# Machine Learning based Predictive Algorithm for Air Traffic Control and Landing Assistance

*Abstract*—**The most crucial phases of a flight are taking off, reaching its destination safely, and landing. It's in the hands of the captain and the first officer to do all these things perfectly and take us to our destination safely, but nature is always unpredictable. There are adverse weather conditions that can happen after taking off and can put thousands of lives in danger. Therefore, we have come up with a solution that helps pilots make decisions instantly and find the route that is optimal for their fuel capacity and can help them escape the bad weather and reach their destination with protection. There is something called Air Traffic Flow and Capacity Management (ATFCM), which belongs to the Air Traffic Management (ATM). ATFCM will likely make airspace and air terminal limit satisfy traffic need and, whenever limit openings are depleted, advance traffic streams to meet the accessible proportions. Any aircraft that is in the air will be connected with one or the other ACT (Air traffic controller) throughout the journey, but if in an adverse weather condition if the connection is lost our model will help the pilot to have the situation under benchmark. Having that in mind, we planned to bring in automation in route prediction for flight. The way the aircraft is routed is a pattern and these pattern designs are called "rotations". Network design is both a science and an art and machine learning (ML) can be applied to network design patterns. ML can analyse aircraft patterns and the commercial, financial and operational implications thereof. Five common patterns are found in airline networks, where each of these patterns follows a certain straightforward approach to calculate Commercial impact, operational impact, and Disruption cost. The proposed methodology helps ATC management and flight pilot to give pre-tactical traffic information in the form of graph visualizations, which will give information about pattern cost for every possible path along with weather updates. So, at any adverse condition(weather changes) the ATC can help the pilot choose the best pattern for the travel, or in the unfortunate case where the pilot lost the connection with ATC can make the decision himself with graph view and the most crucial phase of a flight is landing, the safe landing of an aircraft is in the hands of the captain and the first officer, there can be adverse weather conditions during which people in the cockpit may not be able to view the runway clearly or may not even locate it precisely. Therefore, we provide a model which helps locate the runway in bad weather conditions.**

**To make it automated we bring in ML and DL together to find out the best suitable algorithm which will represent the graph with distance information using graph neural networks and to calculate the overall impacts regression algorithms can be implemented. Once the path is decided and the aircraft is safely approaching its destination we use in machine learning algorithm YOLOv3 to detect the runway in bad weather, we provide satellite images of the destination and train the algorithm in such a way that it detects the runway in which the aircraft must land from a distance itself. It can be further extended to detect the lane, when the aircraft approaches the runway, for this, we first do the pre-processing of the video seen from the cockpit and use the video for edge detection using Kalman filter.**

***Keywords—Air Traffic Flow Management, Flight Route Prediction, Runway Prediction, Air transportation, Airline industry, Aviation sector, Origin-destination, Airline networks***

## 1. Introduction

With the increase in air travel around the world and the emergence of new airlines and runways, there has been a significant increase in air traffic in the past decade. Thus the work of air traffic controllers increases day by day. More than 1,000 scheduled airlines were operating more than 15,000 aircraft at the dawn of the 21st century. The commercial airline industry carried 1.6 billion passengers and 22 million tons of cargo annually, about 40% of the world's manufacturing exports based upon value. There were 4.5 billion scheduled passengers carried in 2019 contributing to a two-thirds ratio of the whole population during 2019. If this continues it is expected to grow to about 10.0 billion by 2040.

Aviation is intermingling disparate economies and cultures, stimulating cultural and social cross-fertilization, diversity, and economic growth for a global environment. The aviation industry supports $2.7 trillion (3.6%) of the world's gross domestic product (GDP). If aviation were a country, it would rank 20th in size by GDP. Since they offer regular services for transporting passengers or goods, these companies make a root for the airline industry, which is a sub-sector of the aviation sector. Air travel demand is not a flight leg in an airline network, it is an origin‑destination market, scrutinizing airline networks create complications of market demand and supply. So a properly estimated demand model allows airline networks to more accurately forecast demand in an origin-destination market. The change in weather made lots of flights canceled at airports or even in the sky in recent years. The climatic change had fallout even in sea level leading to declining in crop yielding and affecting aircraft even in the departure of aircraft from runways affecting costs. Nowadays heat waves are happening due to a rise in temperature and natural calamities with greater frequency. If the same continues, over the decades this would increase enormously. Since the runways are built near the sea, this causes a major risk of flooding often. Stronger winds often change the directions and lengths of a journey which leads to an increase in fuel consumption.

To make the final new route from one of the planned routes earlier and get to the correct runway path is a challenge. This work proposes a methodology for solving runway prediction and direction among the routes in an air travel network using the factors like cost, distance, direction, and weather parameters like humidity, rainfall, wind, and storm. This study uses clustering techniques for traffic planning, YoloV5 for detecting runway efficiently, and can be extended to use Kalman-filter which annotates runway paths.

