Journal of Harbin Engineering University
ISSN: 1006-7043
Vol 44 No. 9
September 2023

Anomaly Detection in Iranian airport's Aviation using Deep learning

S. M. KamaliFiroozabadi[1], M. Poorahmadi[2*] and A. M. Latif[3]
[1] Department of Electrical Engineering, Yazd Branch, Islamic Azad University, Yazd, Iran Email : kamali239@iauyazd.ac.ir
[2] Department of Electrical Engineering, Yazd Branch, Islamic Azad University, Yazd, Iran
Email : pourahmadi@iauyazd.ac.ir
[3] Department of Electrical Engineering, Yazd University, Yazd, Iran
Email : alatif@yazd.ac.ir
*Corresponding author: M. Poorahmadi, Email: pourahmadi@iauyazd.ac.ir.

Abstract- Aviation is one of the essential transportation elements in each country, which has always been the center of attention. One of the greatest damages to this industry, is the events that happen for the airplanes, as the most important element of this industry. These events might occur due to human error, unwanted hardware failure, felonious and terroristic deeds, which will change the usual behavior of the airplane; thus, it is expected that the anomaly detection methods might be useful in finding these events. The anomaly detection methods based on human supervision are very constrained. One of these constraints is human error. Another constraint is the heavy cost imposed on the aerospace industry. Finally, the most important constraint is the analysis of human force performance in checking all flights, which causes major problems in airports with heavy traffic. The automatic anomaly detection methods resolve the constraints and obstacles based on human supervision. The proposed method, which employs deep convolutional neural networks and data of three large airports, including Mehrabad (Tehran), Hasheminezhad (Mashhad), and Shahid Sadooghi (Yazd), can significantly reduce the events resulting from human error; an accuracy of 99.13% verifies the above claim.

Keywords- Aviation, anomaly detection, automatic anomaly detection, feedforward multilayer neural networks.

Mathematics Subject Classification : 68T05
Computing Classification System : ccs (Artificial intelligence)

1. INTRODUCTION

One of the problems that most countries and governments have to deal with is caused by air traffic. For instance, heavy losses of the events caused by airplanes collision or collision with airport walls and so on, can be mentioned. A solution to this problem is to use automatic anomaly detection techniques along airplanes' route to find the abnormal cases, and control and prevent possible events (Rose and Cohen, 1994).

An anomaly detection system is based on training data collected from behavior of the target case in a specific period. The training data can be used to train the algorithm of interest and achieve a system that receives a new route as input, checks it to see if it is normal or not, and returns the route's being normal or abnormal at the output.

However, several problems might occur. First, the mentioned process requires a large number of tagged data for the training process. Second, since there are a variety of anomalies, the system should obtain an understanding of the above process beyond the test and training process using the training data to detect the newanomalies. The data of the target's route over time can be used to solve the first problem. Since this data is accessible over time, the first problem is resolved to a great extent. Neural networks can be used to solve the second problem.

Among various methods that have been proposed in this context, machine learning methods have some shortcomings. Machine learning methods only employ training data to design the methods based on machine learning, the designed system is not flexible enough to deal with the events requiring an understanding of the event. Hence, they do not meet our requirements. Thus, machine learning methods are not optimal. Various conventional methods are compared here. Therefore, deep neural networks (DNNs) can be used to solve the problem. DNNs are employed in a wide range of applications, including bioinformation technology, machine vision, speech recognition, data mining, information recovery, and pattern recognition, indicating that DNNs are powerful tools for various applications.

Among the outstanding studies in this context, Barratt et al. (2018) can be mentioned, which has employed the mixed gaussian model and conventional clustering methods and the airplane behavior in the airport to classify and detect abnormal routes. The advantage of this method is that it is compact (compressed) and detects abnormal routes fast.

In Gabriel et al. (2018), the principle component analysis has been used to analyze airplane behavior while landing and taking off. The advantage of this method is that it uses the components effective in anomaly detection, as a result of which, an efficient anomaly detection system is achieved. In Oehling and Barry (2019) encoder NNs have been used to design an efficient and optimal anomaly detection system, which operates fast in the training step and results in a high accuracy anomaly detection process.

In this study, the second section presents the proposed method. Section 3 simulates the proposed method on the employed dataset. The obtained results are evaluated in section 4. Finally, the paper is concluded section 5 and the results are given.

