

Performance Evaluation of Fire and Smoke Detection with Object based DNN Algorithm

Jisang Moon¹, Seongjin Lee²

¹Dept. of Aerospace and Software Engineering Gyeongsang National University, Jinju, Korea

²Dept. of AI Convergence Engineering Gyeongsang National University, Jinju, Korea

Abstract

The consequence of forest fire is so great that we need to prevent it from happening with all of our resources. However, the time and human resources required to cover all the forest is very high. In this paper, we use object based Deep Learning detection algorithm, YOLOV5, to address the issue of detecting forest fire and smoke. We collected 13,924 images of forest fire and smoke and manually labeled them. We used YOLOV5n to learn the features. We solved overfitting problem via mosaic data augmentation, non-transfer learning, and hyperparameter evolution. YOLOV5n shows that mAP 0.5 is about 14.1% higher than the official YOLOV5n model.

Keywords

YOLOV5, Fire and Smoke Detection, Data Augmentation, Non-Transfer Learning, Hyperparameter Tuning

1. Introduction

The scale of damage done by a forest fire depends greatly on the time between identification of the fire and initial response to mitigate the fire. The faster the report is propagated to the authorities, the higher the chances to contain the fire from spreading the whole forest. Although the main cause of the fire is from the mountain trackers, many of the cases are from the unidentified sources. The surveillance area is so wide that the authorities cannot afford the cost to monitor the area. The research community have come up with methods to detect fire with CCTVs, infrared cameras, IoT fire detection sensors. Behind such methods is image processing techniques. These techniques show alarm delays and high false alarm rates. In this paper, we exploit object based Deep Learning Detection method, YOLOV5n, to address the issue of detecting forest fire and smoke. We collected 13,924 images of fire and smoke and manually labeled the images. Then, we used YOLOV5n to learn the features. To solve the overfitting problem, we adopted Mosaic augmentation [1], non-transfer learning [2], and hyperparameter evolution [3]. Overfitting was solved, and mAP 0.5 is about 14.1% higher than the official YOLOV5n model.

2. RELATED WORK

Lee et al [4] said the occurrence of forest fires in Korea is higher in seasons other than summer, with an increase in frequency in winter in the north due to differences in plant growth periods. Human misfire is the primary cause, but the diversity of causes is increasing, particularly arson. The frequency of forest fires is increasing, and the damage is greater due to the accumulation of forest resources. Large forest fires are increasing in the West Coast area, with no safe zone in the country.

Seo et al [5] proposed fire detection algorithm exploiting image information. They changed the RGB to HSV channel to distinguish and identify flames. Since their method is color dependent, it shows vulnerability in objects with red color. They avoided detecting wrong object by setting the threshold high to ignore the small contours. The obvious issue with the approach is that it cannot detect small fires or the fire in early stage.

Lim et al [6] makes use of CCTV to detect fire. They used CCTV to capture the image and CNN to detect the fire. They used horizontal and vertical flipping, and cropping to prevent the overfitting. They show 90% of test accuracy but with unstable false alarm rate.

Ahmad A. A. Alkhatib [7] discusses the use of wireless sensor networks for forest fire detection. This technology provides real-time monitoring and accurate information with less delay, making it ideal for the application. Key issues in the network for this application include localization, coverage, network lifespan, and fire detection method. The previous work

ISE2022: 1st International Workshop on Intelligent Software Engineering, December 6-9, 2022, Gyeongsang National University, Korea

0000-0003-4239-4661 (J. Moon); 0000-0003-0760-1880 (S. Lee)



© 2023 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

used other sensors and integrated information with databases and models to reduce false alarms. Each node was equipped with multiple sensors to detect fire incidents. Overall, sensor networks are the best solution for accurate and reliable forest fire detection.

Lee et al [8] discusses the development of a wildfire detection system utilizing unmanned aerial vehicles (UAVs). The system uses deep convolutional neural networks to detect wildfires in aerial still photographs. The evaluation of the system shows that GoogLeNet and the modified VGG13 achieved high accuracies. The system can be used for early detection of wildfires and is capable of detecting fires on both ground and onboard UAV systems using a mobile GPU.

3. METRIC

3.1. LOSS

Loss is an evaluation metric used in deep learning models to measure the difference between the predicted output and the actual output of the model. YOLOv5 has three different loss functions that are used during training to optimize the model parameters and improve its performance. These losses are:

1. Objectness loss: This loss function measures the difference between the predicted and ground truth objectness scores. The objectness score is a measure of how confident the model is that there is an object in a particular location in the image. The objectness loss penalizes the model for incorrect objectness predictions and encourages it to correctly identify objects in the image.

