ISBN: 978-981-94-1156-6
2024 the 14th International Workshop on Computer Science and Engineering (WCSE 2024)
doi: 10.18178/wcse.2024.06.008

Identification of Smoke and Fire from a Low-Light Enhanced Synthetic Dataset Using YOLOv7

John Paul Q. Tomas [1+], Von Derwin R. Caibigan [2], Erin Beatrice Micaela G. Reyes [3]

[123] School of Information Technology, Mapua University, Philippines

Abstract. This paper investigates the utilization of YOLOv7, a state-of-the-art object detection system, in identifying instances of smoke and fire using a synthetic low-light enhanced dataset. With fire and smoke incidents causing significant casualties annually, the study aims to enhance the efficacy of fire detection systems. Through extensive experimentation and analysis, the study compares the performance of YOLOv7 on original and low-light enhanced datasets. Results indicate that while the model shows promise, the low-light enhancements, particularly Contrast Limited Adaptive Histogram Equalization (CLAHE), outperform deep learning-based enhancements. Further research avenues include exploring more sophisticated low-light image enhancement techniques, employing diverse and balanced datasets, and deploying the model in real-world settings for comprehensive evaluation and practical applications. This study contributes to advancing non-urban fire detection systems, thereby bolstering public safety measures and disaster management strategies.

Keywords: Fire Detection, Smoke Detection, Wildfire Detection, Computer Vision, YOLOv7, Vision Sensor-Based Techniques, Low-Light Conditions, Synthetic Dataset, Object Detection.

1. Introduction

Statistics from 2021 show that fire and smoke have caused the deaths of 3,389 people in the United States alone [10]. With the rapid development of computer vision technologies, it is now possible to accurately and swiftly identify smoke and fire incidents to minimize potential damage and save lives. While smoke detectors are currently the most popular means of detecting fire, a study done in 2019 found that vision sensor-based techniques could be able to decrease response time, increase the probability of an enhanced detection process, and cover a larger area as vision sensors can collect and describe more data regarding the direction, growth, and size of fire and smoke [1].

As the demand for more robust and efficient systems grows, so does the need for innovative approaches that are made to operate in less-than-ideal lighting conditions effectively. YOLOV7 is a state-of-the-art object detection system that stands out due to its real-time processing speed and accuracy, an ideal candidate for situations in which timely identification is crucial. The research will also be utilizing a synthetic low-light enhanced dataset to provide a controlled environment for training and testing the model. The benefit of a synthetic dataset over a traditional one is that it allows for the tracking and identification of dynamic elements such as smoke and fire that may be challenging to collect data on in real-world scenarios.

This paper aims to utilize YOLOv7 to identify instances of smoke and fire using a synthetic dataset that has been low-light enhanced to contribute to the ongoing efforts to enhance the efficacy of fire detection systems, ultimately bolstering the capabilities of emergency response teams and promoting public safety.

2. Review of Related Literature

This chapter presents an overview of related literature and studies on synthetic datasets and fire-smoke detection. Specifically, it aims to provide insight into the following concepts: The use of synthetic datasets in computer vision and visual-based fire detection systems.

+ Corresponding author. Tel.: + 639178695257
E-mail address: jpqtomas@mapua.edu.ph (J.P.Q. Tomas).

2.1. Synthetic Datasets in Computer Vision

In a paper by Manettas et al. [6], the author discusses the main motivation in the use of synthetic datasets for deep learning models. Utilizing synthetic datasets addressed the challenges of acquiring a large volume of real-world data, which may be expensive and time-consuming. The paper generated synthetic CAD-based images, which they utilized for training image classification machine-learning algorithms. The study found that synthetic data for classification was successful and highlighted the approach's efficiency and simplicity in developing machine learning models.

A review by Man and Chahl in 2022 [5] supports the idea that synthetic data can be recognized as an alternative to real data, especially in cases where collecting real data can be expensive. Furthermore, the study talks about the potential applications of synthetic data, with one of them being for generating specific environmental conditions that may be challenging to capture in the real world. The study also highlights that the main gap in using synthetic data is the quality of generated data and a potential lack of diversity or photorealism.

Another study brought up some of the previously mentioned points while proposing three main directions for synthetic data [8]. Due to legal or privacy-based issues, there is the possibility of using synthetic data to augment existing datasets, using fully synthetic datasets for training, and employing synthetic data.

2.2. Vision-Based Fire Detection Systems

The paper "Fire and Smoke Detection Without Sensors: Image Processing Based Approach" [2] outlines the novel approach for using image processing as a means of fire and smoke protection. The paper focuses on color information as a pre-processing step for the model. YCbCr color space. The paper also proposed using fuzzy logic to improve accuracy and avoid false alarms.