2. THE PROPOSED METHOD

In this section, the proposed anomaly detection method is presented and investigated.

The proposed method is comprised of training and test sections. In the training section, the system is trained using the training data; next, the data is classified using the designed model. In other words, anomaly detection is performed.

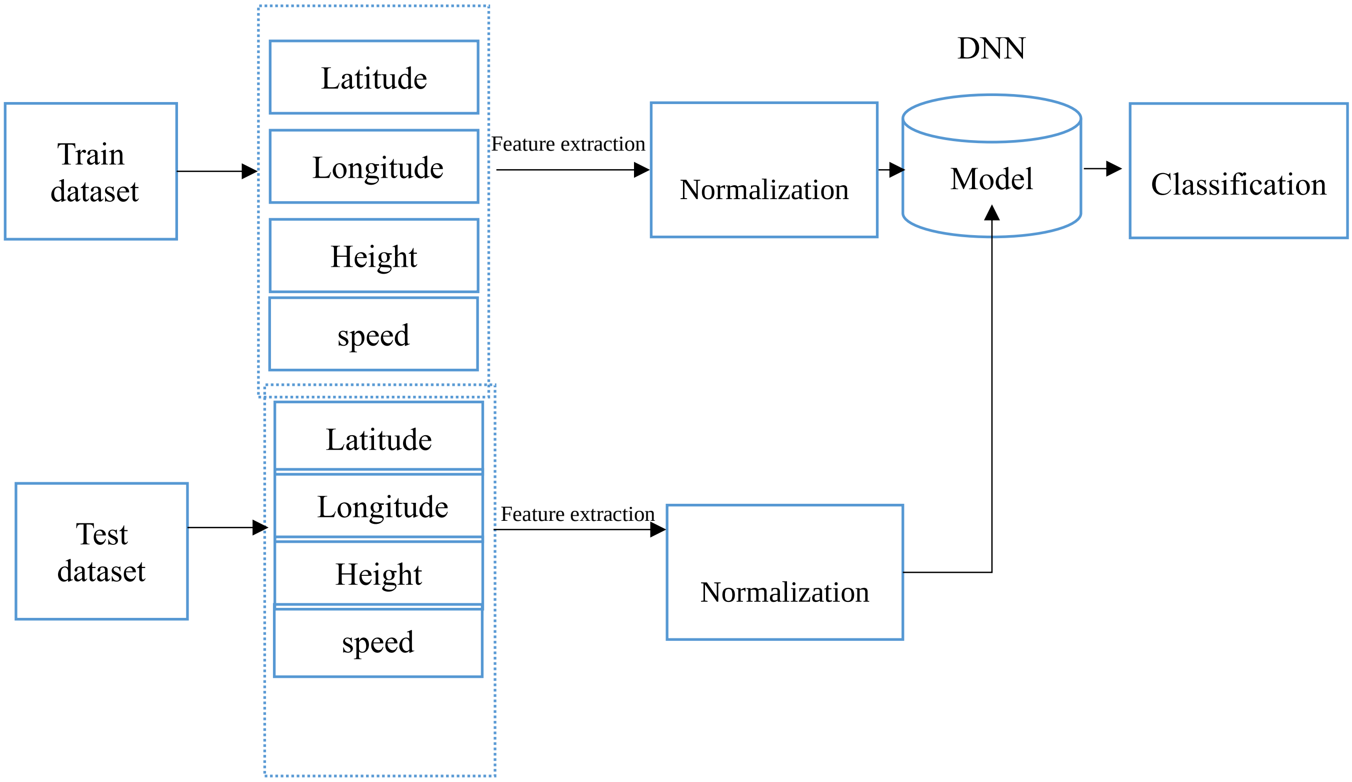

Figure 1. the block diagram of the proposed method.

The first step of the training process is to extract data regarding geographical latitude and longitude, speed, and height from ground level from the data of the database used for anomaly detection.

The second step is to extract the data features. In the third step, the extracted features are normalized. Then, the classifier is trained using the normalized features. Finally, the classification results are evaluated using the test data.

In this section, the test and training sections are studied in details.

2.1. Dataset

The dataset used in this study is collected from airplanes of the IRI airline fleet recorded in one month (Zarei and Sharifabadi, 2018).