2. Classification loss: This loss function measures the difference between the predicted and ground truth class probabilities. YOLOv5 can detect multiple objects in a single image, and the classification loss is used to classify each of the detected objects. The classification loss penalizes the model for incorrect class predictions and encourages it to correctly classify objects.

3. Localization loss: This loss function measures the difference between the predicted and ground truth bounding box coordinates. The bounding box is a rectangle that surrounds an object in the image, and the localization loss penalizes the model for incorrect bounding box predictions and encourages it to accurately locate objects in the image.

By optimizing these three losses during training, YOLOv5 can accurately detect and classify objects in

images with high precision and recall.

3.2. PRECISION

Precision is a performance metric commonly used in machine learning and information retrieval to evaluate the accuracy of a model's positive predictions. It is defined as the ratio of true positive predictions to the total number of positive predictions made by the model. The precision can be expressed mathematically as follows:

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP})$$

3.3. RECALL

Recall is a metric that measures the ability of a machine learning model to correctly identify all relevant instances of a class. It is calculated as the ratio of true positives to the sum of true positives and false negatives, as shown below:

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN})$$

3.4. PR-curve

The PR (Precision-Recall) curve is a graphical representation of the precision and recall of a deep learning model for different decision thresholds. Precision is the proportion of true positive predictions among all positive predictions made by the model, and recall is the proportion of true positive predictions among all actual positive cases. The PR curve is created by plotting precision on the y-axis and recall on the x-axis for different decision thresholds, and a high AUC-PR indicates good model performance in identifying positive cases while minimizing false positives. The PR curve is often used to evaluate the performance of binary classification models.

3.5. MAP

MAP (Mean Average Precision) is an evaluation metric for object detection models that combines precision and recall to measure the overall performance of a model at detecting objects of different classes in an image. The AP (Average Precision) for a particular class is calculated by plotting the precision-recall curve for that class and calculating the area under the curve, which is then averaged across all classes to get the MAP. A high MAP score indicates good performance in detecting objects of different classes in the image, while a low MAP score indicates poor performance. MAP is a useful evaluation metric because it takes into account the precision and recall of each class separately and combines them into a single score.

4. PROPOSED SCHEME

To detect forest fire and smoke, we used object-based Deep Learning algorithm. We collected total of 13,924 images. We manually labeled the fire and smoke in the dataset; there were 17,500 fire labels and 5,000 smoke labels. We additionally acquired 1,000 fire-like/irrelevant images to reduce false positive rate. The acquired images include red tail lights of automobiles, sunset, other red lights, etc. We used about 83.7% of the data as train data and the rest as the verification data. We used YOLOV5n as the model to detect the forest fire and smoke. To reduce the overfitting of the model, we used Mosaic Augmentation [1], Non-Transfer Learning [2], and Hyperparameter Evolution [3].

4.1. Mosaic Augmentation

Although we have managed to collect over 22,000 labels to learn, it is still not enough to learn all the various scenarios. There is also an issue of difficulties in detecting small objects in YOLOV5. To provide remedy to the issue, we used Mosaic augmentation [1]. Mosaic Augmentation is one of the image augmentation techniques, which makes four images into one image. It also has advantage of working well with small batch size. The size of each image is chosen at random and the images are cropped while making a mosaic to fit into the dimension.

4.2. Transfer Learning vs Non Transfer Learning

Transfer Learning [2] allows you to quickly reach high accuracy based on pretrained features and applying those weights to the given dataset. Although YOLOV5 provides pretrained model based on COCO dataset, the dataset does not include forest fire and smokes. When there are similarities in the pretrained dataset and the target dataset, the transfer learning approach could result in high accuracy but when there is little or no similarities then, it results in relearning of the dataset from the beginning losing all of its merits. Figure. 1(a) and Figure. 1(b) show the comparison between adopting the transfer learning and without the transfer learning. Although the *val/box_loss* and *val/obj_loss* is slightly increased, the overall result shows that adopting non-transfer learning is better than adopting the transfer learning.

4.3. Hyperparameter Evolution

Hyperparameter Evolution [3] is a method of optimizing Hyperparameters using Genetic Algorithm. In this work, mutations that generate new offspring are used based on the best combination of parents of all previous generations. Generation is conducted 300 times, 10 epochs at a time. Figure. 1(c) shows the result of 200 epochs

of learning with non-transfer learning and adopting hyperparameter evolution. It shows that the results are stabilized compared to the Figure. 1(a) and 1(b). It also shows that overfitting is reduced.