## II. RELATED WORKS

Various factors have been used previously for air traffic control and travel assistance.

### *A. Route Prediction :*

The strategy for discovering the path with the least energy costs utilizing A* search algorithm Rodrigo Marcos anticipated the carriers' course decisions between two air terminals as a component of the attributes of each route with the help of Air Traffic Flow and Capacity Management (ATFCM) was proposed by Xie Z.. Sukrit Sriratanawilai used logistic regression, support vector machine, and neural network are used to predict existing or non-existing routes in the United States. Craig Freudenrich proposed how air traffic control works that are run by the Federal Aviation Administration (FAA). Using a set of real-time data comprised of, network, demographic, traffic, land use, and weather features, Amir Bahador Parsa works on using eXtreme Gradient Boosting (XGBoost), a Machine Learning (ML) technique was used to detect the occurrence of accidents. Reduces the risk of runway incursion, confusion, and excursion, resulting in a more efficient taxi, takeoff, approach, and landing operations, and safer figured out by Sam Clark was done using flight deck design improvements.

### *B. Runway Prediction :*

Ö. Aytekin used Adaboost-based selected feature subset used for detecting runways and identifying their textural characteristics. Andrew J. Moore∗ applied a Speeded-Up Robust Features feature detector for the problem of identifying a particular runway from a universe of known runways. Changjiang Liu 1 introduced ongoing runway recognition implanted in engineered vision and a ROI (Region of Interest) based level set strategy. Guanhao Yang o replaced manual inspection with a deep learning method and use YOLOV5 to recognize face masks. Techzizou performed custom YOLOv4-tiny object detector design training their custom detector for mask detection using YOLOv4-tiny and Darknet. Pritul Dave1 analyzed that YOLOv4 with the XGBoost the algorithm produces the most precise outcomes with a balance of accuracy and inference time.

### *C. Lane Prediction :*

An end-to-end three-task convolutional neural network (3TCNN) helps to localize the lane and guidance to road was given using local and global Hu moments which are segmented lane objects, proposed by Der-Hau Lee. Kalman filter for detecting Indian road lanes at night was done after gamma correction for each video frame to set the light intensity. A linear estimator was also applied for the clustered lines for the lanes to make them stable and free from any offset errors. Deep learning and Cascade Particle Filter for Indian road lane detection was proposed for Histogram Equalization on Hue channel of HSV color space looking good but still dark.

Most of the previous works were focused on estimating air travel demand for existing routes and route existence in air travel networks, based on socioeconomic factors. Nonetheless, this work studies air traffic control and landing assistance. This helps airlines find out more about nonexisting routes that can be added more for ease of the assistance.

## III. METHOD

### *A. Route Prediction Models*

This work proposes to use three models from machine learning to predict route existence in air travel networks. The models are a naive Bayes classifier, support vector machine (SVM), and clustering methods. Each model is described as follows:

- *Naive Bayes classifiers*:They are a group of classification algorithms based on Bayes' Theorem. It is a family of algorithms where all of them share a common principle, i.e. features being classified that are independent of each other.

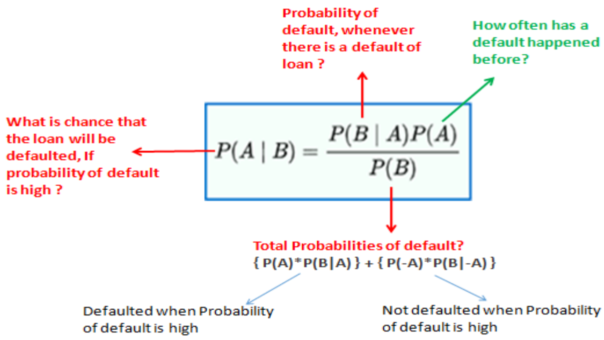



Fig.1. Bayes Theorem

- *Support Vector Machines(SVMs):* SVMs are a supervised learning strategy that is broadly utilized for classification and regression procedures. They have a place with a group of generalized linear classification. A unique property of SVM is, SVM simultaneously minimizes the empirical classification error and maximizes the geometric margin. So SVM is termed as Maximum Margin Classifiers.On each side of the hyperplane, two hyperplanes that are parallel are produced that differ the information. The one that amplifies the distance between the two parallel hyperplanes is the separating hyperplane. A thoery is made that the bigger the edge or distance between these parallel hyperplanes the better the speculation mistake of the classifier will be. We consider data points of the structure. The following Equation(1) and (2) describes about the parallel hyperlanes,

$$w.x + b = 1 \quad (1)$$

$$w.x + b = -1 \qquad (2)$$

Then, at that point SVM tracks down a linear isolating hyperplane with the greatest edge in this higher measurement-space. There are different kernel functions that are used in SVM which are listed below.

