

Vision-based Sensing Framework for Automatic Estimation of Smoke Opacity Levels

Fay Huang,^{1*} Hong-Yi Jang,¹ Yuan-Jen Ou-Yang,² and Bao-Jen Pong³

¹Department of Computer Science and Information Engineering, National Ilan University,
No. 1, Sec. 1, Shennong Rd., Yilan City, Yilan County 26047, Taiwan

²Department of Electronics and Electrical Engineering, National Yang Ming Chiao Tung University,
No. 1001, Daxue Rd., Hsinchu City 300, Taiwan

³Industrial Technology Research Institute,
No. 195, Sec. 4, Zhongxing Rd., Zhudong Township, Hsinchu County 310401, Taiwan

(Received March 27, 2026; accepted August 26, 2026)

Keywords: smoke opacity estimation, smoke segmentation, vision-based sensing, visible-light imaging, smoke-type classification

Smoke opacity is an important indicator for evaluating emissions from stationary pollution sources, but conventional visual inspection is still affected by observer subjectivity, illumination conditions, and background complexity. In this study, we propose a vision-based framework for the automatic estimation of smoke opacity levels from visible-light images. The proposed method consists of four sequential stages: smoke-region segmentation, smoke-type classification, local patch-based opacity estimation, and final decision making. First, a U-shaped convolutional network (U-Net) model is used to extract smoke regions and suppress irrelevant background interference. Next, a support vector machine classifies the segmented smoke as black smoke or white smoke. In accordance with the classification result, sampled local patches are processed by the corresponding convolutional neural network model to predict patch-level opacity classes. Finally, for each sampled patch, a top- k weighted strategy selects the k highest-scoring opacity classes and combines them using normalized score-based weights to obtain a continuous opacity estimate; the patch-level estimates are then averaged to derive the final image-level opacity value. To evaluate the proposed framework, experiments were conducted on a collected dataset with physically measured opacity labels. The results show that the classification-based model achieved average testing accuracies of 90.16% for black smoke and 92.55% for white smoke. When the proposed top- k weighted method was used to estimate continuous opacity values, the accuracies reached 90.6% for black smoke and 96.0% for white smoke within an error tolerance of $\pm 10\%$. These results indicate that the proposed framework provides an effective and practical approach for automated smoke-opacity estimation under real-world environmental conditions.

*Corresponding author: e-mail: fay@niu.edu.tw
<https://doi.org/10.18494/SAM6355>

1. Introduction

Air pollution monitoring has long been an important issue in environmental protection and industrial regulation. Among the indicators used for evaluating emissions from stationary pollution sources, smoke opacity is one of the most direct and practical measures because it reflects the visual density of emitted smoke and is closely related to combustion conditions and possible air pollution violations. In many countries, smoke opacity has traditionally been assessed by trained observers through visual inspection under specified observation conditions.^(1,2) Although this approach has been widely used in regulatory practice, its reliability is still affected by human subjectivity and field conditions.

Previous studies and regulatory practices have established several representative approaches for smoke observation. In the United States, the Environmental Protection Agency (EPA) introduced Method 9 for the visual determination of smoke opacity and Method 22 for the observation of visible emissions.⁽¹⁾ In addition, digital-image-based approaches such as ASTM D7520-16 have been adopted to improve the objectivity and traceability of opacity evaluation.⁽²⁾ These methods indicate that smoke assessment is gradually moving from manual visual judgment to digital and image-assisted analysis. In Taiwan, smoke inspection for stationary pollution sources is also regulated by relevant environmental laws and emission standards, and opacity criteria are used as an important basis for compliance evaluation and enforcement.^(3–5)

In practical inspection, smoke is commonly categorized into black smoke and white smoke. Although the grading basis is the same, the interpretation process differs because the optical characteristics of black smoke and white smoke are not identical. Black smoke is generally produced by incomplete combustion and often appears as dark regions with relatively clear boundaries in images. By contrast, white smoke is more easily affected by condensation, scattering, high brightness, and background interference, and therefore usually exhibits weak boundaries and lower visual distinguishability. As a result, white smoke is often more difficult to interpret than black smoke in both manual inspection and image-based analysis.⁽¹⁾

Despite the availability of regulatory procedures and digital monitoring concepts, several problems remain. First, conventional visual inspection still depends strongly on observer experience, viewing angle, illumination condition, and environmental background. Therefore, the inspection results may vary among observers or across monitoring conditions. Second, although some digital or optical systems have been proposed for smoke monitoring, many of them require costly instruments or are not sufficiently flexible for large-scale deployment. Third, existing image-based approaches are still limited in dealing with white smoke, transparent plumes, irregular smoke shapes, and complex outdoor backgrounds. These problems reduce the robustness and practical applicability of automated smoke assessment systems.

In our previous studies, we investigated image-based smoke analysis and related opacity estimation techniques.⁽⁶⁾ Other researchers also reported that digital imaging and computer vision can provide a feasible alternative to conventional visual inspection for environmental monitoring. However, most existing methods mainly rely on brightness contrast, fixed analysis regions, or simplified smoke conditions, and they do not adequately address the difficulties caused by white smoke and background interference. Therefore, a more robust and practical framework is still needed for automatic smoke opacity estimation under diverse field conditions.

In this study, we propose a vision-based method for the automatic estimation of smoke opacity levels from visible-light images. The proposed framework integrates smoke segmentation, smoke-type classification, and patch-based opacity estimation into a sequential learning pipeline. A U-shaped convolutional network (U-Net) model is first used to identify the smoke region, after which a support vector machine (SVM) is applied to distinguish black smoke from white smoke. Subsequently, local smoke patches are processed by the corresponding convolutional neural network (CNN)-based opacity estimation model, and the final smoke opacity value is determined through top- k aggregation. By separating smoke regions from complex backgrounds and accounting for the distinct visual characteristics of black smoke and white smoke, the proposed method is expected to improve the stability and reliability of smoke-opacity estimation, particularly in challenging cases involving weak boundaries or background interference. The aim of this study is to develop a modular and regulation-aligned image-based sensing framework that can support future automated environmental monitoring and smoke inspection systems.

2. Related Work

Recent studies related to smoke analysis can be broadly divided into three categories: smoke-region segmentation in images, real-time smoke and fire detection, and image-based smoke-opacity measurement. For the first category, the main aim is to identify smoke regions under complex visual conditions. The second is focused on detecting the presence of smoke or fire for early warning and real-time monitoring. The third category is concerned with estimating plume opacity for environmental monitoring or regulatory purposes.