This dataset includes instantaneous fuel consumption, instantaneous weight, geographical latitude and longitude, speed and height from ground level. Among this data, only geographical latitude and longitude, speed and height from ground level are used for feature extraction.

2.2. Feature Extraction

This section is an essential section of this study, which includes extracting features with best description of the input data.

First, some basic components of the dataset regarding geographical latitude and longitude, speed and height from ground level are extracted as described in the following.

The first extracted component is the displacement that is calculated as follows.

$$D=\sqrt{(y_i-y_{i+1})^2+(x_i-x_{i+1})^2} \quad (1)$$

In the above equation, x_i and x_i represent longitude and latitude at instant i.

The second basic component is speed, which is obtained through dividing Eq. (1) by the time interval between two instances. The third basic component is acceleration, which is obtained through dividing speed by the time interval between two instances. The next basic component is the route's angle, which is calculated using Eq. (2).

$$A=\arctan(\frac{y_{i+1}-y_i}{x_{i+1}-x_i}) \quad (2)$$

The fifth basic component is the height changes, calculated as follows.

$$H=y_i-y_{i+1} \quad (3)$$

The next basic component is the time difference that includes the distance between two instances. The main angle is another basic component, which is calculated using Eq. (4).

$$OA=\arctan(\frac{y_i}{x_i}) \quad (4)$$

The last basic component is the route's length, which is the sum of the travelled route before instant i and the route travelled at the mentioned instant. Total flight time is another basic feature, which is calculated. After calculating the basic features, they are used to extract some features from each route with maximum intra-class similarity and maximum inter-class difference. These features include changes on the inputs at two subsequent instants, dynamic speed, acceleration, angle, dynamic displacement, time variations, dynamic angle, total travelled distance, flight time, time interval of different inputs, landing time,the aircraft's initial peaking, and the time that the aircraft has not detached from the ground, maximum speed, mean speed and speed deviation, mean absolute acceleration and absolute acceleration deviation, maximum acceleration, maximum acceleration drop, horizontal speed, and the mean time that takes the airplane to enter the airport, deviation, maximum and minimum of the travelled distance.

After calculating these features, which are a total of 73 features, the next processes are applied.

2.3. Feature Normalization

This step is effective in terms of extendibility and flexibility of constructing a robust model. To normalize the algorithm, moving average is used. Eq. (5) describes this method (Alessio et al., 2002):

$$(y_k)_s=\sum_{i=-n}^{i=n} y_{k+i}/(2n+1) \quad (5)$$

In this equation, n determines the averaging window, $(y_k)_s$ and y_{k+i} are the normalized data and the previous data, respectively.

Another process applied to the raw data is data normalization as given in Eq. (6) (Singh et al., 2015):

$$Y=\frac{X-X_{\min}}{X_{\max}-X_{\min}} \quad (6)$$

In this equation, the normalized value is determined using the maximum and minimum values, which is known as the max-min normalization method.

2.4. Feature Selection

An essential part of pattern recognition methods, which reduces computational complexity and increases accuracy is the feature selection. In this study, fast correlation-based filter is used which demonstrates better speed and efficiency compared to other correlation-based feature selection methods. Therefore, using this method, the features adapt with higher dimensional data and better features are selected. In this method, asymmetrical uncertainty is calculated for all features. This value is obtained through calculating the information gain of x given y divided by the total entropy of all features. Finally, the asymmetrical uncertainty, sorting, and redundant features are removed (Selvi and Pushpa, 2020).

The result of this method is selecting 15 superior features among 73 features.

2.5. Feature Classification

The most important part of the methods based on pattern recognition is classification. In this section, the deep convolutional neural network is used, which is described in details.

Deep Convolutional Neural Network: Extracting features of higher dimension is an important part of content-based pattern recognition. Feature extraction has been applied in medicine, industry, machine vision, and control in the last two decades. Extracting high dimensional features, facilitates data analysis and representation. In recent years, various algorithms have been presented for extracting high-dimensional features. One of these approaches is using deep convolutional networks.

