FireSpot: A Database for Smoke Detection in Early-stage Wildfires

Natthaphol Pornpholkullapat
Sirindhorn International Institute of Technology, *Thammasat University*, Pathum Thani, Thailand

Warit Phankrawee
Sirindhorn International Institute of Technology, *Thammasat University*, Pathum Thani, Thailand

Peraphat Boondet
Falculty of Engineering, Chiang Mai University, Chiang Mai, Thailand

Thin Lai Lai Thein
University of Computer Studies,Yangon, Myanmar

Phoummixay Siharath
National University of Laos, Lao PDR

Jennifer Dela Cruz
Mapua University, Philippines

Ken T. Marata
National Institute of Information and Communications, Tokyo, Japan

Kanokvate Tungpimolrut
NECTEC, *National Science and Technology Development Agency*, Pathum Thani, Thailand
kanokvate.tungpimolrut@nectec.or.th

Jessada Karnjana
NECTEC, National Science and Technology Development Agency, Pathum Thani, Thailand
jessada.karnjana@nectec.or.th

***Abstract*—During a dry season in northern Thailand, typically between November and May, wildfires, either natural or man-made, annually occur and pose severe air pollution, particularly concerning PM2.5 particles. A no-burning period policy has been deployed to prevent and mitigate the effect. However, relying solely on traditional monitoring methods using forest fire lookout towers operated by human staff is insufficient due to limited human resources. Therefore, this project aims to develop a monitoring system based on visual imagery to detect smoke in early-stage wildfires. Based on our survey, no available database is suitable for practically implementing a smoke detection model for an early-stage forest fire warning system. Thus, we construct it by ourselves and report the detail of its construction in this study. We also developed a smoke detection model based on YOLOv5 to demonstrate its practical use. The model's performance evaluation shows that the accuracy is 93.88%, the precision is 97.40%, the recall is 93.75%, and the F1-score is 95.54%, which are promising results initially.**

***Index Terms*—Smoke detection, Forest fire detection, Dataset construction, YOLOv5**

I. Introduction

In a dry season in upper northern Thailand, ranging from November to May, forest fires occur commonly and annually, either natural or human-caused [1], [2]. Natural wildfires or agricultural burning pose severe air pollution, particularly concerning PM2.5 particles [3], [4]. Wildfires are among the significant risks to living organisms in local and global ecosystems, and uncontrollable ones spread rapidly and are challenging to handle and manage [5]–[7]. Even though the government has implemented a no-burning or zero-burning policy to prevent and mitigate the effect [8], relying solely on traditional monitoring methods using forest fire lookout towers operated by human staff is insufficient due to the lack of human resources. Therefore, information and communication technologies, such as remote sensing, computer vision, the Internet of Things, artificial intelligence, and machine learning, have been applied to address many issues concerning forest fire detection systems [5]–[7], [9]–[12].

In general, wildfire detection systems can be divided into three categories: terrestrial-based (e.g., wireless sensor networks and cameras or multi-spectral sensors installed at some buildings or towers), aerial-based (e.g., sensors mounted on unmanned aerial vehicles), and satellite-based systems [6]. The satellite-based system can cover a large area but has less temporal and spatial resolution. The airborne system can access any inaccessible site with more efficiency in accuracy but is limited by workforce budget and not practical for real-time monitoring. The terrestrial-based system is most efficient in terms of accuracy and response time [6].

In this study, we focus on the terrestrial- and machine-vision-based one, not only because of its effectiveness but also because it is a straightforward step from a conventional method of using forest fire lookout towers operated by human staff. That is, cameras and computer vision techniques are used to replace humans for some tedious tasks. More specifically, our project aims to detect a wildfire from its early stage to notify in-charge staff to respond quickly and adequately. The fire life cycle consists of four stages: ignition, growth, maturity, and decay [6]. We target the first or the beginning of the second stage, i.e., before the fire spreads extensively and is fueled by vegetation.