In smoke segmentation, deep-learning methods have been widely adopted to address the irregular shape, blurred boundary, and semitransparent appearance of smoke. Yuan *et al.*⁽⁷⁾ proposed a deep smoke segmentation framework to improve pixel-level smoke extraction. Wang *et al.*⁽⁸⁾ further applied a U-Net-based architecture to smoke segmentation in remote-sensing imagery. Andrade *et al.*⁽⁹⁾ also investigated smoke segmentation in satellite-based wildfire applications, with emphasis on real-time discrimination between smoke and cloud regions. Other studies have investigated uncertainty-aware modeling,⁽¹⁰⁾ noisy-label learning,⁽¹¹⁾ global-context interaction,⁽¹²⁾ and feature-space enhancement for faint smoke regions.⁽¹³⁾ Those studies revealed that smoke segmentation remains challenging when smoke is transparent, when boundaries are weak, or when the background is visually complex.

In addition to segmentation-based studies, recent object-detection approaches based on the You Only Look Once (YOLO) detector have been applied to real-time smoke and fire monitoring. Kim *et al.*⁽¹⁴⁾ developed an enhanced YOLO-based wildfire smoke detection system for a Jetson Xavier embedded environment. Their study was focused on real-time inference and edge deployment, and they showed that weighted loss design and focal loss can improve detection performance for challenging smoke samples. Lin *et al.*⁽¹⁵⁾ proposed a lightweight YOLOv8n-based model for smoke and fire detection in complex environments. Rather than relying on a larger network, their method improved the model structure through lightweight feature extraction, multiscale feature fusion, attention-based feature enhancement, and an

improved loss design, thereby reducing computational complexity while maintaining detection performance. These studies indicate that deep-learning-based smoke and fire detection has become increasingly practical for real-time and edge-device applications.

Practical smoke monitoring or opacity-related applications have also been considered in several studies. Pedrayes *et al.*⁽¹⁶⁾ proposed an automated method for opacity estimation in industrial emission monitoring. Wang *et al.*⁽¹⁷⁾ studied black-smoke detection and class evaluation in ship-emission scenarios. Park *et al.*⁽¹⁸⁾ developed a smoke-opacity measurement framework based on semantic segmentation and edge devices, and demonstrated the feasibility of combining smoke extraction with automated field deployment. In addition, digital optical methods using image devices, such as camcorders and smartphone cameras, have been explored for plume-opacity measurement.^(19,20) These studies are important because they provide more objective alternatives to manual visual observation.

Although significant progress was made in previous studies, several limitations remain. First, many smoke-segmentation studies focus on wildfire, satellite, or maritime scenes rather than visible-light stack smoke in environmental inspection settings. Second, recent YOLO-based smoke and fire detection studies mainly focus on detecting the existence and location of smoke or fire, whereas they do not directly estimate smoke opacity values for regulatory inspection. Third, digital optical measurement methods still depend strongly on imaging conditions, background contrast, and camera setup. Third, relatively few studies explicitly distinguish black smoke from white smoke before opacity estimation, even though the two smoke types have substantially different visual characteristics.

To address these issues, we propose a regulation-oriented framework for visible-light smoke-opacity estimation. The proposed method combines U-Net-based smoke segmentation, smoke-type classification, and CNN-based opacity estimation. By separating smoke-region extraction, black/white smoke identification, and opacity prediction into different stages, the proposed framework is intended to improve robustness and practical applicability for stationary-source smoke monitoring.

3. Data and Methods

In this section, we present the proposed framework and its implementation in detail. The overall workflow is first described to illustrate the structure of the complete system. The models employed in each stage are then introduced, together with their design considerations. Subsequently, the dataset preparation procedure is presented, including data acquisition, annotation, preprocessing, and the division of samples into training, validation, and test sets according to the requirements of the proposed models. Finally, the training strategy and implementation details are described.

3.1 Overall workflow

We propose an automated deep-learning-based framework for smoke opacity estimation, as illustrated in Fig. 1. The proposed framework consists of four sequential stages: smoke-region

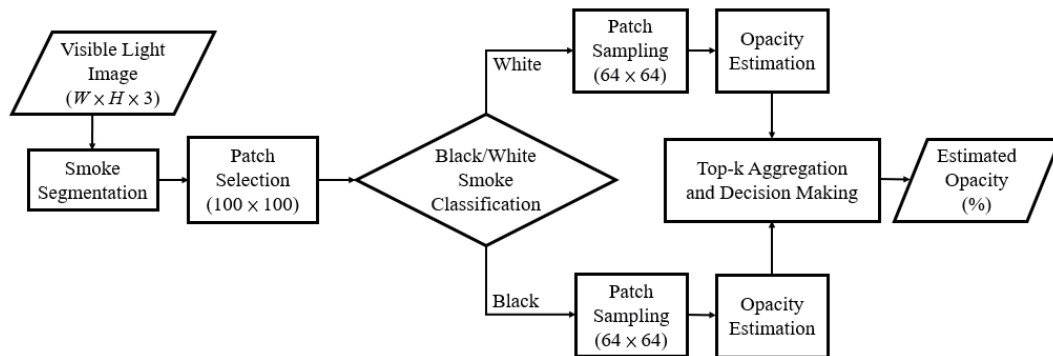


Fig. 1. System architecture of the proposed framework.

segmentation, smoke-type classification, local patch-based opacity estimation, and final decision making. First, the smoke region is extracted from the visible-light image by semantic segmentation. Then, a 100×100 smoke-region patch above the stack outlet is selected for smoke-type classification. According to the classification result, several appropriate 64×64 local patches are randomly sampled from the segmented smoke region and processed by the corresponding convolutional neural network (CNN)-based opacity estimation model. The predicted opacity levels from multiple local patches are further integrated using a top- k strategy to determine the final smoke opacity value.

To reduce the influence of irrelevant background information, the proposed framework first applies a region-oriented segmentation process to isolate the actual smoke area before opacity estimation. A U-Net architecture is employed to generate semantic masks of smoke regions. Owing to its encoder–decoder structure and multiscale feature fusion capability, U-Net can effectively preserve spatial details near smoke boundaries while simultaneously capturing high-level semantic information. This characteristic is particularly advantageous for smoke images in which blurred edges and irregular plume shapes are commonly observed. Using the segmentation result, a 100×100 patch is subsequently extracted from the isolated smoke region and used for smoke-type classification.

The smoke-type classification stage is introduced to address the intrinsic differences between black smoke and white smoke. Although both smoke types are evaluated on the same opacity grading basis, their optical characteristics differ substantially. Black smoke, which is typically rich in carbonaceous particles, exhibits stronger light absorption and higher contrast. In contrast, white smoke, which is mainly composed of water vapor and fine aerosols, often appears semitransparent, has lower contrast, and is more likely to blend with bright or complex backgrounds. If both smoke types are directly combined in a single estimation model, the underlying feature distributions may become inconsistent, thereby reducing prediction accuracy. Therefore, a dedicated classification stage is adopted to separate smoke images into black-smoke and white-smoke categories before opacity estimation.

Finally, two independent CNN-based opacity estimation models are constructed for black smoke and white smoke. Each model predicts patch-level opacity classes from sampled 64×64 local patches. The resulting predictions are then aggregated using a top- k strategy to derive the final smoke opacity value. Through this dual-path design, the proposed framework can better adapt to the distinct optical properties of different smoke types and generate more reliable opacity estimates.