The convolutional networks eliminate the need for manual feature extraction; thus, it is not required to detect the features used for image classification. The convolutional network extracts features directly from the images. The relevant features are not prevented; they learn while the network operates on a set of images. This automatic feature extraction makes the deep learning models more accurate for visual tasks like object classification. Deep learning is a special form of machine learning. The machine learning workflow starts with the relevant features taken manually from the images. Then, these features are used to create a model for classifying the image objects. In the deep learning workflow, the relevant features are automatically extracted from the images.

In the proposed method, a convolutional network is used for extracting high-order features and classification.

To this end, the developers and the programmers can use various tools. One of these tools is the *ALEXNET* library. This library supports the operations of the convolutional networks in MATLAB. The convolutional neural networks are one of the most important deep learning methods, in which multiple layers are trained using a robust method. This method is very efficient and one of the most common methods in machine vision. The general architecture of a convolutional neural network is shown in Figure 1.

In general, a CNN is comprised of four main layers, including the input layer, convolution layer, pooling layer, and the fully-connected layer. Various layers perform different tasks. Feature extraction is performed in the convolution layer (Guo et al., 2009).

The input layer receives data and transfers data to the neural network. It should be noted that the number of neurons of this layer is equal with the input dimension.

The next layer is the convolution layer. The convolution layer can be considered as the eyes of the CNN. Neurons of this layer consider specific features. If they found the features of interest, they generate the up (high) activation. Finally, the output layer generally generates one as the output, this is the probability that the received data belongs to the input of the class of interest and it can be rounded to obtain the classification result (Ossai, 2020).

In this study, in addition to the mentioned classifier, support vector machine (Geng et al., 2008), decision tree classifier (Chen et al., 2020), nearest neighbor classifier (Zhang et al., 2019) and simplified Bayesian (Bosson and Nikoleris, 2018) are used to compare the result and ensure the performance of the proposed method.

In this paper, for better evaluation of the proposed method, the K-fold cross validation method is used.

The method used in this paper, is 10-fold cross validation. If the training data is randomly divided to k layers of the same size, k-1 layers can be considered as training data and one can be considered as test data at each step of the cross-validation process.

3. SIMULATION

For simulation, a hardware with *RAM:8GB, CPU: I9 4500U 2.4GHz* and *NVIDIA GeFORCE* graphical card and *32GB DDR4* is used. *MATLAB R2020a* is used for simulation.

In this section, the simulation of the proposed method is evaluated step by step in details.

3.1. Input data

The first point that should be presented is the input data used to train and test the system. Since a proper evaluation of the proposed method is achieved through applying the method on the selected database, selecting the most proper database is of great importance. The employed database is based on the data collected from the company of the country's airports, including 6 months of air data from Farvardin 1398. This data is collected from Mehrabad (Tehran), ShahidHasheminezhad (Mashhad), ShahidSadooghi (Yazd). In this paper, the data of the three airports is used and the following figure shows a schematic of the position of these airports and their entrance and exit bands. The routes used in this study include all landing and take-off routes of the airports in the eastern band of ShahidHasheminezhad and ShahidSadooghi and the western band of Mehrabad. The total number of routes is 1200.

Figure 2. The plane of HashemiNezhad Airport.

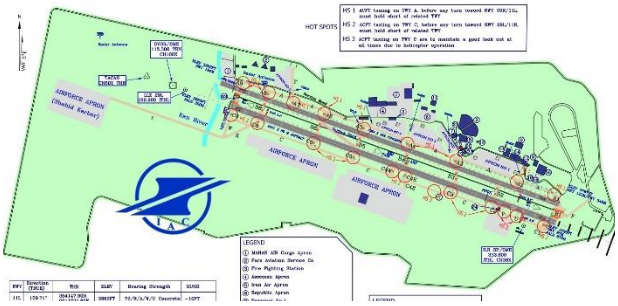

Figure 3. The plane of Mehrabad Airport.

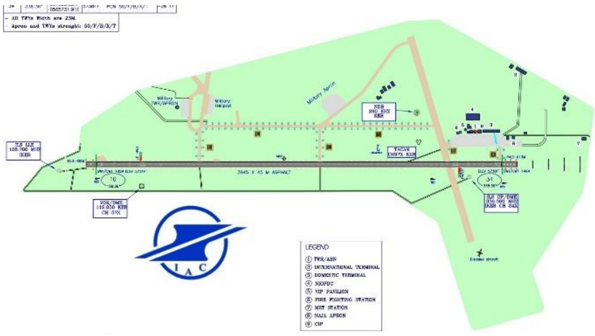

Figure 4. The plane of Shahid Sadooghi Airport.