Another study involves the use of RGB and HSI models to detect smoke and fire at the earliest possible stage [3]. The study explains the RGB models, which add red, green, and blue light in various ways, as well as the significance of separating luminance from chrominance in the extracted YCbCr model. It also introduced the Hue, Saturation, Intensity (HSI) model, which is a fast and practical fire detection method.

YOLOv2, a predecessor of YOLOv7, was used in a paper that presented a real-time video-based fire and smoke detection system [9]. The paper emphasized the potential of YOLOv2 as a suitable and low-cost model that is suitable for real-time surveillance due to its lightweight and state-of-the-art models. The algorithm for the model was described in detail, and the proposed method managed to achieve good results with its classification performance.

3. Methodology

This section discusses the methodology used in developing a model for the Identification of Smoke and Fire from a Low-Light Enhanced Synthetic Dataset Using YOLOv7. Specifically, it discusses the dataset used, data pre-processing and augmentation, and model development and testing.

3.1. Dataset

Based on the review of [13] in regard to open access datasets for fire detection, it shows that the most cited and popularly used datasets do not contain an adequate amount of positive and negative night-time fire videos. With this information, the researchers decided to look towards synthetic datasets to address the scarcity of data. The researchers settled on using MSFFD, a dataset created by [14] containing 3974 synthetic images, available in Roboflow. With the use of Unreal Engine 5, it simulates a diversified forest environment. It comes with eight multi-scale forest scenarios with varying terrain and vegetation, weather scenes, time of day, and different numbers of fire objects. Table 1 shows the fire and smoke distribution details for each multimodal scenario.

Table 1. Details of Fire and Smoke Distribution

Multi-Modal Scenarios	No. of images	Fire	Smoke	No. of annotation boxes

Sunny	3292	8423	7095	15518
Snowy	312	772	547	1319
Rainy & Fog	370	432	494	926
Daytime	2209	4584	3807	8491
Evening	560	2052	1581	3633
Night	1205	2891	2748	5639

3.2. Data Pre-processing and Data Augmentation

One of the challenges of fire and smoke detection using image processing is the varying lighting conditions that may affect the quality and visibility of the images [13, 15]. To address this issue, the study applied two techniques separately to enhance the contrast and brightness of the synthetic images: Contrast Limited Adaptive Histogram Equalization (CLAHE) and Low-Light Image Enhancement using MIRnet. CLAHE is a method that improves the contrast of an image by dividing it into small regions and applying histogram equalization to each region while limiting the amplification of noise. MIRnet is a deep-learning model that enhances low-light images by restoring the scene's illumination, color, and details. The following sections will describe these techniques in more detail and show their effects on the synthetic images.

3.2.1. Contrast Limited Adaptive Histogram Equalization

CLAHE is a traditional technique for enhancing the contrast of images in low-light conditions. It divides the image into small regions, called tiles, and applies histogram equalization to each region. The neighboring tiles are then blended using bilinear interpolation to remove the artificial boundaries. Thereby improving the visibility and quality of images that suffer from poor lighting or uneven illumination. [16] used CLAHE to recognize road signs at night using YOLO Models, providing the best results in comparison to Contrast Stretching (CS) and Histogram Equalization (HE).

Fig 1: Sample image before CLAHE application

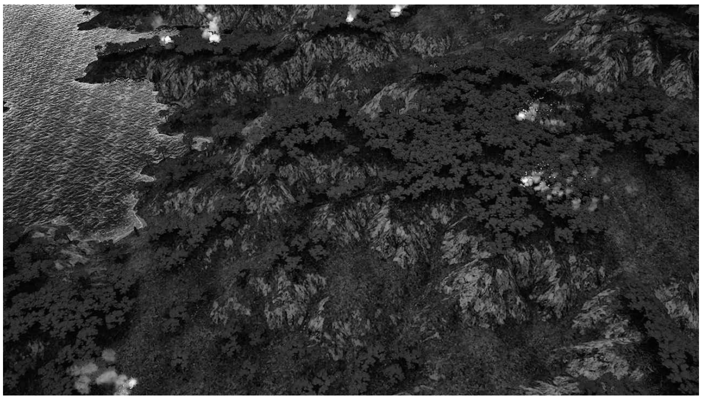

Fig 2: Sample image after CLAHE application

3.2.2. Low-light Image Enhancement Using MIRnet

MIRnet is a deep learning solution for enhancing image contrast, particularly in low-light conditions. MIRnet employs CNNs to process and enhance images efficiently and is inspired by methods such as histogram equalization. This technique not only improves the quality and visibility of an image but also addresses the issue of artificial boundaries, which is a common issue in traditional techniques. In a study [17], it is emphasized that MIRnet performs comparatively better in metrics such as PSNR, SSIM, NIQE, BRISQUE, and LOE compared to other models. This suggests its effectiveness at preserving detailed textures while maintaining the naturalness of the image.