- Linear kernel: $K(x_i, x_j) = x_i^T x_j$.
- Polynomial kernel: $K(x_i, x_j) = (\gamma x_i^T x_j + r)^d, \gamma > 0$
- RBF kernel : $K(x_i, x_j) = \exp(-\gamma \| x_i - x_j \|^2), \gamma > 0$
- Sigmoid kernel: $K(x_i, x_j) = \tanh(\gamma x_i^T x_j + r)$

Here, $\gamma$, r and d are kernel parameters

- *Gaussian mixture models (GMMs):*GMMs are frequently used methods for data clustering. GMMs are widely used for performing hard and soft clustering techniques on data points. Interiorly, a Gaussian mixture model is very similar to a k-means clustering algorithm, which uses an expectation-maximization (EM) approach which qualitatively performs the below-given functionality: Initiate by choosing guesses for the location and shape, repeat until it is converged.

  Expectation (E-step): For every data point, find weights encoding the probability of membership in each cluster.

  Maximization (M-step): For each cluster, update its location, normalization, and shape based on all data points, making use of the weights.

  The end result of this is that each cluster is mapped not with a hard-edged sphere, but with a smooth Gaussian model. Merely like in the k-means expectation-maximization approach, this algorithm can also sometimes miss the globally optimal solutions, and thus in practice, multiple random initializations are used.

*B.Runway Prediction Model:*

This method uses the YOLOV5 algorithm to detect the runway where the aircraft is supposed to land after reaching its destination.YOLOV5 is considered by a wide range of people for real-time object detection among many computer vision and machine learning algorithms because of its good performance, and ease of use, and versatility. The main aim of the YOLO algorithm is to annotate the resultant object with a rectangular box, in order to do that, the algorithm divides the input image into the NXN grid system. Each cell predicts the boundary for the detected object and each object has five attributes namely, x and y coordinates, width and height, and a confidence score for the probability that the box contains the required object in Figure(2).

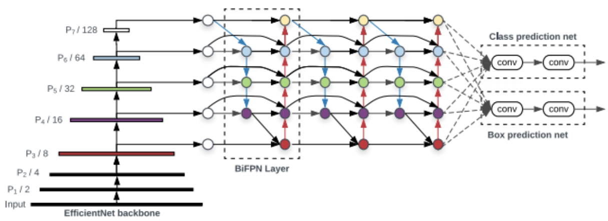



Fig.2. Runway Classification

*C.Data Collection:*

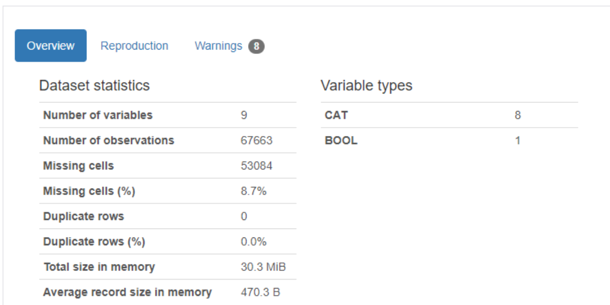
The dataset for air-travel route prediction is focused on the United States of America (US). The Airline Route/OpenFlights Mapper Route Database contains around 59000 routes between 3200 airports on 531 airlines spanning the globe and is collected from Kaggle for an experiment. The data is ISO 8859-1 (Latin-1) encoded. Another database has Fight Take-off data in JFK Airport records. This file contains data about flights leaving from JFK airport between Nov 2019-Dec-2020. The dataset for runway prediction was our own dataset made by us. We have taken the satellite images of different runways and trained the algorithm to detect the runway in Figure(3).

*Satellite image of runway*

Fig.3. Satellite image of runway

*D.Data Pre-Processing and Exploration:*

Once the process of data collection is completed, the next step should be data pre-processing and feature extraction. It is always necessary to train models with features that highly contribute to the output prediction and eliminate features that don't contribute to model training and output. With good features, we can improve accuracy, avoid overfitting risk, speed up the training process and develop better visualization reports. Initially, load the data and explore the dataset thoroughly. It is good to use Profit Report from the pandas' library which gives an overview. The below diagram in Figure(4) shows the overview of the Route dataset.



Overview Reproduction Warnings 8

Dataset statistics

| | |
|---|---|
| Number of variables | 9 |
| Number of observations | 67663 |
| Missing cells | 53084 |
| Missing cells (%) | 8.7% |
| Duplicate rows | 0 |
| Duplicate rows (%) | 0.0% |
| Total size in memory | 30.3 MiB |
| Average record size in memory | 470.3 B |

Variable types

| | |
|---|---|
| CAT | 8 |
| BOOL | 1 |

Fig 4. Dataset Information

Apply feature extraction techniques like Principle Component Analysis (PCA) and extract features that are crucial for model training. In our case, in the route database, we primarily checked for the top 15 airlines and taken data for the clustering Process.

