# Aviation Anomaly Detection Using Deep Convolutional Neural Network

**Seyed Mohsen Kamali Firoozabadi[1], Majid Poorahmadi[2], Ali Mohammad 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 (Corresponding author)
3- Department of Electrical Engineering, Yazd University, Yazd, Iran.
Email: alatif@yazd.ac.ir

**ABSTRACT:**
Due to aviation events caused by human mistakes, unwanted hardware failures, or criminal and terrorism intentions, a huge cost is incurred to governments annually. Due to the wider and more comprehensive tools, automatic detection methods can help us to reduce such problems considerably by analysis and evaluation of the events. These tools also play a key role in decreasing human errors and preventing terrorism and criminal cases. The present research evaluates the behavior of aircraft inside the airport to detect anomaly routes and situations. The proposed approach based on novel innovative Convolutional neural network, and autoencoder neural networks, can be implemented on airplanes with numerous arrival and departure lines, i.e. high-traffic airports such as John F. Kennedy airport. Abnormal behaviors identified using this scenario will play a significant role in reducing accidents within the airports. To assess the validity of the suggested method, this paper utilizes the dataset of aerial routes within the John F. Kennedy airport in the U.S. The method is compared with three conventional approaches in this field. Achieve to the 86 percent true positive rate and 89 percent area under the curve on test data, is an evidence for effectivity of method.

**KEYWORDS:** aircraft route anomaly detection, autoencoder neural networks, deep convolutional neural networks, pattern recognition.

***Tob Regul Sci.™ 2021;7(6-2):* 13-20**
***DOI: doi.org/10.18001/TRS.7.6.2.2***

## 1. INTRODUCTION

One of the major challenges in most developed communities, where a large number of people employ airlines to travel and commute daily, is to detect anomalies along the route to detect and prevent events that can sometimes lead to irreparable consequences. Some of such events include hijacking, technical damage to the fuselage and engine of the aircraft, unforeseen atmospheric phenomena, etc. These events can be assessed by analyzing the route the aircraft travels and measures can be taken to resolve, remedy, or mitigate the severity of the damages. The process consists of several steps, the first step of which is to extract the route based on the data over a long period. In the next step, modeling is performed based on the data. Eventually, the compatibility of the model with the newly extracted data is evaluated and examined.

The paper is organized as follows. The literature review describes some of the recent research conducted in the field. Then, the next section focuses on the proposed method and presents it with more details. The simulation section explains the dataset used in the work and deals with simulation steps. Next, the obtained results are provided and analyzed and, in the end, conclusions are summarized.

## 2. LITERATURE REVIEW

Various pieces of research have been presented in the field of route-based anomaly detection. This part of the paper introduces a report on the wide research conducted concerning anomaly detection techniques focusing on the behavior of vehicles. The research can be categorized into different groups, such as the Bayesian model, support vector machine (SVM), nearest neighborhood, clustering, Markov method, and Gaussian mixture models (GMMs), and sparse reconstruction model.

In the Bayesian model, using prior information and statistical distribution of available data, appropriate classification can be considered for available often single-variable data. In other words, Bayesian model-based methods will act as normative or anomalous data labeling, in which the probability of belonging to the relevant label will be estimated by methods based on Bayes' law.

In [1], a type of nonparametric Bayesian modeling is introduced for route analysis and modeling of the active area of the image using unsupervised method. Active areas consist of parts of the image where there are routes of moving objects. In this method, routes are considered as documents, and routes of objects in the image are assumed as words in the image. The use of these two components leads to the achievement of a desirable model for the dictionary, which can be used to identify routes that do not conform to conventional models in the anomaly detection process. In [2], using the GMM and conventional clustering methods and incorporating the data related to the behavior of aircraft inside the airport, abnormal routes are categorized and identified. The advantage of this method is its compactness and high speed in detecting abnormal routes.

Authors in [3] adopt the Functional Principal Component Analysis to implement a practical approach to analyzing and analyzing the behavior of aircraft during landing and take-off at airports. The advantage of this method is the use of effective components in detecting abnormal processes, which will ultimately lead to the achievement of an efficient system in anomaly detection. Coding neural networks have been adopted in [4] to design an efficient and optimal anomaly detection process, which in addition to operating the system much faster in the system training phase, provides much higher accuracy for the anomaly detection process. It should be noted that in the present paper, deep convolutional neural networks are exploited for this process.

## 3. PROPOSED METHOD

In this section, first, the proposed method for implementing the anomaly detection process will be discussed and its different parts will be examined separately.

The suggested method consists of two main parts: training and testing. In the former part, the system is trained using the available training data, and in the next step, data classification, i.e. the anomaly detection process is addressed using the designed model.

The Fig. 1 illustrates the block diagram of the proposed solution for CNN part.