3.2 Smoke segmentation, smoke-type classification, and opacity estimation

The proposed framework is composed of three machine learning models: smoke segmentation, smoke-type classification, and opacity estimation. In this subsection, the design details and functional roles of these three components are described.

3.2.1 U-Net-based smoke segmentation

To ensure that the estimation model focuses on meaningful smoke regions rather than background interference, a semantic segmentation module based on the U-Net architecture was adopted as the first stage of the proposed framework. This module serves as a preprocessing step for smoke-region extraction and binary mask generation, which are subsequently used for patch selection and opacity estimation.

As illustrated in Fig. 2, the U-Net model follows a standard encoder–decoder architecture. In the encoder, the input RGB image is progressively downsampled through repeated convolution and max-pooling operations to capture high-level semantic features. In the decoder, transposed convolution layers are used to restore spatial resolution, while skip connections are introduced to fuse low-level spatial details from the encoder with high-level semantic information from the

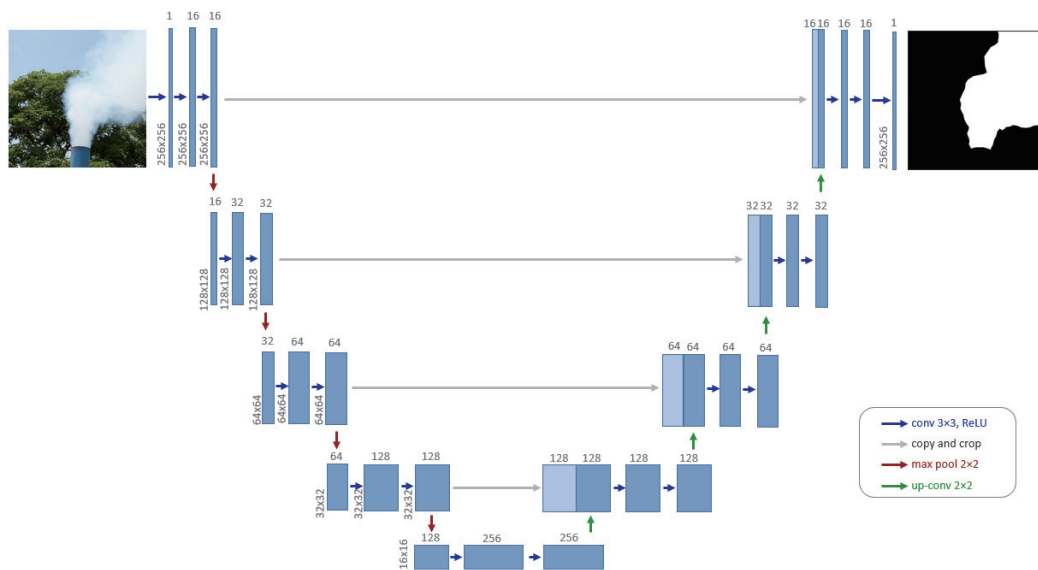


Fig. 2. (Color online) The U-Net architecture.

decoder. This structure enables the model to preserve fine boundary information while maintaining global contextual understanding, which is particularly important for smoke images with blurred edges, irregular shapes, or semitransparent regions.

By using the segmentation mask generated by U-Net, the proposed framework can reduce the influence of irrelevant background patterns and constrain subsequent processing to the smoke region of interest. This design is intended to improve the consistency of the following classification and opacity estimation stages.

3.2.2 Smoke-type classification

After smoke-region extraction, a smoke-type classification step is performed to distinguish black smoke from white smoke before opacity estimation. Although both smoke types are evaluated under the same opacity grading basis, they exhibit substantially different optical characteristics. Black smoke generally shows stronger light absorption and higher contrast, whereas white smoke tends to be semi-transparent, lower in contrast, and more susceptible to interference from bright or complex backgrounds. If both smoke types are directly processed using a single estimation model, the resulting feature distributions may become inconsistent and reduce prediction reliability.

To address this issue, we employ a SVM as the smoke-type classification model. The input to the classifier is a 100×100 smoke-region patch selected from the segmented region above the stack outlet. From this patch, brightness-related features are extracted to characterize the visual appearance of the smoke region. These features are then used by the SVM to construct a decision boundary between black-smoke and white-smoke samples.

The SVM classifier is adopted because it is effective for binary classification problems with relatively low-dimensional features and provides a clear decision boundary for separating classes with distinct visual distributions. In the proposed framework, the classification result is used to determine the subsequent regression path, such that black-smoke and white-smoke samples are processed by their corresponding opacity estimation models.

3.2.3 CNN-based opacity estimation

Following smoke-type classification, smoke opacity estimation is performed using a CNN-based model. Multiple appropriate 64×64 local patches are sampled from the segmented smoke region and used as the input to the network. To make the patch selection procedure more reproducible, the local patches are sampled in accordance with the segmentation mask generated by the U-Net model. Candidate patches are restricted to the smoke region above the stack outlet. A patch is selected only when its center is located within the segmented smoke mask and the patch overlaps the smoke region. For each test image, eight 64×64 local patches are randomly sampled from the valid smoke region for opacity estimation. This mask-guided sampling strategy is used to reduce the influence of background-dominant patches and to ensure that the sampled patches contain representative smoke information. These local patches are intended to capture representative spatial and intensity-related patterns associated with smoke opacity.

As illustrated in Fig. 3, the proposed CNN architecture consists of three convolutional modules for hierarchical feature extraction. Each convolution layer employs a 3×3 kernel, followed by max-pooling operations that progressively reduce the spatial dimensions of the feature maps. Through this hierarchical structure, the network is able to learn local texture, intensity variation, and structural characteristics relevant to smoke opacity. The resulting feature maps are then flattened into a one-dimensional vector and passed through fully connected layers for opacity classification. The final output dimension is defined in accordance with the smoke type.

To account for the substantial visual differences between black smoke and white smoke, two independent CNN models are constructed and trained separately. For black smoke, the final output layer contains four neurons corresponding to the four opacity classes of 20, 40, 60, and 75%. For white smoke, the final output layer contains five neurons corresponding to the five opacity classes of 10, 20, 40, 60, and 75%. In this manner, the proposed framework can better accommodate the distinct optical characteristics and class distributions of the two smoke types.

Because opacity may vary across different local regions of the same smoke plume, the final smoke opacity is not determined from a single patch-level prediction. Instead, the predicted opacity levels obtained from multiple sampled patches are further integrated using a top- k strategy to derive the final smoke opacity value. This patch-based classification and aggregation process helps reduce the influence of locally unrepresentative regions and improves the robustness of the final estimation result.

3.3 Dataset preparation

In this subsection, we describe the preparation of the datasets used for U-Net-based smoke segmentation, smoke-type classification, and opacity estimation. It includes the data acquisition setup, the ground-truth measurement procedure, the label definition, the dataset composition, and the data augmentation and splitting strategies adopted for model development and evaluation.