Given these two Low-Light Image enhancement methods, the dataset's images containing the label ‘night’ will be modified while the rest of the images remain the same. Thus, two datasets, one augmented with CLAHE and the other with MIRnet, are to be compared against the original dataset. These augmentations will help train a model capable of detecting fire and smoke during low-light scenarios.

Fig 3: Sample image before MIRnet application

Fig 4: Sample image after MIRnet application

3.3. Model Development and Testing

YOLOv7 is a state-of-the-art real-time single-stage object detector that exceeds existing object detectors in both accuracy and speed [18]. With the latest addition of the 'bag of freebies' training technique, the model is able to learn from level features much better than previous iterations. Recently released in 2022, it has been applied in several object detection tasks such as corn pests identification [19], automotive running light defect detection [20], and license plate detection [21] with promising results. Additionally, [22-25] are a handful of studies that utilize YOLOv7 for low-light detection tasks, particularly night time vehicle detection, foggy traffic environment, object detection, and people & vehicle detection. Given the nature and format of the data (as Roboflow includes a YOLOv7 format), the problem of low-light detection, and the promising performance, the researchers selected YOLOv7 as the model for low-light detection of synthetic fire and smoke.

Table 2. Performance Evaluation of Baseline Model in Comparison with Augmented Datasets.

Dataset		**Evaluation Index**				
		mAP@.5	mAP@.95	Precision	Recall	F1-Scores
MSFFD	All	0.848	0.434	0.823	0.769	0.795
	Fire	0.858	0.381	0.848	0.786	0.815
	Smoke	0.838	0.487	0.799	0.751	0.774
MSFFD CLAHE	All	0.619	0.232	0.707	0.551	0.619
	Fire	0.674	0.242	0.707	0.637	0.67
	Smoke	0.564	0.223	0.708	0.564	0.627
MSFFD MIRnet	All	0.607	0.219	0.651	0.582	0.614
	Fire	0.717	0.247	0.672	0.709	0.69
	Smoke	0.496	0.191	0.63	0.455	0.528

4. Results and Discussion

In this section, the researchers present and discuss the results of the study where they applied the YOLOv7 model for fire and smoke detection on a synthetic dataset. The approach is novel in that the researchers incorporated Low Light Image (LLI) enhancement techniques to improve the model's performance in low-light or nighttime environments. The results provide insights into the effectiveness of YOLOv7 in detecting non-urban fire and smoke under these challenging conditions and highlight the benefits of using LLI enhancements in such scenarios. The researchers will delve into the specifics of their findings, discuss the implications, and explore potential avenues for future research.

4.1. Results

Trained on 2802 synthetic images, with 564 of them classified as a low-light scenario or night time, Table 2 presents the results validated on the test sets of the original dataset and the two datasets with LLI enhancement on the 564 images. The baseline model presents the highest result across all evaluation metrics when trained under 10 epochs and with 12 batch sizes. With an 84.8% mean Average Precision with a 50% threshold for fire and smoke is 85.8 % and 83.8%, respectively. Meanwhile, the precision, recall, and f1-scores for fire are as follows: 84.8%, 78.6%, and 81.5%. The metrics for smoke in order of precision, recall, and f1-scores are 79.9%, 75.1%, and 77.4%.

Based on the table, the augmented datasets paled in comparison with the baseline, with the LLI enhancement through MIRnet proving to be the worst model among the three instances when comparing each metric. When compared to the precision, recall, and f1- scores of the baseline model, the fire percentage decrease is 20.75%, 9.79%, and 15.33%, respectively. For smoke, there was a decrease of 21.15%, 39.41%, and 31.78% in the same order. Meanwhile, its mAP@.5 and mAP@.95 were 71.7% & 24.7% for fire and 49.6% & 19.1% for smoke detection. When compared to the CLAHE augmented model, it outperformed the mAP@.5 and mAP@.95 for fire detection scores slightly by 0.043 and 0.005, respectively. However, it performed worse for smoke detection, being outperformed by the CLAHE-augmented model by 0.068 and 0.032 in the same order

4.2. Discussion

The results demonstrate the effectiveness of YOLOv7 in detecting fire and smoke instances in addition to low-light or night time environments. Given the results, it wasn't as promising as the researchers had hoped, as the augmented models achieved a lower precision, recall, F1-score, and mean average precision in the detection tasks when compared to the original dataset.

