International Journal of Aviation Science and Engineering
e-ISSN: 2715-6958 p-ISSN: 2721-5342
AVIA

Original Article

Aircraft Detection in Low Visibility Condition Using Artificial Intelligence

Khairul Ummah [1,*], M. Dhiku Widyosekti [1], Yanuar Zulardiansyah Arif [2], Rizal Adi Saputra [3], Akhmad Riszal [3], and Javensius Sembiring [1]

[1] Faculty of Mechanical and Aerospace Engineering, Institut Teknologi Bandung. Labtek II, Jl. Ganesa 10, Bandung, 40132 Indonesia
[2] Department of Electrical & Electronic Engineering, Faculty of Engineering, Universiti Malaysia Sarawak, Jln. Datuk Mohammad Musa, 94300 Kota Samarahan, Sarawak, Malaysia
[3] Department of Mechanical Engineering, Faculty Engineering, Universitas Lampung, Jl. Prof. Soemantri Brojonegoro no. 1, Bandar Lampung 35143 Indonesia
* Correspondence: khairulu@ae.itb.ac.id

Received: 28 January 2023; Accepted: 15 April 2023; Published: 12 June 2023

Bad weather often interferes with the functioning of the air transport system. One example is the frequent flight delays for commercial aircraft, resulting in losses for both the airline and passengers. Artificial Intelligence (AI) technology can now minimize delays caused by bad weather, especially in low visibility conditions. This paper discusses AI modeling that can detect aircraft in a low visibility weather condition, especially in the airport area. The employed method is the deep learning approach with the YOLOv4 algorithm (single-stage detection), which is regarded as one of the optimal platforms in this field. There are 600 images used in this work to create and train three different models. Image Dehazing filter is employed on the training data before it is trained to produce the detection model. The result shows that the model has a good performance in terms of performance metrices. Thus, this model is suitable to be used to detect aircraft in low visibility conditions.

Keywords: low visibility, artificial intelligence, aircraft, air traffic control, performance matrices.

1. Introduction

In 2019, nearly 70% of flight delays in the world were caused by bad weather including fog, thunderstorms, snowstorms, wind shear, and icing. Delays on flights result in cost for passengers and airlines [1]. The study conducted previously shown that in 2007, US suffered losses due to flight delays of 31.2 billion dollars, where the components of the loss consisted of Airline, Passengers, Lost Demand, and Indirect Cost [2]. To minimize losses, it is necessary first to know the cause of the delay. Based on Zamkova (2017) there are three main factors causing delay: delay caused by air traffic control, delay caused by airport limitation, and delay caused by delay on the previous flight [3]. Meanwhile, in Indonesia, stated that flight delay factor in Indonesia is dominated by airline factors and airport operations such as Air Traffic Control (ATC) [4].

Airlines and airports have made efforts in dealing with delays, and there is one airport that can reduce delays with the help of artificial intelligence (AI), namely Heathrow, London. This airport could reclaim up to 20% of lost capacity due to delay caused by low cloud. AI is a supporting tool to assist the ATC in making decisions and enhancing the controller's performance [5].

Practically, no other airport has implemented this artificial intelligence technology. A similar technology applied in other airports is Airport Surveillance Radar (ASR) which functions as a monitor for aircraft movements both on ground and airside. However, in Indonesia, no airports are

e-ISSN: 2715-6958 p-ISSN: 2721-5342

equipped with this technology. This paper attempted to further investigate the mechanism of this system and its applicability in the airport environment. AI that helps ATC in the future can improve the Controller's performance in monitoring aircraft movements visually and in real-time, especially at large airports, often constrained by low visibility.

2. Materials and Methods

2.1. Data Collection

The author takes secondary data from open dataset sources on the Google APIs and other open sources on the Internet. The dataset used in this study includes jet commercial aircraft images of various angles and backgrounds in a typical and foggy condition [6]. The amount of data used in this research is 600 data divided into two datasets, namely the training dataset and the validation dataset. According to research that has been done previously, the training dataset has a percentage of 80% training data and 20% validation data. The following figure shows some of the image data in the dataset used for developing artificial intelligence models in this study.

2.2. Data Preprocessing

The data obtained is still in raw data, which must be processed first to fit the model. In this study, each of the raw data will go through two stages of processing: labeling and filtering the data.