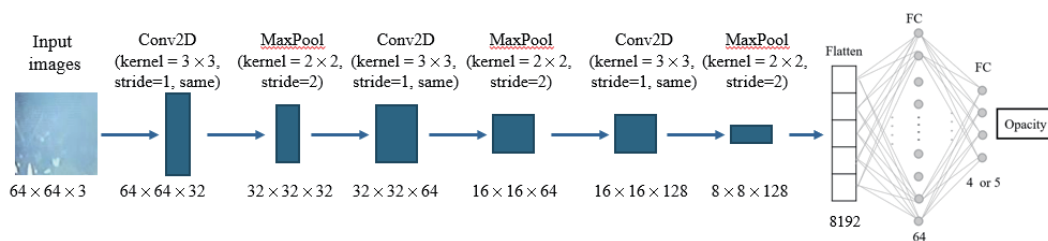


Fig. 3. (Color online) Smoke opacity estimation model architecture.

3.3.1 Data collection setup

To investigate the feasibility of automating the visual smoke-opacity assessment procedure used in environmental inspection, representative smoke-image data were collected under standardized observation conditions. The data collection process was designed to establish a reliable correspondence between visible-light smoke images and physically measured opacity values.

Two outdoor experimental sites were established in Taiwan, located in Taoyuan and Tainan. The observation setup followed the image-acquisition principles specified in ASTM D7520, while also considering the regulatory requirements of Taiwan's Ministry of Environment for stationary-source smoke-opacity evaluation. This design was intended to ensure that the collected image data were both representative and regulation-relevant.

As illustrated in Fig. 4(a), the observer and the camera were positioned at a horizontal distance greater than three times the stack height ($D > 3H$), such that the elevation angle to the stack outlet remained below 18° . In addition, to reduce overexposure and improve the visibility of the smoke boundary, the observation direction was arranged under side-lighting to back-lighting conditions, with the sun located within 70° behind the observer. These geometric and illumination constraints were adopted to improve image consistency and reduce variations caused by unstable viewing conditions.

To simulate smoke emissions under controllable yet realistic outdoor conditions, a smoke-generation device was installed at both experimental sites. The device was able to produce different smoke types, including black smoke and white smoke, with controllable opacity levels. During data acquisition, smoke images were recorded under natural illumination and varying background conditions to better reflect practical outdoor inspection scenarios.

3.3.2 Ground-truth acquisition

To provide objective reference values for opacity labeling, a transmissive optical sensing device was installed at the middle section of the stack, as shown in Fig. 4(b). The sensing

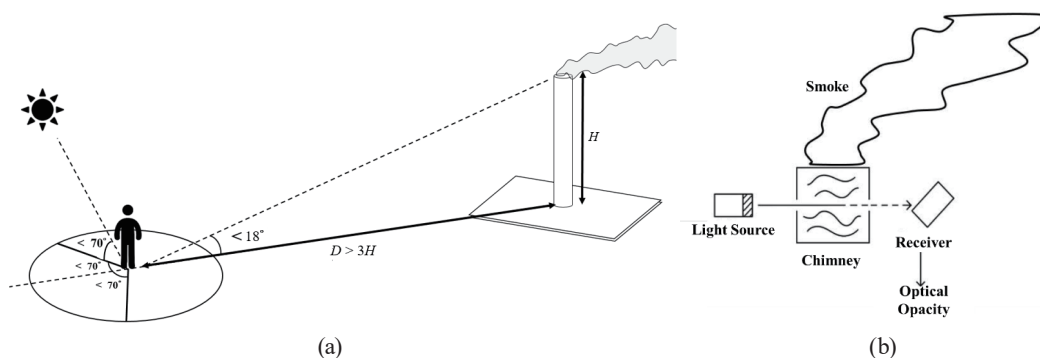


Fig. 4. (a) Observation geometry and illumination configuration for smoke-image acquisition. (b) Transmissive optical sensing device used for opacity measurement.

principle was based on the Beer–Lambert law. A light beam emitted from the source passed through the smoke medium and was attenuated by particle absorption and scattering. The remaining light intensity was measured by a photodetector placed on the opposite side, and the corresponding smoke opacity value was calculated from the difference between the emitted and transmitted light intensities.

This optical measurement system provided synchronized physical opacity values for the recorded smoke images and was used as the ground-truth source for dataset labeling. Compared with manual visual labeling, this approach reduces subjectivity and improves the consistency and traceability of the dataset.

3.3.3 Label definition

Each recorded smoke image was assigned an opacity label on the basis of the measured value obtained from the optical sensor. Because black smoke and white smoke exhibited different data distributions in this study, separate class settings were adopted for the two smoke types. For black smoke, the opacity labels were grouped into four discrete classes: 20, 40, 60, and 75%. For white smoke, the opacity labels were grouped into five discrete classes: 10, 20, 40, 60, and 75%. These class definitions were used as the target labels for the corresponding CNN-based opacity estimation models.

Although practical smoke-opacity assessment standards may have finer grading intervals, smoke is a dynamic fluid phenomenon and its measured opacity can fluctuate considerably over time, even under the same nominal emission condition. In actual inspection practice, opacity is therefore not judged from a single instantaneous observation, but from measurements collected over a period of time and summarized as a representative average value. Following this consideration, the present dataset adopted a coarser class design than the finest grading interval, so as to improve label stability and reduce the effect of instantaneous fluctuation in smoke appearance and measurement.

3.3.4 Dataset construction

After image acquisition and opacity labeling, the dataset was organized for subsequent smoke segmentation, smoke-type classification, and opacity estimation tasks. From each recorded image, a 640×640 square region containing the smoke plume was first extracted from the original scene for smoke segmentation. Before being input to the U-Net model, these cropped images were resized to 256×256 pixels. Each cropped smoke image was then manually annotated to generate a binary smoke mask for U-Net training. Representative black-smoke and white-smoke images together with their corresponding binary masks are shown in Fig. 5.

Before data augmentation, the original cropped dataset contained 161 images in total, comprising 80 black-smoke images and 81 white-smoke images. To increase data diversity and improve model robustness, three augmentation operations were applied to each image: horizontal flipping, angular variation, and angular variation followed by horizontal flipping. After augmentation, the dataset size was increased fourfold to 644 images, including 320 black-smoke