In literature, fire pixel detection in machine vision can be done in three ways: using color rules or models or spaces, motion analysis, and machine learning [5]. The last one has gained popularity recently due to its efficiency [7], [9]–[11], [13]. However, machine learning algorithms have a characteristic that their performance strongly depends on the quality of training data. Unfortunately, the publicly available wildfire dataset is limited [6]. For example, a web database called ForestryImages.org has no ground truth [14]. A social-media-based database contains about $2,000$ images in different environments [15]. Still, the settings or backgrounds are too different to represent forest fires. Even though many public fire databases are available [16]–[19], it should be noted that most prioritize images with bonfires or flames, which are unsuitable for practically implementing a smoke detection model for an early-stage forest fire warning system. Therefore, at the beginning step of this project, we constructed a smoke database that reasonably represents wildfires in an early stage, and the construction detail is reported in this paper with a showcase.

Wildfires are a common problem worldwide, and the topography and forests of ASEAN countries are similar. For this reason and from the initial idea that techniques and databases developed among ASEAN countries can be shared, we initiated a collaborative research project among four ASEAN countries (Thailand, Myanmar, Lao PDR, and the Philippines) and Japan under the ASEAN IVO framework to address the problem. The rest of this paper is organized as follows. Section II provides detail of the database construction methodology. Section III describes the implementation and evaluation of smoke detection to showcase the practical use of our database. Discussion and conclusion are made in Section IV and Section V, respectively.

II. Database Construction Method

The FireSpot database is developed based on a collaboration among National Electronics and Computer Technology Center (NECTEC) and three local municipalities, including Pa Miang, Nong Yaeng, and Choeng Doi, in Chiang Mai, Thailand. In the current release, it consists of $4,000$ images. Half of them contain smoke in the early burning stages, and another half do not. Smoke areas in those images are labeled with bounding boxes by the authors, as shown in Fig. 1. The bounding box values are a quadruple (x_c, y_c, w, h), where (x_c, y_c) is the Cartesian coordinates of the point at the box center, w is the box width, and h is the box height. The bounding box values are normalized to the maximum of 1, i.e., range from 0 to 1, of which the point at $(0, 0)$ represents the top-left corner, and the point at $(1, 1)$ represents the bottom-right corner of the image.

The FireSpot database construction procedure consists of the following activities. First, the seasoned local government staff responsible for managing and handling wildfires in the area listed locations where actual wildfires had occurred and suggested prospective areas for our experiment and database construction. Five sites around mountainous areas were chosen at the end of the first activity: Huai Huk (18°55′29″N, 99°05′32″E), Pa Miang (18°54′50″N, 99°13′42″E), Sala Pang Sak (18°53′59″N, 99°11′55″E), Doi Koo (plain side) (18°53′07″N, 99°10′10″E), and Doi Koo (mountain side) (18°53′07″N, 99°10′15″E), as shown in Fig. 2.

Second, for each site, the authors and the local government staff discussed locations for setting fires and places for capturing smoke photos. There were five to eight fire spots for each site, each located about two to five hundred meters from one another. A team of experts and local rescuers intentionally set and managed the fire. The bonfire was approximately a meter in diameter and burned for about 15 minutes per fire spot. Straw bales, twigs, and dry leaves were used as fuel. Before, during, and after the burning, six or seven camerapersons distributed separately took photos from different angles for the database. Each cameraperson was asked to capture at least 50 shots at each fire spot. The camerapersons and the fire-setting team communicated using walkie-talkies. An example of fire spots and cameraperson positions is shown in Fig. 3, and an example of activities is presented in Fig. 4.

It should be noted that these activities were conducted from the 6th to the 16th of June, 2023, which falls outside the dry season in Thailand. The period is not ideal for capturing smoke images that may represent an early stage of forest fires in actual conditions. The dry season in northern Thailand is typically between November and May, during which wildfires occur commonly and annually [1], [2]. However, June marks the beginning of heavy rain. Nevertheless, due to a no-burning policy deployed in Chiang Mai, Chiang Rai, Lamphun, Lampang, Mae Hong Son, and Nan, as measures to reduce air pollution, setting up the experiment in the dry season is legally prohibited [20]. For this reason, our database construction activities were conducted in June.

In the end, there were 29 fire spots, and more than 14,642 photos with and without smoke were captured. For the first release of our database, $4,000$ images were randomly selected from 29 fire spots equally. The ratio between smoke and non-smoke images is approximately $68\% : 32\%$. Hence, out of $4,000$, there are 2,511 smoke images and $1,189$ non-smoke images.