2.2.1. Labelling (annotate)

Labelling provides information on the object to be detected. The data label is in the form of a bounding box around the object, which is then referred to as ground truth. The output of this step is a file with (.txt) format that fits to YOLO format. These parameters describe the class of the object and the coordinates of the bounding box [7]. Secondary data from Googleapis is already labeled. In contrast, data obtained from other open sources do not yet have a label, so a web-based application called Roboflow [8] is used to carry out the labelling process.

2.2.2. Filtering Data

After labeling the data, filtering the data is carried out to add new image variations. The filtering process carried out in this study is by using Image Dehazing and Edge Enhancing sourced from Open CV [9] which is depicted in Figure 1.

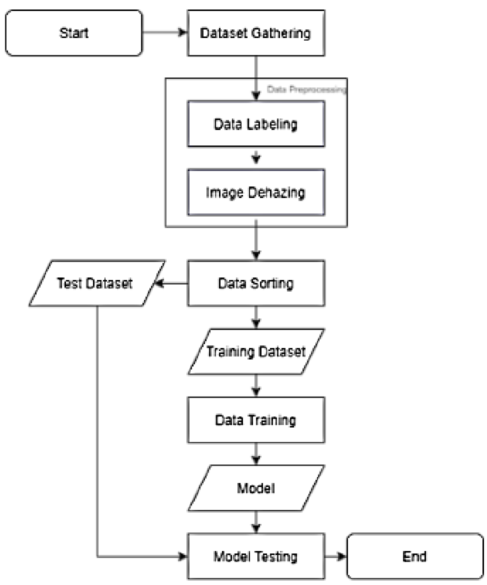


Figure 1. Model without Image Dehazing Filter

There are three models that will build with different filters as seen on figure above:

1. First model is a model without filter which will be the baseline.
2. Second model is a model using Edge Enhancing filter.
3. Third model is a model using Image Dehazing filter.

2.3. Train the Model

The transfer learning method is used to train the model used in this study. Transfer learning is the process of applying previously learned knowledge and skills tasks to the new target task, which mean the model used has been trained [10]. Pretrained models are usually trained in large data sets, by retraining the model, the accuracy will be increasing even with a limited number of datasets [11]. Transfer learning was carried out using the YOLOv4.conv.137 pre-trained model provided Alexey's GitHub Page [12].

The Google Collab platform is used to train the model in this research. It provides a virtual machine with a Graphics processing unit (GPU) which can be used for free and limited by Google users.

Here are the parameters needed in model training:

Table 1. Parameter setup

Parameter	Value
Batch size	64
Subdivision	16
Width	416
Height	416
Channel	3
Class	1

2.4. Model Evaluation

The evaluation process is carried out every 100 iterations and will automatically be saved. The results of the evaluation itself consist of several metrics, such as true positive (TP), false positive (FP), false negative (FN), precision, recall, F1-score, and mean average precision (mAP). The metric that will used to compare between model is mAP.

Model testing is done with a local machine in a laptop equipped with a GPU and assisted by software such as Anaconda, Git Bash, and Visual Studio Code. Anaconda is used to create a virtual environment to install programming packages needed, such as OpenCV, TensorFlow, CUDA Toolkit, cudnn, and Python. Visual Studio Code is used to modify the code as needed, and then Git Bash to run the commands. The settings for the local machine and the main source code for running the model are obtained Adrian Rosebrock, (2017) [13] and The AI Guy's GitHub page [12].

3. Results and Discussions

3.1. Experiment Results

The three models will be used in this analysis to detect aircraft in low visibility conditions. The performance of the three models will be compared from the evaluation metrics and visual detection results. Visual testing experiments were carried out using images in aircraft in low visibility conditions.

In the first experiment, the 600 datasets that had gone through the labeling process were divided into training datasets and 80% and 20% validation datasets. The dataset is directly used in the training process without going through any image filtering process. Table 2 shows that the first model produced 148 predictions with a precision of 0.76 and a recall/sensitivity of 0.83, resulting in F1 and mAP scores of 0.80 and 86.10%, respectively. Figure 1 shows that the training process is going through 3000 iterations.