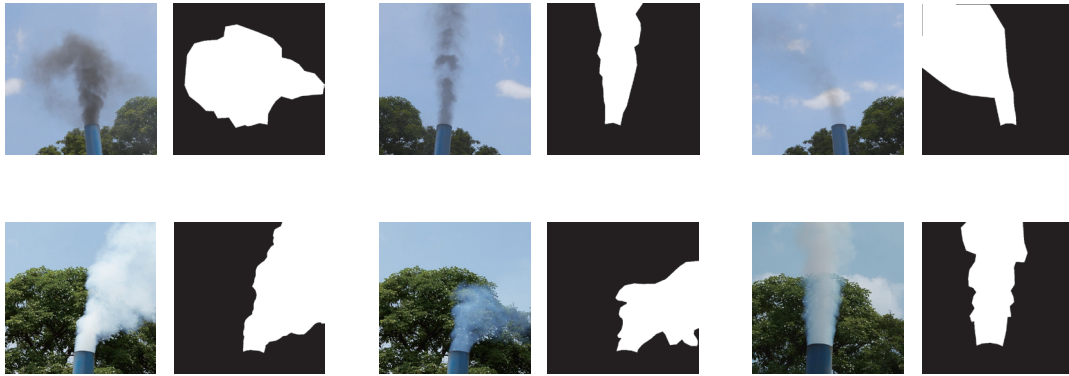


Fig. 5. (Color online) Representative smoke images and their corresponding manually annotated binary masks used for smoke segmentation. The upper row shows black-smoke examples, and the lower row shows white-smoke examples.

images and 324 white-smoke images. For model development, the augmented dataset was further divided into training and validation sets using an 8:2 split. In addition, an independent test set containing 43 images was reserved for the final performance evaluation, made up of 18 black-smoke images and 25 white-smoke images.

For smoke-type classification, the same cropped smoke-region images were further processed to extract 100×100 patches located above the stack outlet. These patches were used as the input to the SVM classifier for distinguishing black smoke from white smoke.

For smoke opacity estimation, a substantially larger patch-level dataset was constructed. After smoke-type identification, multiple 64×64 local patches were sampled from the segmented smoke region of each image. Because multiple local patches were extracted from each smoke image, the total number of samples used for opacity estimation was much larger than that used for smoke segmentation and smoke-type classification.

Separate class settings were adopted for black smoke and white smoke in the opacity estimation stage. For black smoke, the CNN model was trained using four opacity classes: 20, 40, 60, and 75%. For white smoke, the CNN model was trained using five opacity classes: 10, 20, 40, 60, and 75%. The numbers of patch-level samples in each opacity class before and after sampling and augmentation are summarized in Table 1. Representative examples of five patches from each opacity level are shown in Fig. 6.

3.4 Training strategy

Here, we describe the training strategies adopted for the three components of the proposed framework, namely, U-Net-based smoke segmentation, smoke-type classification, and CNN-based opacity estimation. Because these components address different tasks and operate on different data representations, separate training procedures were designed for each of them. The corresponding strategies were developed to improve model robustness, enhance adaptability to

Table 1
Training image set for smoke opacity estimation.

Opacity Level	Black smoke (original)	White smoke (original)	Black smoke (after sampling and augmentation)	White smoke (after sampling and augmentation)
10	0	20	0	215
20	25	20	184	203
40	15	20	160	201
60	35	20	229	203
75	25	21	200	202
Sum	100	101	773	1024

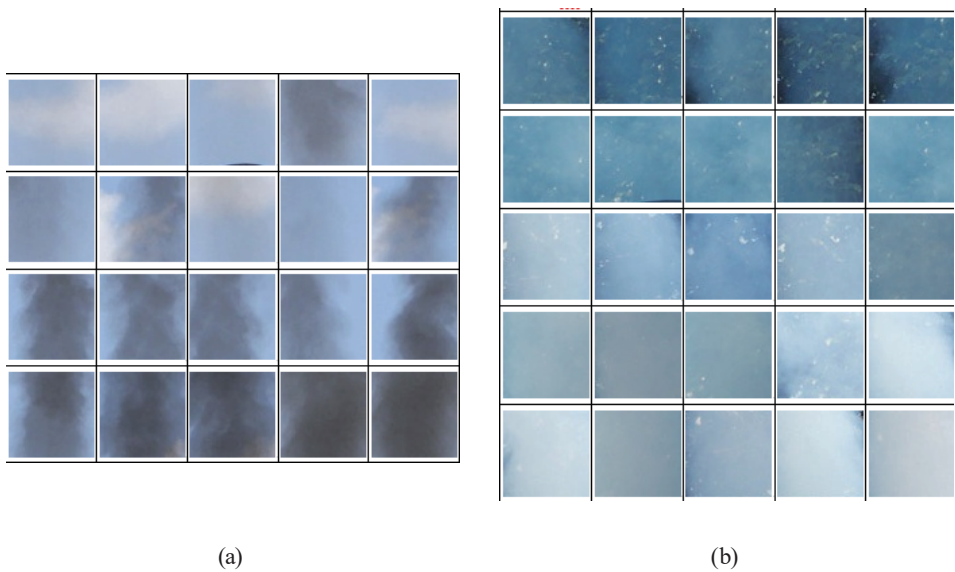


Fig. 6. (Color online) Representative image patches used for opacity estimation training. (a) Black-smoke samples at opacity levels of 20, 40, 60, and 75%, arranged from top to bottom. (b) White-smoke samples at opacity levels of 10, 20, 40, 60, and 75%, arranged from top to bottom.

real-world environmental conditions, and ensure stable performance throughout the overall smoke-opacity estimation pipeline.

3.4.1 U-Net fine-tuning strategy

A pretrained open-source U-Net model was adopted as the initial smoke-segmentation model in this study. Although the pretrained model showed satisfactory performance on its original training data, its segmentation quality degraded when applied to the real-world smoke images collected in this study. This performance drop was mainly caused by the domain gap between the original training images and the target scenes, which contained more complex backgrounds, including trees, buildings, shadows, and substantial variations in illumination. .

To improve the adaptability of the segmentation model to real-world conditions, the pretrained U-Net was further fine-tuned using the smoke-segmentation dataset constructed in

this study. The training images consisted of 640×640 cropped smoke-region images, which were resized to 256×256 before being input to the network. Each image was manually annotated to generate a binary smoke mask, which was used as the ground-truth label for supervised training.

This fine-tuning step was intended to improve mask quality and reduce false detections in complex scenes. It was particularly important because the subsequent smoke-type classification and opacity estimation stages relied on accurate smoke-region extraction. Qualitative examples before and after fine-tuning are shown in Fig. 7.

3.4.2 Training strategy for smoke-type classification

For smoke-type classification, SVM was trained to distinguish black smoke from white smoke using a 100×100 patch extracted from the segmented region above the stack outlet. Two brightness statistics, namely, the mean brightness and the standard deviation of brightness, were used as input features because they effectively represent the major visual differences between the two smoke types. A linear SVM was adopted to construct the decision boundary in the resulting two-dimensional feature space.