III. Implementation and Evaluation of a Smoke Detection Based on YOLOv5

This section presents an example of implementing and evaluating a smoke detection based on YOLOv5 for early-stage wildfire detection to demonstrate the practical use of

Fig. 1: Example of images with bounding boxes in FireSpot: images with smoke (top-left and top-middle), images without smoke (bottom-left and bottom-center), and image with a bounding box quadruple (x_c, y_c, w, h).

Fig. 2: Five experimental sites in Chiang Mai, Thailand.

our FireSpot database. We first briefly review YOLOv5 and then describe our implementation and performance evaluation of the detector.

A. YOLOv5

YOLO, which stands for You Only Look Once, is a real-time object detection algorithm that has gained immense popularity among researchers and developers in the computer vision community [21], [22]. Its key advantages include being extremely fast and globally reasoning about the image when making predictions [21]. Also, it exhibits an impressive balance between speed and accuracy, achieving state-of-the-art performance in real-time scenarios [22]. For this reason, this study started with YOLO to showcase an application developed on our database. Over the years since its origin, YOLO has been developed continuously. There are many versions, and the eighth version (YOLOv8) has recently been released. In this study, we implemented our model based on the fifth version (YOLOv5) because it is already well-established and has a large community of users and developers compared to the newer ones. The framework and architecture of YOLOv5 are shown in Fig. 5 [23].

B. Implementation and Evaluation

The YOLOv5 was trained, validated, and tested on our FireSpot database to demonstrate a smoke detection model. The database was split into training, validation, and test sets with a ratio of $60\%:20\%:20\%$, respectively, and 4-fold cross-

Fig. 3: Eight fire spots (red-shaded area) and six cameraman positions (yellow-shaded area) at the Huai Huk site.

Fig. 4: FireSpot database construction activities.

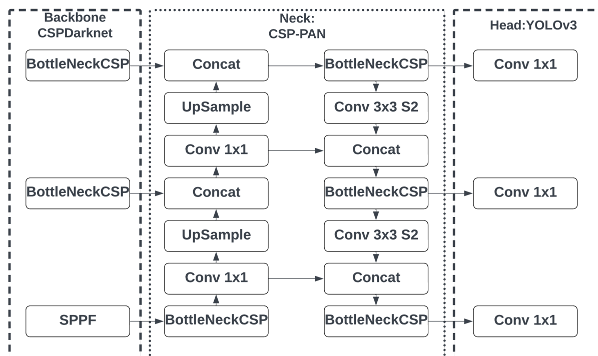

Fig. 5: Framework and architecture of YOLOv5.

validation was applied. The hyperparameters of YOLOv5 were set as follows. The initial learning rate and the final learning rate are both 0.01. The optimizer in YOLOv5 is a stochastic gradient descent with momentum, of which the optimizer weight decay is 0.0005, the momentum factor is 0.937 with a warmup momentum of 0.8, and the warmup epoch is 3.0. The warmup initial bias learning rate is set to 0.1. The gains for the box loss, classification loss, and objectiveness loss are 0.05, 0.5, and 1.0, respectively. The positive weights in the classification and objectiveness cross-entropy loss functions are both set to 1.0. The IoU training threshold is 0.2. The value of the anchor-multiple threshold used is the default one, which is 4.0. The rest are set to default values. The model was trained for 40 epochs.

An example of plots of box loss, objectiveness loss, precision, recall, and mean average precision (mAP) over training epochs for training and validation sets is shown in

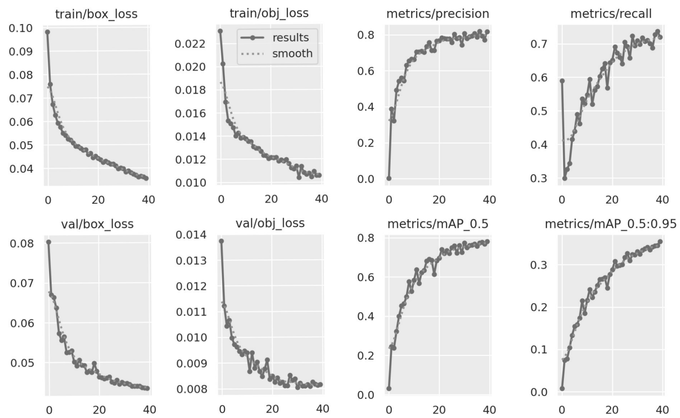

Fig. 6: Plots of box loss, objectiveness loss, precision, recall, and mean average precision (mAP) over training epochs for training and validation sets.