4.4. Verification

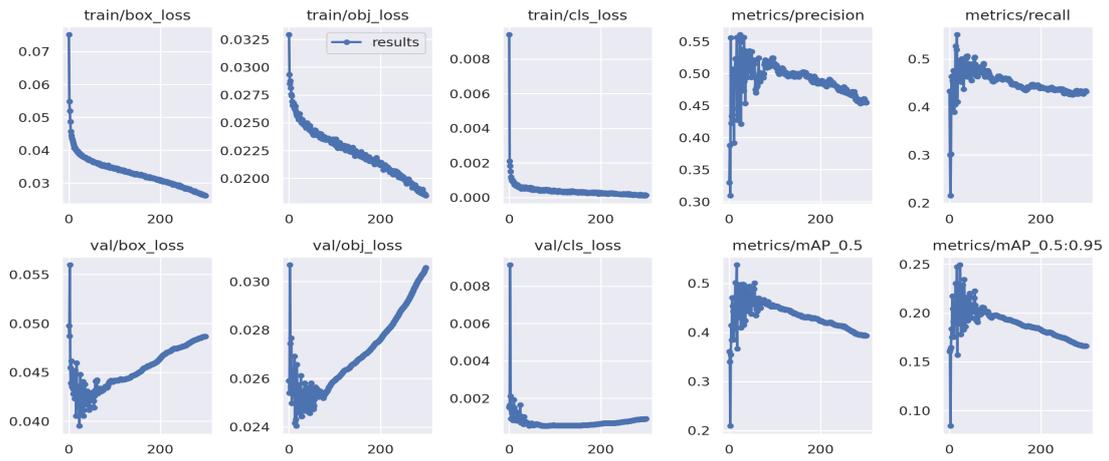
Table I shows the evaluation index for fire and smoke data. The area below the precision-recall curve becomes AP (Average Precision), and mAP is the mean of AP values according to the value of IOU (Intersection Over Union). The higher the mAP, the better the performance model. According to the performance indicators, fire objects have a high value of mAP (0.674), but low for smoke (0.389). The reason behind the low mAP for the smoke data is mainly due to imbalance of data. There were only 5,000 labels of smoke whereas there were 17,500 labels of fire in the dataset. Table II shows the Evaluation Index for official YOLOV5n model using transfer learning the features from COCO dataset, and YOLOV5n used in this paper. Measurement of mAP averaged over IOU thresholds in [0.5:0.05:0.95] shows 28.0 and 25.9 for official YOLOV5n and the proposed, respectively. The proposed is about 7.5% lower. In the case of mAP 0.5 for the official YOLOV5n and the proposed shows 45.7 and 53.2, respectively. The proposed is about 14.1% higher than the official.

Table 1
THE EVALUATION INDEX FOR FIRE AND SMOKE

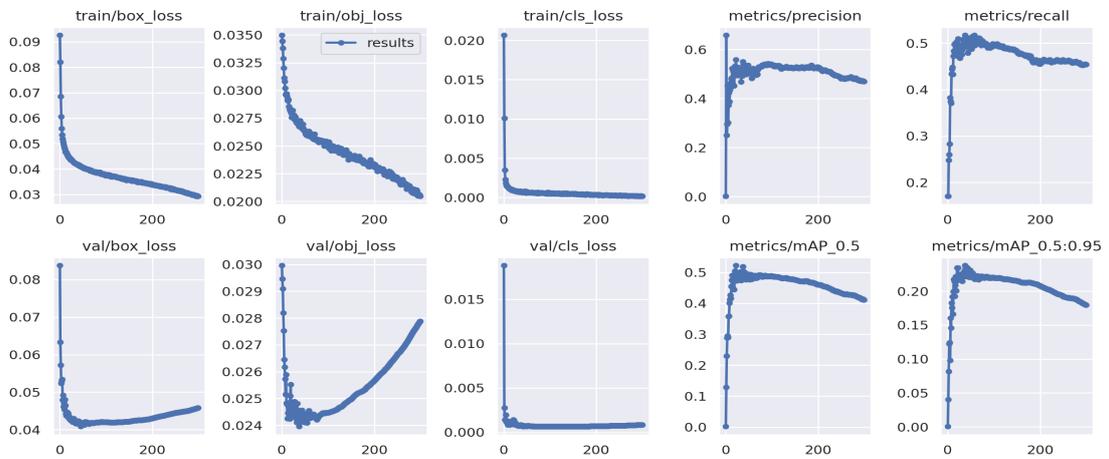
Class	Evolution Index		
	Image	Instances	Precision
All	2,268	5,932	0.570
Fire	2,268	4,281	0.595
Smoke	2,268	1,651	0.545
	Recall	mAP 0.5	mAP 0.5:0.95
All	0.513	0.532	0.259
Fire	0.670	0.674	0.377
Smoke	0.356	0.389	0.142