For model development, the available smoke-type classification samples were randomly divided into training and validation sets using an 8:2 split. The training set was used to construct the classifier, while the validation set was used to examine the stability of the classification performance. Using the mean brightness x and the standard deviation of brightness y , the linear SVM decision boundary is given by

$$y = -0.8099x + 116.7812. \quad (1)$$

The distribution of the extracted brightness features and the corresponding linear SVM decision boundary are shown in Fig. 8. In this feature space, black-smoke samples are mainly distributed in the lower-brightness and lower-variance region, whereas white-smoke samples generally occupy regions with higher brightness and greater variability. This feature distribution supports the use of brightness statistics as an effective basis for smoke-type classification. The

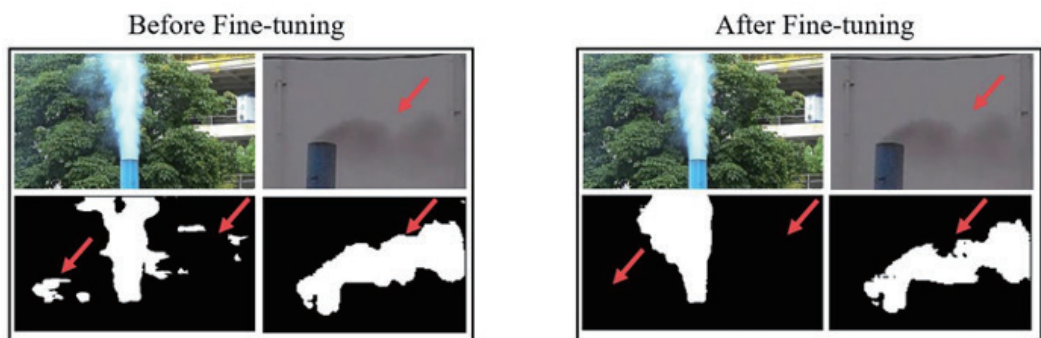


Fig. 7. (Color online) Smoke segmentation results before and after fine-tuning.

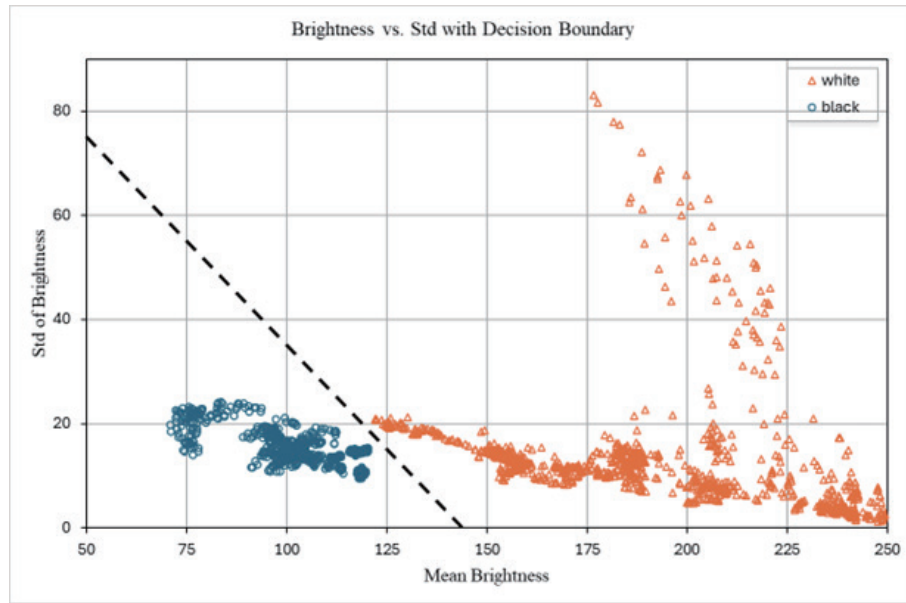


Fig. 8. (Color online) Smoke-type classification result.

trained classifier was subsequently used to determine the appropriate opacity estimation path for each smoke sample.

3.4.3 Training strategy for opacity estimation

For opacity estimation, two CNN models were trained separately for black smoke and white smoke. The input patches were resized to 64×64 pixels and normalized before being fed into the network. The black-smoke model was trained using four opacity classes, namely, 20, 40, 60, and 75%, whereas the white-smoke model was trained using five opacity classes, namely, 10, 20, 40, 60, and 75%. The models were trained using Adam with a learning rate of 0.001 and a batch size of 20. The classification loss was computed using categorical cross-entropy. To determine suitable training durations, the number of training epochs was varied from 50 to 350 in increments of 50. The final reported results were obtained using 350 epochs for the white-smoke model and 300 epochs for the black-smoke model.

4. Experimental Results

The experimental results of the proposed smoke-opacity estimation framework using the newly collected dataset and two evaluation strategies are presented. The first strategy is for the evaluation of the classification performance of the CNN-based opacity estimation models for black smoke and white smoke. The second strategy further converts the discrete classification outputs into continuous opacity estimates using the proposed top- k weighted method, so as to assess the practical applicability of the model in real-world smoke monitoring scenarios.

4.1 Classification-based evaluation

To evaluate the classification performance of the proposed method, overall classification accuracy was adopted as the primary performance metric. The dataset samples were randomly divided into training and testing sets in accordance with the experimental setting, and the training process was repeated ten times under the same configuration to examine performance stability across different runs. Training loss was also monitored to assess whether the model converged to a stable and sufficiently low level.

The testing set contained 160 black-smoke images and 200 white-smoke images. For white smoke, the best overall classification performance was obtained at 350 epochs. Under this setting, the training accuracy ranged from 93.28 to 98.12%, with an average of 96.17%, while the average testing accuracy reached 92.55%. For black smoke, a better trade-off was obtained at 300 epochs. Under this setting, the training accuracy ranged from 93.36 to 97.85%, with an average of 95.84%, and the average testing accuracy reached 90.16%. The training accuracy, training loss, and testing accuracy obtained from ten repeated experiments are summarized in Table 2. The accuracy values are reported as mean ± standard deviation to provide a concise overview of model performance and stability.

These results indicate that the proposed CNN-based opacity estimation models achieved stable and relatively high classification accuracy on the newly collected dataset. In particular, the white-smoke model achieved a testing accuracy above 92%, whereas that of the black-smoke model reached 90%. Overall, the results confirm that the proposed framework is effective for smoke-opacity level recognition under real-world environmental conditions.

4.2 Effect of smoke-type classification

To further examine the effect of the smoke-type classification stage, an additional comparison was conducted between the proposed dual-path model and a single-CNN model trained using both black-smoke and white-smoke samples. In the single-CNN setting, all smoke patches were combined and processed by one opacity estimation model without smoke-type separation. The same testing protocol was used for both settings.

For the single-CNN model, the mean accuracies were 94.46% for training and 91.04% for testing. In comparison, the proposed dual-path model achieved average accuracies of 96.00% for training and 91.36% for testing. Although the improvement was moderate, the dual-path design still achieved higher accuracy than the single-CNN setting. This result supports the design

Table 2
Summary of ten repeated experiments.

Smoke type / phase	Accuracy (%)	Average Loss
White-smoke training	96.17 ± 1.49	0.0952
White-smoke testing	92.55 ± 2.70	—
Black-smoke training	95.84 ± 1.47	0.1166
Black-smoke testing	90.16 ± 1.90	—

Note: The accuracy values are reported as mean ± standard deviation over ten repeated experiments.

choice of separating black smoke and white smoke before opacity estimation. Since the two smoke types exhibit different brightness distributions, boundary characteristics, and optical appearances, using separate opacity estimation models can better capture their respective visual patterns and improve the stability of the final opacity estimation.