TABLE I: Evaluation metrics of 4-fold cross-validation.

Round	Precision	Recall	$\mathrm{mAP}_{0.5}$	$\mathrm{mAP}_{0.5:0.95}$	F1-score
1	**0.8171**	0.7198	**0.7809**	**0.3547**	**0.9688**
2	0.7749	**0.7228**	0.7693	0.3435	0.9605
3	0.8055	0.7207	0.7772	0.3332	0.9700
4	0.7631	0.6948	0.7364	0.3359	0.9616

Fig. 6. This example is from the first fold. There are two mean average precisions, i.e., $\mathrm{mAP}_{0.5}$ and $\mathrm{mAP}_{0.5:0.95}$, where the former measures the mean average precision at an intersection over union (IoU) of 0.5 and the latter measure that at a range of IoU thresholds from 0.5 to 0.95. It can be seen that both box and objectiveness losses (i.e., **train/box_loss**, **train/obj_loss**, **val/box_loss**, and **val/obj_loss**, where '**train**' and '**val**' denote the training and validation sets, respectively) properly converge toward zero as the number of epochs increases. The classification loss, which is not shown in the figure, is always zero because there is only one class. The evaluation metrics for all four rounds from the 4-fold cross-validation, including the F1-score, defined as the harmonic mean between precision and recall, are presented in Table I. It can be seen that the performances of all four models are comparable, and the model from the first round of 4-fold cross-validation performs slightly better than the others.

Since, in practice, the surveillance camera is installed at the top of a tower and captures images at every predefined interval, the classification task takes precedence over the detection task. It is more critical to correctly classify an input image containing smoke than to detect the exact location of smoke precisely. Therefore, for evaluation purposes, we assess the model's performance on the test set using classification evaluation metrics with the following conditions:

1) Given an image with smoke as the input, it is counted as a true positive if the model detects at least one smoke area with an IoU value larger than a predefined threshold; otherwise, it is considered a false negative.

TABLE II: Performance evaluation of our YOLOv5-based smoke detection.

IoU Threshold	Accuracy	Precision	Recall	F1-score	Balanced Accuracy
0.3	**0.9388**	**0.9740**	**0.9375**	**0.9554**	**0.9396**
0.4	0.9375	0.9740	0.9357	0.9545	0.9387
0.5	0.9375	0.9740	0.9357	0.9545	0.9387
0.6	0.9363	0.9739	0.9339	0.9535	0.9378
0.7	0.9338	0.9738	0.9304	0.9516	0.9360

2) Given an image without smoke, it is counted as a true negative if the model detects no smoke objects; otherwise, it is considered a false positive.

We evaluated the model's performance using different IoU thresholds ranging from 0.3 to 0.7 with a step of 0.1, and the results are presented in Table II. It can be seen that the performance does not significantly vary with the IoU threshold. At the IoU threshold of 0.3, the accuracy is 93.88%, the precision is 97.40%, the recall is 93.75%, the F1-score is 95.54%, and the balanced accuracy is 93.96%. The accuracy decreases as the IoU threshold increases since the number of true positives reduces. In contrast, the number of true negatives does not change. The number of true positives reduces because the criterion for being a true positive is stricter, i.e., the IoU value must be large enough to be counted as a true positive. On the other hand, the more stringent criterion has nothing to do with those being counted as true negatives since the actual input image does not contain smoke. By the same token, as the IoU threshold increases, the number of false negatives increases because of the same more stringent criterion. Consequently, the recall reduces.

An example of predicted bounding boxes is shown in Fig. 7. It can be seen that, in some cases, the input image has at least one smoke area, and the model predicts more than one smoke area. In such cases, the image is a true positive according to the criterion.

IV. Discussion

This section discusses some limitations of the FireSpot database and the general idea of detecting smoke in images captured in the visible light spectra. First, it's essential to note that this database was constructed based on the initial concept of addressing specific forest fire prevention and mitigation issues, such as the operational cost of forest fire lookout towers and the scarcity of human resources in certain areas. As a result, the database includes images from actual environments where wildfires commonly occur in Chiang Mai, Thailand. For this reason, the database may not be representative of other topography and climate.