The data is normal, because when the data is recorded, no abnormal event has occurred in the flights. However, 300 routes are added as abnormal trajectories by applying some changes in the speed and height so that the performance of the proposed method is evaluated better. It is tried to add a diversity of data to obtain a better performance of the proposed method.

3.2. Feature Extraction

In this step, 73 features are extracted from all routes. These features are described in the corresponding section; they are selected so as to provide a better description of the route. Next, it is time for feature selection. In this process, 15 features are selected from 73 features.

3.3. Classification

In this study, the deep convolutional network of ALEXNET is used for classification. One of the layers is the output layer and the others are the middle layers. Output of the first layer is the input of the second layer and the output of the second layer is the input of the third layer. Finally, the output of the third layer is the real response of the network. These neural networks are trained in two steps. First, the feedforward step and then the back-propagation error. In the feed forward step, the input enters the first layer of the network and dot product is applied to the input and the components of each neuron, and finally the convolution is applied in each layer. Finally, the network output is calculated. In this study, the network components including the neurons' weights are adjusted

to measure the network error. To this end, the network output is compared with the correct response using a loss function, and the error is determined after calculating the distance between the two results. In the next step, the back-propagation process starts based on the calculated error. In this step, the derivative of each parameter is calculated considering the chain rule and all parameters change considering their impact on the resulting error. After updating the parameters, the feedforward step starts and after a suitable number of iterations, the network training is finished.

In this study, for better comparison of the results, in addition to the neural network classifier, other classifiers like SVM, decision tree, nearest neighbor and simple Bayesian are used to ensure the performance of the proposed method.

4. RESULTS

In this section, to evaluate the proposed method, the obtained results are compared with other common machine vision methods. First, the true positive and false positive rates are calculated as follows.

$$\mathrm{TP}(TD,\mathrm{GT}) = \frac{\#(TD \cap ED)}{\#(ED)} \tag{8}$$

In the above equation, TD represents the number of anomalies obtained using the proposed method and ED represents the number of anomalies detected correctly in the database by an expert. Also, FPR is calculated using Eq. (9).

$$\mathrm{TN}(TD,\mathrm{GT}) = \frac{\#(TD \cap \overline{ED})}{\#(ED)} \tag{9}$$

Finally, using these two rates and false positive rate which is the complement of TP, and false negative rate which is the complement of TN, accuracy is calculated using Eq. (10). In this study, this measure is used to compare the proposed method with other methods.

$$Accuracy = \frac{TN + TP}{TN + TP + FN + FP} \tag{10}$$

Also, 10-fold cross validation is used in the proposed process. For summary, the average of the ten iterations is presented.

Figure 5 shows the results of the tenth iterations for each classifier.

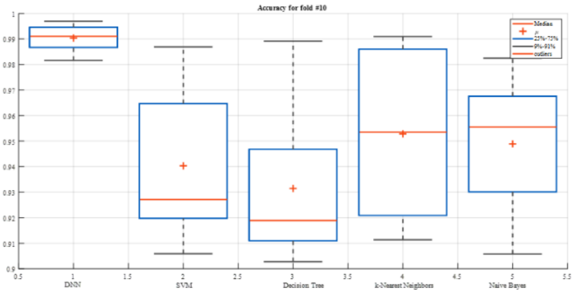

Figure 5. The results obtained for accuracy in the tenth iteration.

Also, the average accuracy of the ten iterations is shown in Figure 6.

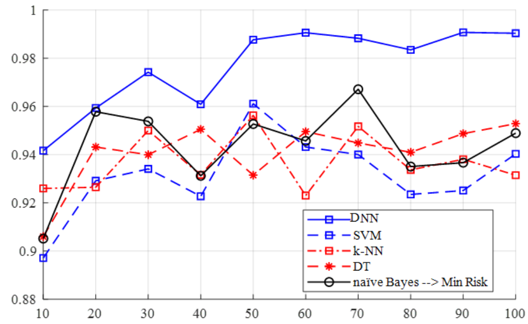

Figure 6. The average accuracy obtained for 10 iterations.