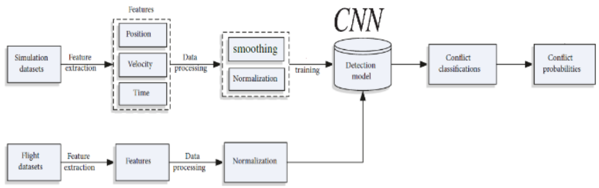


**Fig. 1**. The block diagram of the proposed method.

Preprocessing

The first step in the training phase is to extract the velocity and time characteristics from the data in the database used for the anomaly detection process.

The second step is the preprocessing of raw data, which includes the smoothing and normalization of the extracted time series. The smoothing of the time series resulting from the extracted features assists us to achieve generalization and proper flexibility to build a powerful model. We use the moving average algorithm to smooth the data. The following equation express this method [5]:

$$\left(\mathrm{y}_{\mathrm{k}}\right)_{s}=\sum_{i=-n}^{i=n}\mathrm{y}_{\mathrm{k+i}}\,/(2n+1) \qquad (1)$$

where, n denotes the component of determining the averaging window, and (yk)s and yk+i represent the softened data and the previous data, respectively.

Another process performed on raw data is data normalization, which is implemented using the following formula [6]:

$$Y=\frac{X-X_{\min}}{X_{\max}-X_{\min}} \qquad (2)$$

In this equation, the normalized value is found using the maximum and minimum values, which is known as the Min-Max normalization method.

### 3.1. Convolutional Neural Network

These types of neural networks are one of the most important and practical methods used in the field of deep learning in which several layers are trained using coherent principles and rules. These networks are very efficient and have a variety of applications in many areas of machine vision. In general, a convolutional network consists of three main layers: convolution layer, pooling layer, and fully connected layer. Each of these layers has a specific task [7].

To train the aforementioned neural networks, the training process consists of two phases: the feed-forward phase and the back-propagation phase. In the forward phase, the input enters the initial layer of the network and the point multiplication process takes place between the input and the components of each neuron and finally the convolution process is performed in each layer. In the end, the network output is calculated. In this step, to adjust the network parameters, or in other words, the training process, the output result is incorporated to measure the network error. To this end, the network output is compared with the correct answer using a loss function, and thus the error rate is calculated. In the next step, based on the calculated error, the error back-propagation phase begins. In this step, the derivative of each parameter is calculated according to the chain rule and all parameters are changed according to the effect they have on the error created in the network. After updating the parameters, the feed-forward stage begins, and finally, after repeating a suitable number of these steps, the network training is completed.

In the training phase, after entering the data under test into the input of the pre-trained neural network, the extent to which each data belongs to different classes is determined at the output of the end layer, which is the fully connected layer.

### 3.2. Autoencoder

Autoencoder (AE) neural networks are among the most widely adopted neural network structures for a broad range of pattern recognition applications, including classification, clustering, feature compression, and data reconstruction. The general structure of these networks is shown in Fig. 2.

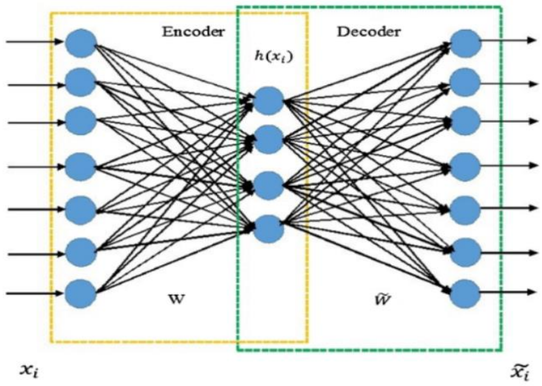


**Fig. 2**. The Structure for Autoencoder Networks [8]

As Fig. 2 shows input ($x_i$) and output ($\widetilde{x_i}$) have a same dimension while $h(x_i)$ has a different dimension. Generally, $h(x_i)$ is in a lower dimension than others. You can find More information in [9].

### 3.3. Anomaly detection process

Once the matrix of the extent to which each data belongs to the existing classes is established, using the analysis performed, it can be inferred whether the new data is within the specified normal classes or not. This process is further examined below.

The first item to analyze is that the maximum probability in the above matrix, which indicates belonging to a certain class, must be greater than the reliable threshold. This threshold is represented by Tr1.

The next item to consider is that data can never belong to two different classes at the same time, which is to be expected given the definition of different classes for training data. It is worth noting that observing a threshold in this criterion is mandatory. This threshold is denoted by Tr2 and this condition is also denoted by the following equation.

$$\text{dependency1} - \text{dependency2} > \text{Tr2} \quad (3)$$

In the above equation, dependency1 and dependency2 represent the maximum belonging and the highest belonging before the fully connected layer is available in the output matrix.