Second, even though achieving more than 90% accuracy is good enough for installing the system in some remote and prone areas for the short-handed forest fire lookout tower, it's important to note that visible-light-spectrum images have limitations, particularly in night-time monitoring. Detecting smoke at night is challenging. To overcome this limitation, the use of multispectral cameras can be helpful. As part of

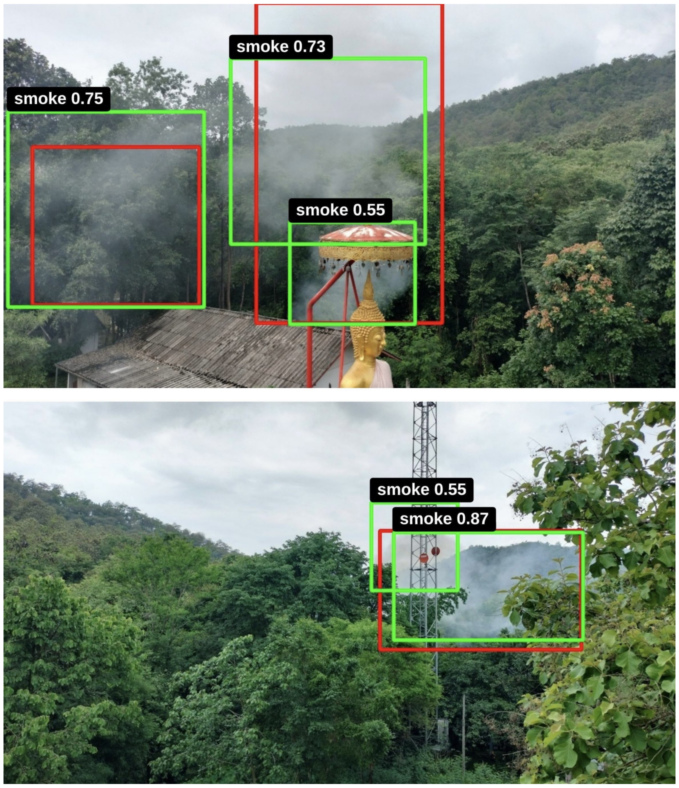

Fig. 7: Example predicted bounding boxes (green) with confidence scores compared to ground truth (red).

our future plans, we aim to expand the FireSpot database to include multispectral images to enhance night-time monitoring capabilities.

V. Conclusion

In this paper, we presented the construction of the FireSpot database, specifically designed for smoke detection in early-stage forest fires. The first version of the database consists of $4,000$ images, of which $2,511$ are smoke images, and $1{,}189$ are non-smoke images. Those images were captured from 29 fire spots around five experimental sites in Chiang Mai, Thailand. We developed a smoke detection model based on YOLOv5 to showcase the practical use of our FireSpot database. The performance evaluation demonstrated promising results, with an accuracy of 93.88%.

Acknowledgment

The ASEAN IVO project (http://www.nict.go.jp/en/asean_ivo/index.html), titled 'Visual IoT Network for Environmental Protection and Disaster Prevention,' was involved in the production of the contents of this work and financially supported by NICT (http://www.nict.go.jp/en/index.html).

References

[1] M. Sukitpaneenit and N. T. Kim Oanh, "Satellite monitoring for carbon monoxide and particulate matter during forest fire episodes in northern thailand," *Environmental monitoring and assessment*, vol. 186, pp. 2495–2504, 2014.

[2] S. Chantara, S. Sillapapiromsuk, and W. Wiriya, "Atmospheric pollutants in chiang mai (thailand) over a five-year period (2005–2009), their possible sources and relation to air mass movement," *Atmospheric Environment*, vol. 60, pp. 88–98, 2012.

[3] A. Uttajug, K. Ueda, X. T. Seposo, A. Honda, and H. Takano, "Effect of a vegetation fire event ban on hospital visits for respiratory diseases in upper northern thailand," *International journal of epidemiology*, vol. 51, no. 2, pp. 514–524, 2022.

[4] P. Prapatigul and S. Sreshthaputra, "Causes and solution of forest and agricultural burning in northern, thailand."