There are several possible reasons for the poor performance of the models. First, the epoch and batch size are unideal and were decided due to time constraints, as an increase in batch size leads to faster convergence, while larger epochs lead to more accuracy as the model allows for more opportunities to learn the intricacies of data or overfitting if increased too much. Additionally, the model downscaled the images to 640x640 as further adjustments weren't suitable to the local training environment.

Second, the dataset used has contributed to the issue due to the imbalance of classes as it was diversified but not balanced, leaning more towards daytime images and providing a smaller emphasis on the contribution of LLI enhancement to the night time images. Moreover, the resolution of the images varied, as the researchers failed to apply pre-processing techniques to account for different sizes, as the model downscaled the images.

Despite the unsatisfactory results, there are a couple of notable observations in Table 2. 1) The researchers used the CLAHE version used for grayscale or black and white images, rather than RGB, yet it outperformed the deep learning model MIRnet for LLI on average across all metrics, but particularly in smoke detection. 2) Despite being the worst model out of the three, the MIRnet-enhanced model was able to perform quite reasonably for fire detection tasks, outperforming the CLAHE across all metrics except precision.

5. Conclusion and Recommendations

This study proposed and utilized YOLOv7 for the identification of smoke and fire using a synthetic low-light enhanced dataset to test its efficacy. The model has offered insight into the challenges of limited real-world data, low-light or challenging lighting conditions, and visual-based fire and smoke detection systems.

The findings show that YOLOv7 with the LLI-enhanced models was unable to perform comparably to the baseline model. The results show that the traditional CLAHE enhanced model outperformed the deep learning-enhanced MIRnet model, albeit minute. Across all metrics for both detection tasks, the CLAHE enhanced model underperformed with an average of 24.09% while 29.19% for MIRnet against the baseline model.

Moving forward, there are several ways to improve and further the research. Firstly, an option to further enhance the model's performance is the exploration of more sophisticated low-light image enhancement techniques and hyperparameter tuning. The utilization of more diverse, balanced, and realistic synthetic datasets may also help the model improve. Furthermore, the integration of data fusion techniques may improve the robustness and accuracy of fire and smoke detecting systems. Deploying and testing the model in real-world settings may also contribute to more accurate evaluations of its effectiveness and practical applications.

In conclusion, the study contributes to the ongoing efforts to develop more efficient non-urban fire and smoke detection systems by leveraging advancements in computer vision technologies and synthetic datasets. The approach holds great potential to enhance public safety and disaster management strategies.

6. Acknowledgements

This work is an output in compliance with the final requirement of CS174 under the Computer Science program of Mapua Institute of Technology. Acknowledgment and gratitude are given to our course adviser, Sir John Paul Tomas, for his support and guidance.

7. References

[1] Bu, F., & Gharajeh, M. S. (2019). Intelligent and vision-based fire detection systems: A survey. Image and Vision Computing, 91, 103803. https://doi.org/10.1016/j.imavis.2019.08.007

[2] Çelik, T., Özkaramanlı, H., & Demirel, H. (2007). Fire and smoke detection without sensors: Image processing based approach. 15th European Signal Processing Conference (EUSIPCO 2007), 1794–1798. https://doi.org/10.5281/zenodo.40571

[3] Chen, T., Wu, P., & Chiou, Y. (2005). An early fire-detection method based on image processing. 2004 International Conference on Image Processing (ICIP). https://doi.org/10.1109/icip.2004.1421401

[4] Dewangan, A., Pande, Y., Braun, H., Vernon, F. L., Perez, I., Altıntaş, I., Cottrell, G. W., & Nguyen, M. D. (2022). FIgLib & SmokeyNet: Dataset and Deep Learning Model for Real-Time Wildland Fire Smoke Detection. Remote Sensing, 14(4), 1007. https://doi.org/10.3390/rs14041007

[5] Man, K., & Chahl, J. (2022). A review of synthetic image data and its use in computer vision. Journal of Imaging, 8(11), 310. https://doi.org/10.3390/jimaging8110310

[6] Manettas, C., Nikolakis, N., & Alexopoulos, K. (2021). Synthetic datasets for Deep Learning in computer-vision assisted tasks in manufacturing. Procedia CIRP, 103, 237–242. https://doi.org/10.1016/j.procir.2021.10.038

[7] Mohammed, R. K. (2022). International journal of nonlinear analysis and applications. https://doi.org/10.22075/ijnaa.2022.5899