Table 2. Metric Evaluation on First Experiment

Metric	Value
TP	113
FP	35
FN	23
Precision	0.76
Recall	0.83
F1-Score	0.80
mAP	86.10%

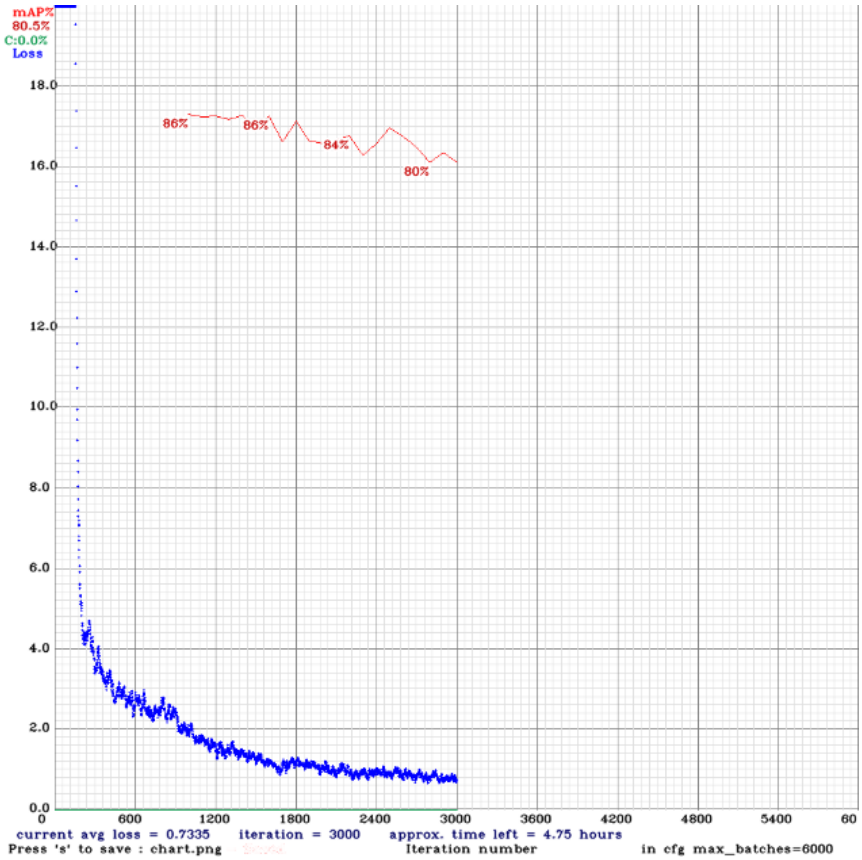


Figure 2. mAP Graph of First Experiment

In the second experiment, with the same number dataset with first experiment, after the labeling process the dataset is going through Edge Enhancing as part of the filtering process before the training is done. Table 3 shows that the second model produced 142 predictions with a precision of 0.82 and a recall/sensitivity of 0.86, resulting in F1 and mAP scores of 0.84 and 86.65%, respectively.

Table 3. Metric Evaluation on Second Experiment

Metric	Value
TP	117
FP	25
FN	19
Precision	0.82
Recall	0.86
F1-Score	0.84
mAP	86.65%

