0.08mg/m³. The experimental results show that HCHODetector can effectively detect the Formaldehyde concentration in the environment.
5. Conclusion
In this paper, we propose a novel Formaldehyde detector namely HCHODetector. HCHODetector can detect formaldehyde concentration in the environment by CNN instead of traditional detection instrument. Specifically, this detector is based on deep learning and HSV colour space augmentation. Moreover, we propose a novel Mask-guided module and a novel pre-training network to enhance the colour discrimination ability of HCHODetector. As a consequence, the experimental results show that the detection error is within 0.08 mg/m$^3$, which provides a new idea for Formaldehyde concentration detection.

Acknowledgments
We acknowledge funding received from the Key Research Program of the Science Foundation of Shandong Province (ZR2020KE001), and the General Research Project of Zaozhuang University.

References
[1] LeCun, Y., Bengio, Y., Hinton, G. (2015) Deep learning. nature, 521(7553): 436-444.
[2] He, K., Zhang, X., Ren, S., Sun, J. (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. Las Vegas. pp. 770-778.
[3] Graves, A., Mohamed, A., Hinton, G. (2013) Speech recognition with deep recurrent neural networks. In: 2013 IEEE international conference on acoustics, speech and signal processing. Vancouver. pp. 6645-6649.
[4] Mombaerts, P. (2004) Love at first smell—the 2004 Nobel Prize in Physiology or Medicine. New England Journal of Medicine, 351(25): 2579-2580.
[5] Krizhevsky, A., Sutskever, I., Hinton, G. E. (2017) Imagenet classification with deep convolutional neural networks. Communications of the ACM, 60(6): 84-90.
[6] Sabour, S., Frosst, N., Hinton, G. E. (2017) Dynamic routing between capsules. In: Advances in neural information processing systems. Long Beach. pp. 3856-3866.
[7] Ganesh, P., Volle, K., Burks, T. F., Mehta, S.S. (2019) Deep orange: Mask R-CNN based orange detection and segmentation. IFAC-PapersOnLine, 52(30): 70-75.
[8] Chi, C., Zhang, S., Xing, J., Lei, Z., Stan, Z. L., Zou, X.D. (2020) PedHunter: Occlusion Robust Pedestrian Detector in Crowded Scenes. In: The AAAI Conference on Artificial Intelligence. New York. pp. 10639-10646.
[9] He, K., Gkioxari, G., Dollár, P., Ross, G. (2017) Mask r-cnn. In: Proceedings of the IEEE international conference on computer vision. Venice. pp. 2961-2969.