Table 1 shows the final results for five classifiers as the average of ten iterations. Evaluation of the diagrams and the above table shows that using the proposed method for detecting abnormal trajectories performs well, and its comparison with other classifiers indicates its desired performance on new data entering the system.

Table 1: Final results for five classifiers.

classifier	NB	DT	KNN	SVM	AE[4]	Alexnet
Accuracy (Avrragz)	94.24%	94.08%	93.68%	93.16%	94.12%	99.13%
Accuracy (varience)	3.02×10^{-4}	1.88×10^{-4}	1.39×10^{-4}	2.82×10^{-4}	2.74×10^{-4}	1.89×10^{-4}

5. CONCLUSION

In this study, a novel method is presented for detecting anomalies in air ways using neural networks and their proper performance in modelling complex processes. The proposed method is implemented on the database of three internal airports and the results are compared with four common pattern recognition methods. The simulation results indicate that the proposed method outperforms the other methods in detecting anomalies of air ways. An accuracy of 99.13% verifies this claim. After the proposed method, the simple Bayesian method is ranked the second, due to its excellent performance in employing previous information for creating future information.

REFERENCES

[1] Alessio, E., Carbone, A., Castelli, G., Frappietro, V., 2002, Second-order moving average and scaling of stochastic time series, *Eur. Phys. J. B*, **27**, 197-200.

[2] Barratt, Sh.T., Kochenderfer, M.J., Boyd, S.P., 2018, Learning Probabilistic Trajectory Models of Aircraft in Terminal Airspace from Position Data, *IEEE Trans. Intell. Transp. Syst.*, **20**, 3536-3545.

[3] Bosson, C.S., Nikoleris, T., 2018, Supervised learning applied to air traffic trajectory classification. in *2018 AIAA Information Systems-AIAA Infotech@ Aerospace*. https://doi.org/10.2514/6.2018-1637

[4] Chen, J.H., Wei, H.H., Chen, C.L., Wei, H.Y., Chen, Y.P., Ye, Z., 2020, A practical approach to determining critical macroeconomic factors in air-traffic volume based on K-means clustering and decision-tree classification, *J. Air Transp. Manag.*, **82**, 101743.

[5] Geng, R., Cui, D., Xu, B., 2008, Support vector machine-based combinational model for air traffic forecasts [J]., *J. Tsinghua Univ. (Sci. Technol.)*, **7**, p.35.

[6] Guo, Z., Zhang, D., Zhang, L., Zuo, W., 2009, Palmprint verification using binary orientation co-occurrence vector, *Pattern Recognit. Lett.*, **30**, 1219-1227.

[7] Jarry, G., Delahaye, D., Nicol, F., Eric Féron, E., 2018, Aircraft Atypical Approach Detection

using Functional Principal Component Analysis, *J. Air Transp. Manag*., **84**, 101787.

[8] Oehling, J., Barry, D.J., 2019, Using machine learning methods in airline flight data monitoring to generate new operational safety knowledge from existing data, *Saf. Sci*., **114**, 89-104.

[9] Ossai, C.I., 2020, Corrosion defect modelling of aged pipelines with a feed-forward multi-layer neural network for leak and burst failure estimation, *Eng. Fail. Anal*., **110**, 104397.

[10] Rose, D.K., Cohen, M.M., 1994, The airway: problems and predictions in 18,500 patients, *Can. J. Anaesth*., **41**, 372-383.

[11] Selvi, U., Pushpa, S., 2020, Big Data Feature Selection to Achieve Anonymization, in *International Conference on Communication, Computing and Electronics Systems* (pp. 59-67). Springer, Singapore.

[12] Singh, B.K., Verma, K., Thoke, A.S., 2015, Investigations on impact of feature normalization techniques on classifier's performance in breast tumor classification, *Int. J. Comput. Appl*., **116**, 11-15.

[13] Zarei, K.K., Sharifabadi, R., 2018, Feasibility study of creating an archive in the accident investigation office of the Civil Aviation Organization, *Document Treasure*, **28**, 125-165.

[14] Zhang, F., Kodituwakku, H.A.D.E., Hines, J.W., Coble, J., 2019, Multilayer data-driven cyber-attack detection system for industrial control systems based on network, system, and process data, *IEEE Trans Industr Inform*., **15**, 4362-4369.