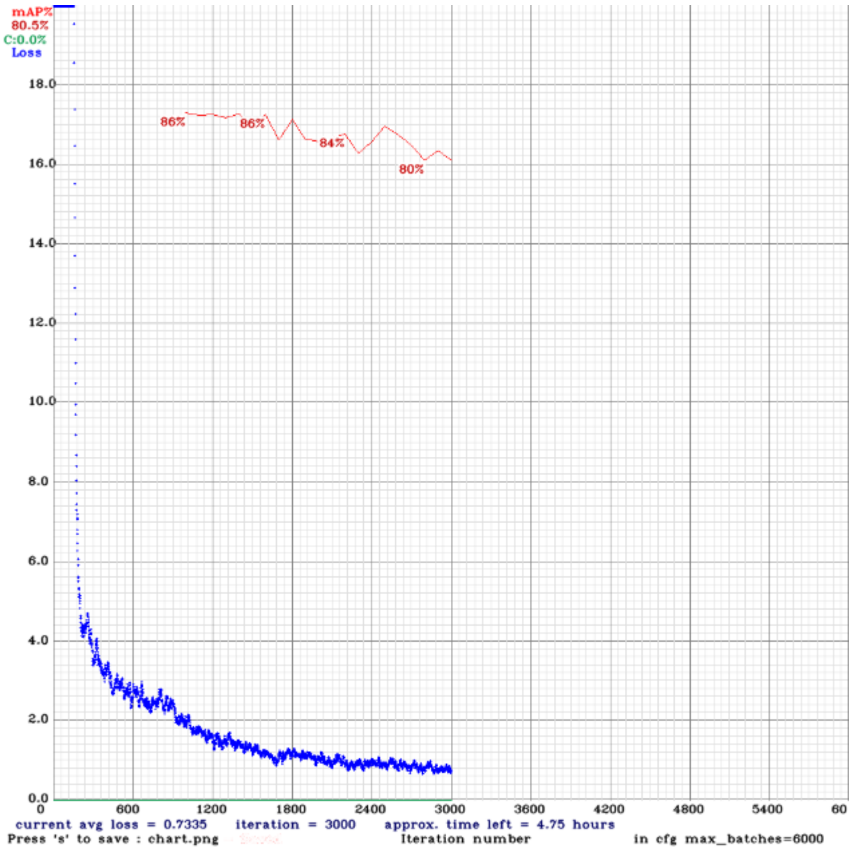


Figure 3. mAP Graph of Second Experiment

In the third experiment, it has the same process as second experiment, except for the filtering process, this dataset is using Image Dehazing filter in the filtering process. Table 4 shows that the third model produced 162 predictions with a precision of 0.72 and a recall/sensitivity of 0.87, resulting in F1 and mAP scores of 0.79 and 87.11%, respectively.

Table 4. Metric Evaluation on Third Experiment

Metric	Value
TP	118
FP	46
FN	18
Precision	0.72
Recall	0.87
F1-Score	0.79
mAP	87.11%

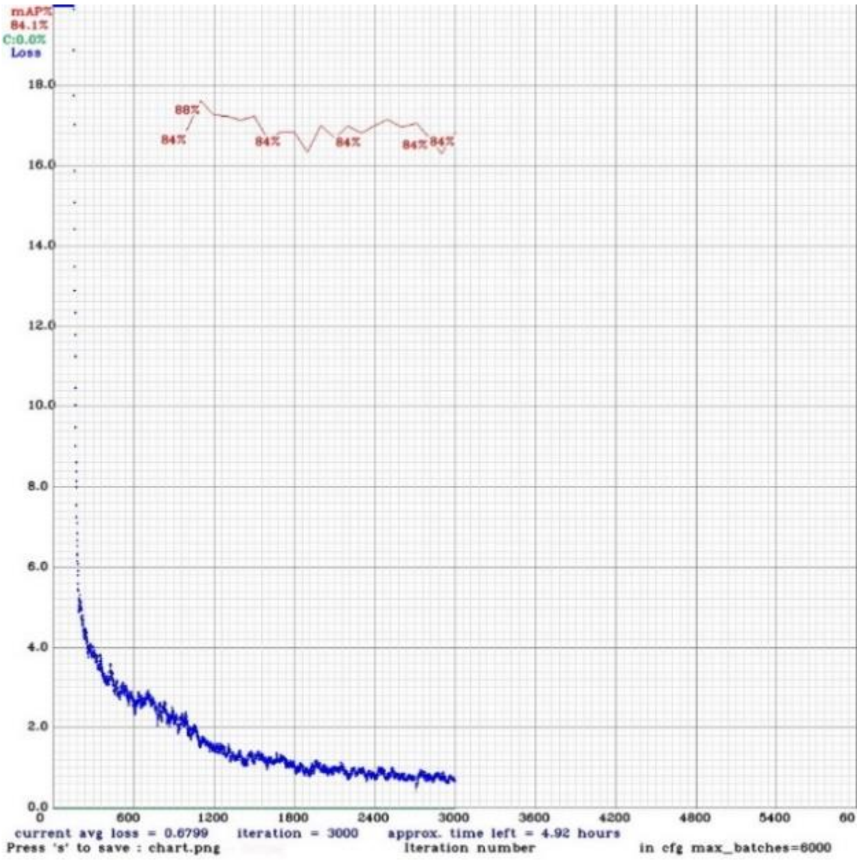


Figure 4. mAP Graph of Third Experiment.