Finally, if the two thresholds for the test data are met, the data is normal; otherwise, if even one of these two conditions is not met, the data is introduced as abnormal.
With this process we can compute a score for anomaly.

For second part we compute anomaly with AE. AE networks offer excellent results in unsupervised conditions, which can be utilized in the best manner in this article. Moreover, in the structure of such networks, the number of input and output neurons is equal. Plus, an AE is trained to adjust the output to the input as much as possible. Thus, the other advantage of autoencoder networks is that an AE is forced to generalize data and seek common patterns. Therefore, by comparing the reconstructed data in the output with the input data using specific criteria, we can determine whether a given data can be classified as training data or not. From this session also we can compute another score for abnormality of trajectory.

Finally with averaging process we can find weather trial trajectory, is normal or anomaly.

## 4. SIMULATION

For simulation, hardware with features CPU: I7 4500U 2.4GHz, RAM: 8GB, NVIDIA GeForce 820M graphics card, and 6GB dedicated memory with a dedicated 2 GB DDR3 for the graphics card is used. The software is Python and the TENSORFLOW-JUPITER environment.

In this part, the simulation of the proposed method is presented step by step and its details are evaluated.

### 4.1. Input Data

The first item that should be discussed in this part is the input data used to train and test the system. Since the proper evaluation of the proposed method is shown only by applying it to the selected database, selecting the most appropriate database is very important. The database used includes the data collected by the US Federal Aviation Administration (*FAA*), known as *FAA radar* [10], which contains aerial data 6 months from March 2012. The data was collected from three airports in northern and southern California, as well as New York City. In this paper, only the data of New York John F. Kennedy Airport is used. The Fig. 3 depicts its location along with its landing and take-off runways.

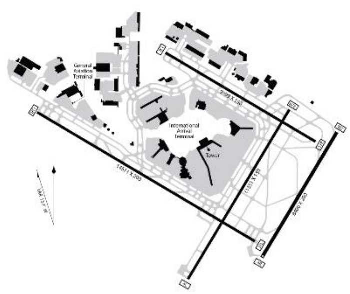

**Fig. 3.** New York John F. Kennedy Airport

The routes used in this paper are all take-off and landing routes on the eastern runway of this airport. In the Fig. 4, some of these trajectories can be seen.

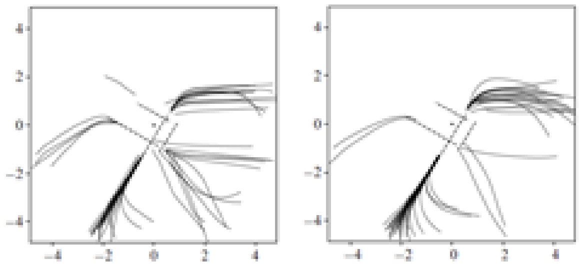

**Fig. 4.** Some of trajectories in dataset

### 4.2. Convolutional Neural Network and AE Network

The second step is to determine the components of the deep neural network for modeling. In the following, information about the components of convolutional neural network training is provided:

- The size of the filters used are 7, 8, and 9, each of which is implemented with 100 different weights (300 feature maps in total)
- The dropout rate is 0.5
- The size of each mini-batch is 50
- SGD 19 and the Adaddelta rule have been used to update the weights.
- RELU activation function on all layers (except the last layer)
- Training rate: 0.03120

In the following, the simulation results of the proposed method are presented on the desired base.

It should be noted that a total number of 500 epochs are considered for the training of AE network.

### 4.3. Results

At this stage, to evaluate the proposed method, the results obtained with other conventional machine vision methods are evaluated according to common criteria such as false positive rate (FPR), true positive rate (TPR), and the area under the curve (AUC). The first criterion used, i.e. TPR, is calculated using the Eq.4.

$$\mathrm{TPR}(TD, \mathrm{GT}) = \frac{\#(TD \cap ED)}{\#(ED)} \tag{4}$$

In the above equation, TD indicates the number of anomalies obtained by the proposed method, and ED also indicates the number of anomalies detected in the database by an expert. Additionally, another criterion to be evaluated is FPR according to Eq. 5.

$$\mathrm{TPR}(TD, \mathrm{GT}) = \frac{\#(TD \cap \overline{ED})}{\#(ED)} \tag{5}$$

Another tool that can be used to evaluate the performance of the proposed method graphically and visually is the ROC curve. The closer the values of the vertical axis of this curve are to one, the better the performance of the method. Using it, another component is introduced as the AUC, which indicates the overall performance of the system, and the closer this value is to one, it can be inferred that the proposed method has better performance. Table 1 tabulates the comparison of the results between the suggested method with two SVM [11] and (K-nearest neighbor) KNN [12] modeling methods and also hidden Markov model (HMM) [13]. The Fig. 5 shows the ROC curve for the proposed method and the other three methods.