4.3 Top- k weighted estimation of continuous opacity values

Although the CNN models were trained using discrete opacity classes, practical smoke monitoring often requires a continuous opacity estimate rather than only a class label. Therefore, we further introduced a top- k -based weighted estimation method that converts the output of the classification model into a continuous opacity prediction.

Let the set of opacity levels be denoted as $\{L_1, L_2, \dots, L_m\}$, where each level L_i corresponds to a fixed opacity percentage. For an input sampled patch, the CNN model produces a score vector $(r_1, r_2, \dots, r_m)$, where r_i denotes the output score associated with opacity level L_i . In this study, the rectified linear unit (ReLU), defined as $\text{ReLU}(x) = \max(0, x)$, was adopted as the output-layer activation function instead of softmax. This design was chosen because the exponential normalization of softmax tends to overemphasize the dominant class, which is less suitable for the subsequent weighted estimation of a continuous opacity value.

The top- k classes with the highest scores are first selected, and their corresponding indices are denoted as $\{c_1, c_2, \dots, c_k\}$ with scores $\{r_{c_1}, r_{c_2}, \dots, r_{c_k}\}$. To produce a continuous opacity prediction, only the selected top- k classes are retained, and their scores are renormalized within the selected subset. The normalized weight for the j -th selected class is defined as

$$w_{c_j} = \frac{r_{c_j}}{\sum_{l=1}^k r_{c_l}}, \quad j = 1, 2, \dots, k. \quad (2)$$

The final continuous opacity estimate $\hat{O}$ for a given image patch is then computed as

$$\hat{O} = \sum_{j=1}^k w_{c_j} L_{c_j}. \quad (3)$$

In this way, if the model assigns similarly high confidence to two adjacent opacity levels, the final estimated value can fall between them, thereby better reflecting the continuous variation of smoke opacity.

In this study, different values of k were used for black smoke and white smoke. For black smoke, the top two classes ($k = 2$) were used, whereas for white smoke, the top three classes ($k = 3$) were used because the white-smoke model involved more opacity classes. For the weight assignment to better reflect the actual opacity distribution, the selected top- k scores were renormalized using Eq. (2), and the final continuous opacity value was then calculated as the weighted sum of the corresponding opacity levels using Eq. (3).

Table 3 presents representative examples of the proposed top-*k* weighted estimation results for black smoke and white smoke. For each test image, the table lists the true opacity level, the final predicted continuous value, the selected top-*k* opacity labels, and their corresponding normalized weights. The testing set contained 200 white-smoke images and 160 black-smoke images, corresponding to 40 images for each opacity level. The results show that, within an error tolerance of $\pm 10\%$, the accuracies of the continuous opacity estimation reached 90.6% for black smoke and 96.0% for white smoke. This indicates that the proposed method can effectively capture the overall scale of smoke-opacity variation and that most estimated values fall within a practically acceptable range.

The results further suggest that the two smoke types exhibit different estimation characteristics under the current setting. White smoke showed more stable continuous-opacity predictions, whereas black smoke exhibited greater variability in several cases. Nevertheless, the proposed top-*k* weighted method provides a practical way to convert discrete class outputs into continuous opacity estimates, making the results more suitable for real-world smoke monitoring and interpretation.

For each test image, eight 64×64 patches were sampled from the segmented smoke region above the stack, as illustrated in Fig. 9. The proposed top-*k* weighted method was first applied to each patch to obtain a patch-level continuous opacity estimate. The final image-level opacity value was then computed as the average of the eight patch-level estimates.

Table 3
Representative results of the proposed top-*k* weighted estimation for 10 black-smoke and 10 white-smoke images.

Black				White			
True Opacity	Predicted	Labels	Weights	True Opacity	Predicted	Labels	Weights
60	57.99	[60, 40]	[0.8996,0.1004]	40	40.27	[40, 75, 60]	[0.991,0.0062,0.0028]
20	20.00	[20, 40]	[0.9998,0.0002]	75	68.89	[75, 60, 40]	[0.6603,0.2891,0.0505]
20	20.00	[20, 40]	[1,0]	20	18.14	[20, 10, 40]	[0.8135,0.1865,0.]
20	20.25	[20, 40]	[0.9876,0.0124]	10	10.00	[10, 20, 75]	[1,0,0]
40	28.06	[20, 40]	[0.5971,0.4029]	10	10.00	[10, 20, 75]	[1,0,0]
75	67.15	[60, 75]	[0.5234,0.4766]	20	20.49	[20, 40, 10]	[0.9748,0.0248,0.0003]
40	40.06	[40, 60]	[0.9972,0.0028]	10	17.96	[20, 10, 40]	[0.7953,0.2046,0.0001]
20	20.12	[20, 40]	[0.994,0.006]	20	18.87	[20, 10, 40]	[0.8863,0.1135,0.0002]
75	75.00	[75, 60]	[1,0]	75	70.65	[75, 60, 40]	[0.8109,0.1136,0.0755]
40	47.79	[40, 60]	[0.6103,0.3897]	20	20.16	[20, 40, 10]	[0.9919,0.008,0.0001]



Fig. 9. (Color online) Eight small image patches are randomly sampled from the smoke region above the stack for opacity estimation.

4.4 Discussion

Although the experimental results demonstrate the feasibility of the proposed framework, the current dataset still has some limitations. The images were collected from two outdoor experimental sites under controlled observation settings; therefore, the diversity of backgrounds, stack structures, weather conditions, and illumination variations may still be limited compared with fully uncontrolled field environments. Although the use of physically measured opacity labels improves label reliability, the number of original smoke images remains relatively modest. Therefore, the generalization ability of the trained models to other industrial sites, camera devices, and environmental conditions should be further verified. In future work, the dataset will be expanded to include more diverse monitoring sites, broader weather and illumination conditions, and additional smoke-generation scenarios.

5. Conclusions

We proposed a vision-based framework for the automatic estimation of smoke opacity levels from visible-light images for stationary-source monitoring. To reduce the subjectivity and environmental sensitivity of conventional visual inspection, the proposed method integrated smoke-region segmentation, smoke-type classification, patch-based opacity estimation, and final top- k aggregation into a sequential pipeline. By separating black smoke and white smoke before opacity estimation, the framework was designed to better accommodate the different visual characteristics of the two smoke types and to improve robustness under complex outdoor conditions.

The experimental results demonstrated that the proposed framework achieved stable and effective performance on the collected dataset. In the classification-based evaluation, the white-smoke model achieved an average testing accuracy of 92.55%, while the black-smoke model achieved 90.16%. In the top- k weighted evaluation, the continuous opacity estimation accuracies reached 96.0% for white smoke and 90.6% for black smoke within an error tolerance of $\pm 10\%$. These results indicate that the proposed method can provide reliable smoke-opacity estimation under real-world environmental conditions.