[5] T. Toulouse, L. Rossi, A. Campana, T. Celik, and M. A. Akhloufi, "Computer vision for wildfire research: An evolving image dataset for processing and analysis," *Fire Safety Journal*, vol. 92, pp. 188–194, 2017.

[6] P. Barmpoutis, P. Papaioannou, K. Dimitropoulos, and N. Grammalidis, "A review on early forest fire detection systems using optical remote sensing," *Sensors*, vol. 20, no. 22, p. 6442, 2020.

[7] R. Xu, H. Lin, K. Lu, L. Cao, and Y. Liu, "A forest fire detection system based on ensemble learning," *Forests*, vol. 12, no. 2, p. 217, 2021.

[8] I. Kumar, V. Bandaru, S. Yampracha, L. Sun, and B. Fungtammasan, "Limiting rice and sugarcane residue burning in thailand: Current status, challenges and strategies," *Journal of Environmental Management*, vol. 276, p. 111228, 2020.

[9] S. Khan, K. Muhammad, T. Hussain, J. Del Ser, F. Cuzzolin, S. Bhattacharyya, Z. Akhtar, and V. H. C. de Albuquerque, "Deepsmoke: Deep learning model for smoke detection and segmentation in outdoor environments," *Expert Systems with Applications*, vol. 182, p. 115125, 2021.

[10] S. Saponara, A. Elhanashi, and A. Gagliardi, "Real-time video fire/smoke detection based on cnn in antifire surveillance systems," *Journal of Real-Time Image Processing*, vol. 18, pp. 889–900, 2021.

[11] C. Li, B. Yang, H. Ding, H. Shi, X. Jiang, and J. Sun, "Real-time video-based smoke detection with high accuracy and efficiency," *Fire Safety Journal*, vol. 117, p. 103184, 2020.

[12] K. Muhammad, S. Khan, V. Palade, I. Mehmood, and V. H. C. De Albuquerque, "Edge intelligence-assisted smoke detection in foggy surveillance environments," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 2, pp. 1067–1075, 2019.

[13] A. B. Abdusalomov, B. M. S. Islam, R. Nasimov, M. Mukhiddinov, and T. K. Whangbo, "An improved forest fire detection method based on the detectron2 model and a deep learning approach," *Sensors*, vol. 23, no. 3, p. 1512, 2023.

[14] Forestry images. (Accessed 4-Aug-2023). [Online]. Available: https://www.forestryimages.org/

[15] M. Bedo, G. Blanco, W. Oliveira, M. Cazzolato, A. Costa, J. Rodrigues, A. Traina, and C. Traina Jr, "Techniques for effective and efficient fire detection from social media images," *arXiv preprint arXiv:1506.03844*, 2015.

[16] Bowfire dataset. (Accessed 4-Aug-2023). [Online]. Available: https://bitbucket.org/gbdi/bowfire-dataset/downloads/

[17] D-fire. (Accessed 4-Aug-2023). [Online]. Available: https://github.com/gaiasd/DFireDataset

[18] Forest fire. (Accessed 4-Aug-2023). [Online]. Available: https://www.kaggle.com/datasets/kutaykutlu/forest-fire

[19] roboflow: collect computer vision project. (Accessed 4-Aug-2023). [Online]. Available: https://universe.roboflow.com/forest-fire-journal/collect-nrv3o

[20] P. Junlapeeya, T. Lorga, S. Santiprasitkul, and A. Tonkuriman, "A descriptive qualitative study of older persons and family experiences with extreme weather conditions in northern thailand," *International Journal of Environmental Research and Public Health*, vol. 20, no. 12, p. 6167, 2023.

[21] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, *You only look once: Unified, real-time object detection*, 2016. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_2016/html/Redmon_You_Only_Look_CVPR_2016_paper.html

[22] J. Terven and D. Cordova-Esparza, "A comprehensive review of yolo: From yolov1 to yolov8 and beyond," *arXiv preprint arXiv:2304.00501*, 2023.

[23] I. S. Isa, M. S. A. Rosli, U. K. Yusof, M. I. F. Maruzuki, and S. N. Sulaiman, "Optimizing the hyperparameter tuning of yolov5 for underwater detection," *IEEE Access*, vol. 10, pp. 52 818–52 831, 2022.