Table 2
PERFORMANCE COMPARISON OF TWO MODELS

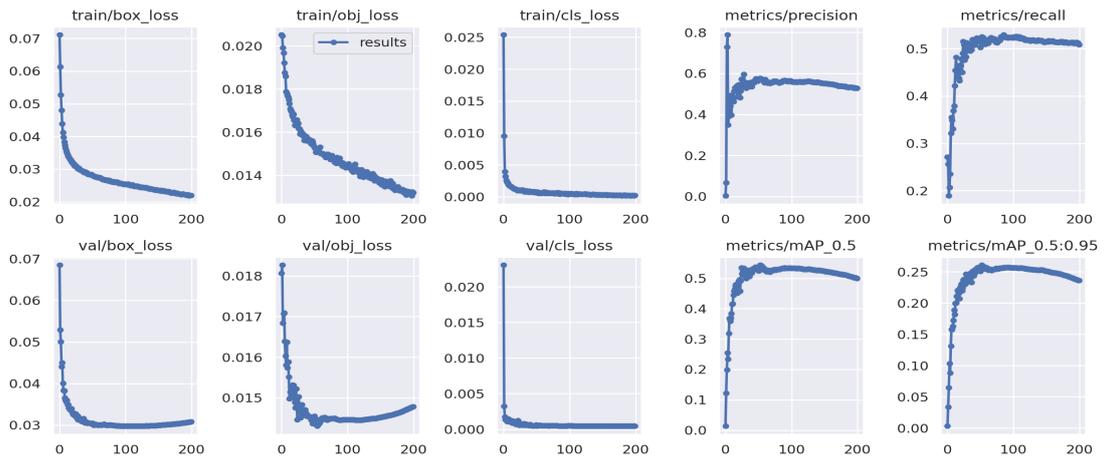
Model	Evolution Index		
	Size	mAP 0.5:0.95	mAP 0.5
YOLOV5n (coco data)	640	28.0	45.7
YOLOV5n (Fire&Smoke)	640	25.9	53.2
	Speed GPU b1(ms)		Speed GPU b32(ms)
YOLOV5n (coco data)	6.3		0.6
YOLOV5n (Fire&Smoke)	6.5		1.2



(a)



(b)



(c)

Figure 1: Result of Transfer Learning, Non-Transfer Learning, and Hyperparameter Evolution

5. CONCLUSION

In this paper, a deep learning model, YOLOV5n was used for fire and smoke detection. YOLOV5n typically have overfitting issue, and we applied three techniques to address the issue. We exploit Data Augmentation [1], Non-Transfer Learning [2], and Hyperparameter Evolution [3]. We compared the performance of the official YOLOV5n model that uses coco dataset and our model. The result shows that mAP averaged over IOU thresholds in [0.5:0.05:0.95] in the proposed model is about 7.5% lower than the official model, but it is about 14.1% higher for mAP 0.5. We find that the mAP of the smoke class is low. The reason behind the low score is that the number of labels for the smoke is about 1/3 of the fire labels. For the future work, we need to acquire smoke datasets and work on improving the mAP score for smoke objects using various data augment methods.

6. ACKNOWLEDGEMENT

This study was carried out with the support of 'R&D Program for Forest Science Technology (Project No. 2021344A00-2223-CD01)' provided by Korea Forest Service(Korea Forestry Promotion Institute).

References

- [1] A. Bochkovskiy, C.-Y. Wang, H.-Y. M. Liao, Yolov4: Optimal speed and accuracy of object detection, arXiv preprint arXiv:2004.10934 (2020).
- [2] H. Tercan, A. Guajardo, J. Heinisch, T. Thiele, C. Hopmann, T. Meisen, Transfer-learning: Bridging the gap between real and simulation data for machine learning in injection molding, *Procedia Cirp* 72 (2018) 185–190.
- [3] YOLOv5, Hyperparameter evolution, 2020. URL: <https://github.com/ultralytics/yolov5/issues/607>.
- [4] S.-Y. Lee, H.-P. Lee, Analysis of forest fire occurrence in korea, *Fire Science and Engineering* 20 (2006) 54–63.
- [5] M.-S. Seo, C. H. Lee, A forest fire detection algorithm using image information, *Journal of the Institute of Convergence Signal Processing* 20 (2019) 159–164.
- [6] J. Yim, H. Park, W. Lee, S. Kim, Y.-T. Lee, Deep learning based cctv fire detection system, in: *Proceedings of the Korean Society of Broadcast Engineers Conference, The Korean Institute of Broadcast and Media Engineers, 2017*, pp. 139–141.
- [7] A. A. Alkhatib, A review on forest fire detection techniques, *International Journal of Distributed Sensor Networks* 10 (2014) 597368.
- [8] W. Lee, S. Kim, Y.-T. Lee, H.-W. Lee, M. Choi, Deep neural networks for wild fire detection with un-

manned aerial vehicle, in: *2017 IEEE international conference on consumer electronics (ICCE), IEEE, 2017*, pp. 252–253.