3.2. Analysis

The results of experiments that have been carried out in this research will be compared in two ways. First, by comparing the value of the evaluation metric of each model and secondly, by visually comparing the detection results of each trial. Comparison with the detection results visually becomes an assessment key to the performance of artificial intelligence models.

3.2.1. Comparison with Evaluation Metrics

In this table, first Model is defined as the model of the first experiment, which was only going through labeling without any image filtering. Second model is defined as the model of the second experiment, which has been filtered with Edge Enhancing. Third model is defined as the model of the third experiment, which has been filtered with Image Dehazing. The following table shows the type of image filtering used in each model as part of data preprocessing.

Table 5. Comparison of Metric Evaluation

Metric	First Model	Second Model	Third Model
TP	113	117	118
FP	35	25	46
FN	23	19	18
Precision	0.76	0.82	0.72
Recall	0.83	0.86	0.87
F1-Score	0.80	0.84	0.79
mAP	86.10%	86.65%	87.11%

 doi: 10.47355/AVIA.V5I1.84

e-ISSN: 2715-6958 p-ISSN: 2721-5342

Based on the table above, by comparing the mAP metrics model without filters (first model) and model with filters (second model and third model) that addition of filter has improved the model. The performance's difference between second model and third model is the former has a better precision, and the latter has better recall in detecting objects. But in this case, the metrics recall is preferrable because in case of detecting aircraft in low visibility condition it is better to have more False Positive (FP) than higher False Negative (FN) which means the technology more sensitive in detecting objects. This is due to the technology is not the main tool to decides whether the runway is safe to use for the next airplane to lands, it's more focuses on assisting the controllers to enhance their performance in monitoring the runway. Amongst all the models, the third model has the best performance in terms of mAP. To ensure that the model can be implemented, it is necessary to test all the model visually by using data images which is not included in the training model.

3.2.2. Comparison with Visual Metrics

Comparison with visual metrics is carried out to determine the performance of the three models in detecting aircraft in low visibility conditions. Three images will be the input for the models that will detect the object in low visibility conditions. Before the model carries out the detection process, the input in the image above must go through the image filtering stage first adjusted to the Image filter used by each model.

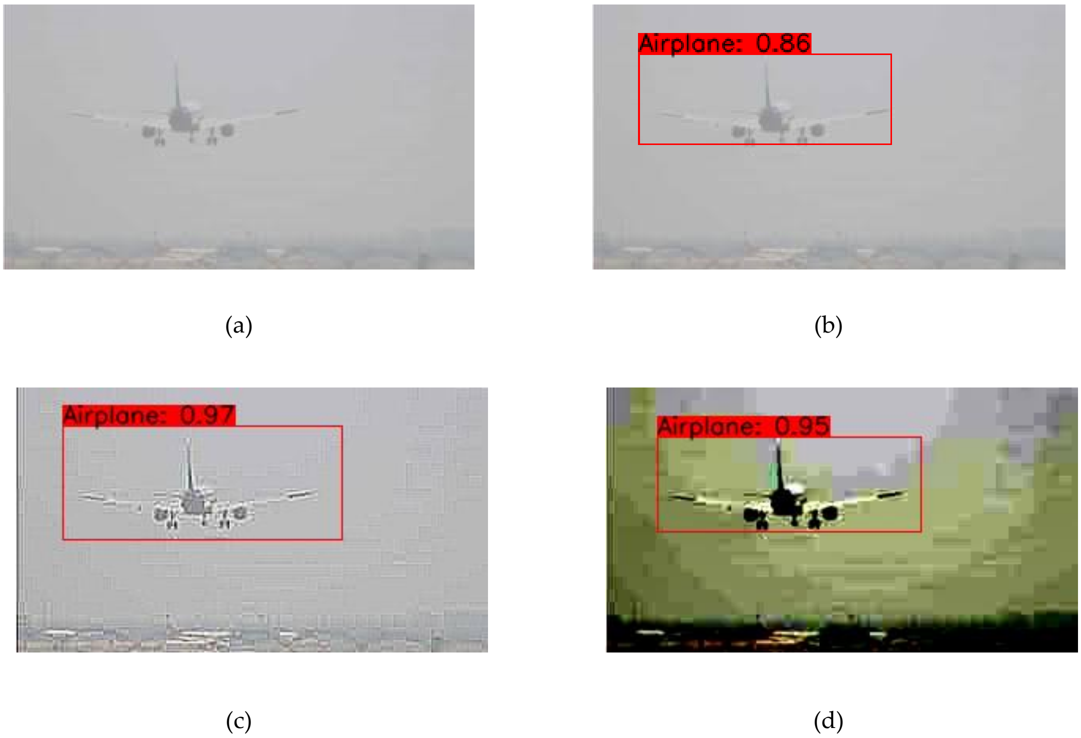
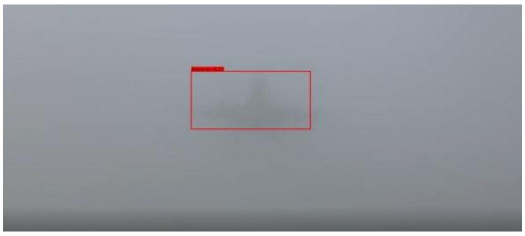