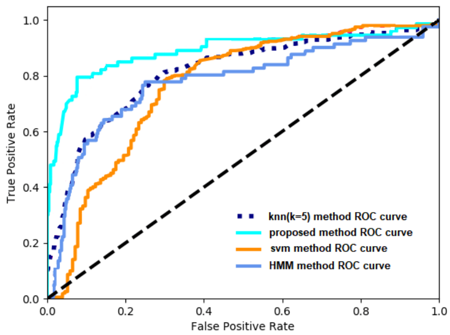


**Fig. 5.** Roc plot in simulation result

The table lists the values for TPR and FPR according to the most appropriate case, which includes the highest TPR rate for the lowest FPR.

**Table 1.** FPR and TPR of simulation result.

| Method | FPR | TPR |
|---|---|---|
| Proposed method | 0.2133 | 0.8622 |
| HMM | 0.2617 | 0.7935 |
| Svm | 0.2681 | 0.7407 |
| Knn(k=5) | 0.2513 | 0.7843 |

As is observed in the following table, the values of the area under the curve are significant.

In the next section, which is dedicated to the conclusion, the above cases are analyzed and a comprehensive evaluation of the performance of the proposed method in comparison with other cases is presented.

**Table 2.** AUC of simulation result.

| Method | AUC |
|---|---|
| Proposed method | 0.8921 |
| HMM | 0.7835 |
| Svm | 0.7683 |
| Knn(k=5) | 0.8172 |

## 5. CONCLUSION

In this paper, using deep neural networks and benefiting their outstanding performance in modeling complex processes  and ability of AE in finding patterns, a novel method is presented for detecting airway abnormalities, which are implemented in the database of John F. Kennedy Airport in New York. We also compared it with three conventional pattern recognition methods. The simulation results highlight that the proposed method provides superior performance in the airway abnormalities detection process. Performance comparison was performed using three TPR, FPR, and AUC components. The proposed method has the best performance with 7% and 4% improvements in TPR and FPR rates, respectively. The other component that was evaluated was AUC, in which an 8% improvement is observed in the proposed method compared to the best method. The results all indicate that the application of the proposed method to the airway abnormalities detection process can help reach suitable results.

## REFERENCES

[1] A. Gopnik, and J. B. Tenenbaum, **"Bayesian networks, Bayesian learning and cognitive development,"** *Dev. Sci.*, vol. 10, no. 3, pp. 281-287, 2007.

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

[3] G. Jarry, D. Delahaye, F. Nicol, and E. Féron, **"Aircraft Atypical Approach Detection using Functional Principal Component Analysis,"** *J. Air Transp. Manag.,* vol. 84, 101787, 2020.

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

[5] E. Isufi, A. Loukas, A. Simonetto, and G. Leus, **"Autoregressive moving average graph filtering,"** *IEEE Trans. Signal Process*, vol. 65, no. 2, pp. 274-288, 2016.

[6] F. Jean, A. Branzan Albu, and R. Bergevin, **"Towards view-invariant gait modeling: Computing view-normalized body part trajectories,"** *Pattern Recognit.*, vol. 42, no. 11, pp. 2936-2949, 2009.

[7] Sh. Zhou, W. Shen, D. Zeng, M. Fang, Y. Wei, and Zh. Zhang, **"Spatial–temporal convolutional neural networks for anomaly detection and localization in crowded scenes,"** *Signal Process. Image Commun.*, vol. 47, pp. 358-368, 2016.

[8] H. O. Ahmed, M. L. Wong, and A. K. Nandi, **"Intelligent condition monitoring method for bearing faults from highly compressed measurements using sparse over-complete features,"** *Mech. Syst. Signal Process.*, vol. 99, pp. 459-77, 2018.

[9] S. Kelathodi Kumaran, D. Prosad Dogra, P. Pratim Roy, and A. Mitra, "**Video Trajectory Classification and Anomaly Detection Using Hybrid CNN-VAE,"** arXiv e-prints, arXiv-1812, 2018.

[10] C. Jagodnik, J. Stella, and D. Varon, **"Fusion tracking in air traffic control,"** *J. Air Traffic Control*, vol. 50, no. 1, May 2008.

[11] C. Piciarelli, and G. L. Foresti, **"Anomalous trajectory detection using support vector machines,"** *2007 IEEE Conference on Advanced Video and Signal Based Surveillance*, pp. 153-158.

[12] V. Chandola, D. Cheboli, and V. Kumar, **"Detecting anomalies in a time series database,"** Technical Report, pp. 12, 2009.

[13] N. P. Cuntoor, B. Yegnanarayana, and R. Chellappa, **"Interpretation of state sequences in HMM for activity representation,"** *Proceedings (ICASSP'05). IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2005, IEEE, vol. 2, pp. 702-709.