The results also showed that the prediction characteristics of black smoke and white smoke were not identical. Under the current experimental setting, white smoke exhibited more stable continuous-opacity estimation results, whereas black smoke showed greater variability in several cases. In addition, the proposed top- k aggregation strategy further improved the evaluation accuracy, demonstrating the importance of this design in the estimation process. The strategy of converting discrete class predictions into continuous opacity estimates makes the proposed framework more suitable for practical smoke monitoring and environmental inspection.

Overall, the proposed framework provides a feasible and regulation-oriented approach for automated smoke-opacity estimation and has potential for future application in environmental inspection and intelligent monitoring systems. In future work, the dataset will be further expanded to include more diverse monitoring sites, weather conditions, illumination settings, and smoke-generation scenarios.

Acknowledgments

This work was supported by the Industrial Technology Research Institute (ITRI), Taiwan. The authors also thank Dr. Pong's research team for providing the training dataset used in this study. This work was partially supported by the National Science and Technology Council (NSTC), Taiwan, under Grant No. NSTC 114-2221-E-197-015.

References

- 1 U.S. Environmental Protection Agency: Visible Emission Field Manual, EPA Methods 9 and 22 (EPA, Washington, DC, 1993).
- 2 ASTM International: ASTM D7520-16, Standard Test Method for Determining the Opacity of a Plume in the Outdoor Ambient Atmosphere (ASTM International, West Conshohocken, PA, 2016).
- 3 Ministry of Environment, Taiwan: Stationary Source Air Pollutant Emission Standards. <https://oaout.moenv.gov.tw/law/LawContent.aspx?id=FL015350> (accessed March 2026).
- 4 Ministry of Environment, Taiwan: Training Guidelines for Air Pollutant and Noise Inspectors. <https://oaout.moenv.gov.tw/law/LawContent.aspx?id=GL006292> (accessed March 2026).
- 5 Ministry of Environment, Taiwan: Executive Yuan Gazette **30** (2024) 2024012.
- 6 Y.-J. Ou-Yang, F. Huang, and B.-J. Pong: Proc. Remote Sensing Technologies and Applications in Urban Environments IX (SPIE, 2024) 131980I. <https://doi.org/10.1117/12.3033766>
- 7 F. Yuan, L. Zhang, X. Xia, B. Wan, Q. Huang, and X. Li: Neurocomputing **357** (2019) 248. <https://doi.org/10.1016/j.neucom.2019.05.011>
- 8 Z. Wang, P. Yang, H. Liang, C. Zheng, J. Yin, Y. Tian, and W. Cui: Remote Sens. **14** (2022) 45. <https://doi.org/10.3390/rs14010045>
- 9 S. Andrade, N. Shafiei, and P. Mehlitz: Comput. Geosci. **204** (2025) 105960. <https://doi.org/10.1016/j.cageo.2025.105960>
- 10 S. Yan, J. Zhang, and N. Barnes: Proc. AAAI Conf. Artif. Intell. **36** (2022) 3009. <https://doi.org/10.1609/aaai.v36i3.20207>
- 11 L. Yao, H. Zhao, J. Peng, Z. Wang, and K. Zhao: Adv. Neural Inf. Process. Syst. **37** (2024) 106743.
- 12 L. Zhang, J. Wu, F. Yuan, and Y. Fang: IEEE Trans. Image Process. **33** (2024) 1175. <https://doi.org/10.1109/TIP.2024.3359816>
- 13 L. Yao, H. Zhao, J. Peng, Z. Wang, and K. Zhao: Proc. The European Conference on Computer Vision (ECCV 2024) 467. https://doi.org/10.1007/978-3-031-72784-9_26
- 14 T.-H. Kim, E.-S. Seo, and S.-H. Choi: Sens. Mater. **37** (2025) 3621. <https://doi.org/10.18494/SAM5716>
- 15 M. Lin, C. Zhang, W. Liu, and H.-W. Tseng: Sens. Mater. **37** (2025) 3339. <https://doi.org/10.18494/SAM5786>
- 16 O. D. Pedrayes, R. Usamentiaga, and D. F. García: Process Saf. Environ. Prot. **170** (2023) 479. <https://doi.org/10.1016/j.psep.2022.12.023>
- 17 S. Wang, Y. Han, M. Yu, H. Wang, Z. Wang, G. Li, and H. Yu: J. Mar. Sci. Eng. **11** (2023) 1945. <https://doi.org/10.3390/jmse11101945>
- 18 M.-H. Park, J.-H. Choi, and W.-J. Lee: J. Environ. Manage. **393** (2025) 127139. <https://doi.org/10.1016/j.jenvman.2025.127139>
- 19 W. Yuen, Y. Gu, Y. Mao, P. M. Kozak, S. Koloutsou-Vakakis, H.-K. Son, K. Mattison, B. Franek, and M. J. Rood: Environ. Technol. Innov. **12** (2018) 43. <https://doi.org/10.1016/j.eti.2018.07.003>
- 20 W. Yuen, Y. Gu, Y. Mao, S. Koloutsou-Vakakis, M. J. Rood, H.-K. Son, K. Mattison, B. Franek, and K. Du: Aerosol Air Qual. Res. **17** (2017) 1281. <https://doi.org/10.4209/aaqr.2016.08.0369>

About the Authors



Fay Huang received her B.S., M.S., and Ph.D. degrees from the University of Auckland, New Zealand, in 1994, 1996, and 2002, respectively. She was a Postdoctoral Fellow at Academia Sinica, Taiwan, in 2003 and also worked at the Industrial Technology Research Institute, Taiwan, in 1997. Since 2006, she has been with the Department of Computer Science and Information Engineering, National Ilan University, Taiwan, where she is currently an associate professor. Her research interests include computer graphics and computer vision. (fay@niu.edu.tw)



Hong-Yi Jang received his B.S. degree in computer science and information engineering from National Ilan University, Taiwan, in 2024. He is currently pursuing his M.S. degree in computer science and information engineering at National Ilan University, Taiwan, where he is a member of the Graphics and Vision Laboratory. His research interests include computer vision and computer graphics. (rl443016@ems.niu.edu.tw)



Yuan-Jen Ou-Yang received his B.S. and M.S. degrees in computer science and information engineering from National Ilan University, Taiwan, in 2023 and 2025, respectively. He is currently pursuing his Ph.D. degree at National Yang Ming Chiao Tung University, Taiwan. His research interests include diffusion models, computer vision, and generative modeling. (oyyj.mail@gmail.com)



Bao-Jen Pong received his B.S., M.S., and Ph.D. degrees from the National Central University, Taiwan in 1986, 1988, and 1994, respectively. From 2000 to 2003, he was a manager in a fiber optics company. From 2006 to 2016, he was a senior engineer at ITRI, Taiwan. Since 2016, he has been a principal engineer at ITRI, Taiwan. His research interests are in co-package optics, displays, microLEDs, and sensors. (bjpong@itri.org.tw)