[8] Nikolenko, S. I. (2019). Synthetic data for deep learning. arXiv (Cornell University). https://doi.org/10.48550/arxiv.1909.11512

[9] Saponara, S., Elhanashi, A., & Gagliardi, A. (2020). Real-time video fire/smoke detection based on CNN in antifire surveillance systems. Journal of Real-time Image Processing, 18(3), 889–900. https://doi.org/10.1007/s11554-020-01044-0

[10] Statista. (2023, April 20). Number of deaths due to fire and smoke in U.S. 1920-2021. https://www.statista.com/statistics/526310/timeline-deaths-number-injury-related-due-to-fire-flames-smoke-in-us/

[11] Umar, M. M., De Silva, L. C., Bakar, M. S. A., & Petra, M. I. (2017). State of the art of smoke and fire detection using image processing. International Journal of Signal and Imaging Systems Engineering, 10(1/2), 22. https://doi.org/10.1504/ijsise.2017.084566

[12] Zhang, Q., Lin, G., Zhang, Y., Xu, G., & Wang, J. (2018). Wildland Forest Fire Smoke Detection Based on Faster R-CNN using Synthetic Smoke Images. Procedia Engineering, 211, 441–446. https://doi.org/10.1016/j.proeng.2017.12.034

[13] Ağirman, A. (2022). Nighttime Fire Detection from Video. http://acikerisim.agu.edu.tr/xmlui/bitstream/handle/20.500.12573/1442/731763%20%281%29.pdf?sequence=1&isAllowed=yS

[14] Yunnan University, (2023). synthetic fire-smoke Dataset. Roboflow Universe. https://universe.roboflow.com/yunnan-university/synthetic-fire-smoke

[15] Park, M., & Ko, B. C. (2020). Two-Step Real-Time Night-Time Fire Detection in an Urban Environment Using Static ELASTIC-YOLOv3 and Temporal Fire-Tube. Sensors, 20(8), 2202. https://doi.org/10.3390/s20082202

[16] Chen, R., Dewi, C., Zhuang, Y., & Chen, J. (2023, August 28). Contrast limited adaptive histogram equalization for recognizing road marking at night based on Yolo models. IEEE Xplore. https://ieeexplore.ieee.org/abstract/document/10232983

[17] Cui, H., Li, J., Hua, Z., & Fan, L. (n.d.). Attention-Guided Multi-Scale feature fusion network for Low-Light image enhancement. Frontiers in Neurorobotics, 16. https://doi.org/10.3389/fnbot.2022.837208

[18] Wang, C., Bochkovskiy, A., & Liao, H. M. (2023). YOLOv7: Trainable bag-of-Freebies sets new state-of-the-Art for real-time object detectors. 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). https://doi.org/10.1109/cvpr52729.2023.00721

[19] Zhang, C., Hu, Z., Xu, L., & Zhao, Y. (2023). A YOLOv7 incorporating the Adan optimizer based corn pests identification method. Frontiers in Plant Science, 14. https://doi.org/10.3389/fpls.2023.1174556

[20] Chen, J., Bai, S., Wan, G., & Li, Y. (2023). Research on YOLOv7-based defect detection method for automotive running lights. Systems Science & Control Engineering, 11(1). https://doi.org/10.1080/21642583.2023.2185916

[21] Xia, T., Zhang, R., Zhang, Y., & Chen, B. (2022). Application of YOLOv7 and transformer structures to small object (license plate) detection in complex traffic scenes. 2022 4th International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI). https://doi.org/10.1109/mlbdbi58171.2022.00031

[22] Sun, H., Ye, H., & Zhai, J. (2023). Nighttime vehicle object detection based on improved YOLOv7. Lecture Notes in Electrical Engineering, 667-675. https://doi.org/10.1007/978-981-99-6847-3_58

[23] Qiu, Y., Lu, Y., Wang, Y., & Jiang, H. (2023). IDOD-YOLOV7: Image-dehazing YOLOV7 for object detection in low-light foggy traffic environments. Sensors, 23(3), 1347. https://doi.org/10.3390/s23031347

[24] Al-refai, G., Elmoaqet, H., Ryalat, M., & Al-refai, M. (2023). Object detection in low-light environment using YOLOv7. https://doi.org/10.21203/rs.3.rs-3365905/v1

[25] Wu, Y., Tang, Y., & Yang, T. (2023). An improved nighttime people and vehicle detection algorithm based on YOLO v7. 2023 3rd International Conference on Neural Networks, Information and Communication Engineering (NNICE). https://doi.org/10.1109/nnice58320.2023.10105722