Figure 5. Visual Comparison of (a) first input, (b) first model output, (c) second model output, and (d) third model output

(a) (b)


(c) (d)

Figure 6. Visual Comparison of (a) second input, (b) first model output, (c) second model output, and (d) third model output


(a) (b)


(c) (d)

Figure 7. Visual Comparison of (a) third input, (b) first model output, (c) second model output, and (d) third model output

Figure 5 shows that all the models can detect the first input, and the addition of filters improve the confidence score of the model. Second model gives the best confidence score in detecting the first input. Figure 6 shows the second model unable to detect the object, this is due to the filter unable to enhance the edges of the images and the thick fog makes the machine difficult to extract the edges of the object. Third model gives the best confidence score in detecting the second input. Figure 7 shows that all the models can detect the third input, and third model gives the best confidence score in detecting this input.

Overall, the third Model has the best performance because it has the highest confidence score compared to the other two models in detecting the given input. On the other hand, the first Model is better compared to the second model because it can detect all the object in every given input not like second model which only detect two out of three inputs, even though second model has a better confidence score. Thus, based on the comparative analysis third model is the most suitable model to use in detecting aircraft in low visibility condition.

5. Conclusions

The model we proposed that without filters can detect aircraft in low visibility condition re are three conclusions based on this research. The addition of the *Image Dehazing* filter improves the model's performance by 1% in terms of mean Average Precision (mAP) and successfully detects all objects from given tested data. In contrast, the addition of the *Edge Enhancing* filter improve the model's performance by 0.55% in terms of mean Average Precision (mAP) but only detects two objects out of three from given tested data. The most suitable model is the model that utilizes the *Image Dehazing* filter employed on the training data.

References

1. Federal Aviation Administration, *FAA Facts.* http://www.faa.gov/news/fact_sheets/news_story.cfm?newsId=14153.
2. M. Ball *et al.,* (2010).
3. M. Zamkova, M. Prokop, and R. Stolin, Acta Univ. Agric. Silvic. Mendelianae Brun. **65**, 1799-1807 (2017)
4. I. Setiawan, S. A. Majid, and Y. Tuliantini. J. Manaj. Transp. Dan Logistik. **2**, 365-384 (2015)
5. P. Henttu, T. Grobotek, and I. Prissonotti, (2019)
6. Google LLC, *Google APIs.* https://storage.googleapis.com/openimages/web/index.html.
7. R. Padilla, W. L. Passos, T. L. B. Dias, S. L. Netto, and E. A. B. Da Silva, Electron. **10**, 279-306 (2021)
8. Roboflow. https://roboflow.com/
9. OpenCV. https://opencv.org/
10. Cs231, Gtihub. https://cs231n.github.io/transfer-learning/
11. C. H. Lee and C. W. Lin, Appl. Sci. **11** (2021)
12. A. Grigorev, Github. https://github.com/theAIGuysCode
13. D. A. Rosebrock. Learn Comput. Vis. Using OpenCV (2019)