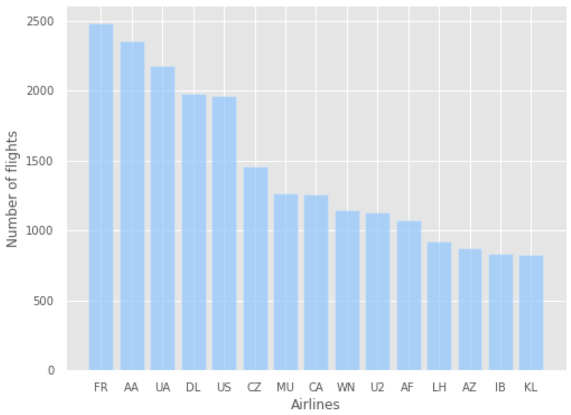



Fig.5.Top 15 Airlines in World Route DB

The possible pre-processing steps for runway detection include using Gamma correction for increasing the visibility of the image while during nighttime, defining the region of interest to detect only the runway and finally we can do noise reduction to make the image smoother.

*E:Performance Evaluation:*

A graph within a graph is an “inset”, not an “insert”. The word alternatively is preferred to the word “alternately” This work uses accuracy and area under the Receiver Operating Characteristic curve (AUC) to compare performances among three learning models. As a classification problem, we compared whether the predicted class (Ƹ) of each route is similar to the actual class () of the same route or not. When the predicted positive (or negative) class is the same as the actual positive (or negative) class, it is counted as true positive (TP) (or true negative (TN)). On the other hand, when the predicted positive (or negative) class is not the same as the actual negative (or positive) class, it is considered as false positive (FP) (or false negative (FN)). Accuracy represents a ratio between the number of correctly predicted routes to all routes.

Accuracy= correct predictions/all predictions

Receiver Operating Characteristic (ROC) is a ratio between false positive rate (FPR) and true positive rate (TPR). FPR and TPR are computed as shown in Equations (3) and (4) respectively.

$$FPR = FP/ FP+TN \tag{3}$$

$$TPR = TP/TP+FN \tag{4}$$

There are different forms in Figure(6). of YOLOV5 namely, YOLOV5s(small), YOLOV5m(medium), YOLOV5l(large), YOLOV5x(xLarge), for predicting the runway we have used YOLOV5s since it is the fastest among the four models and when it comes to predicting where the runway during the bad weather the time taken by the model to predict it is more important the pilot cannot stand still in air or hover over a fixed location before landing, because while landing the aircraft during strong wind the pilots will need to use high airspeed Since the strong wind can change its direction and strength quite frequently if the wind speed suddenly drops the plane might lose lift and crash. So we take the speed of the algorithm as an important factor here.

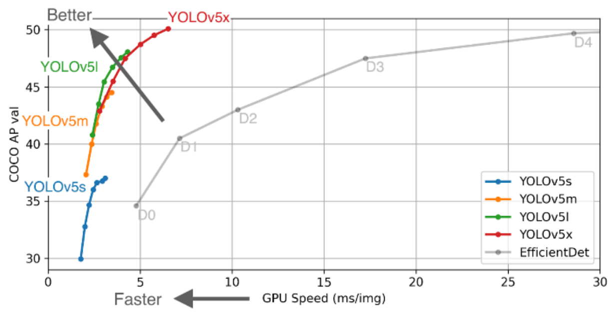



Fig.6.YOLOV5

*F:Experimental Results:*

*Route Prediction Experiment:*

Following the pre-processing step, Target Label is added based on the weather parameters like Wind Speed, Wind Gust, and Air Pressure which majorly contributed to Air Traffic in aviation with appropriate conditions. A typical network view of the route database is pictured below. The dataset is split into an 8:2 ratio, 80% training dataset, and 20% testing dataset

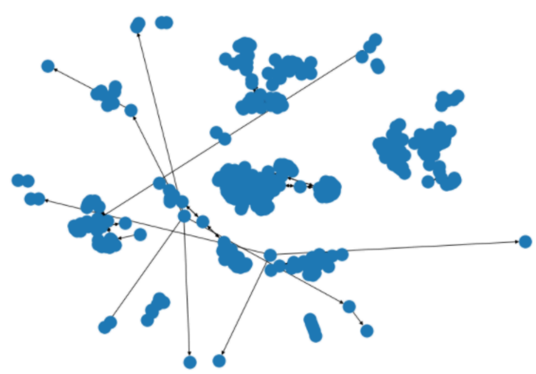
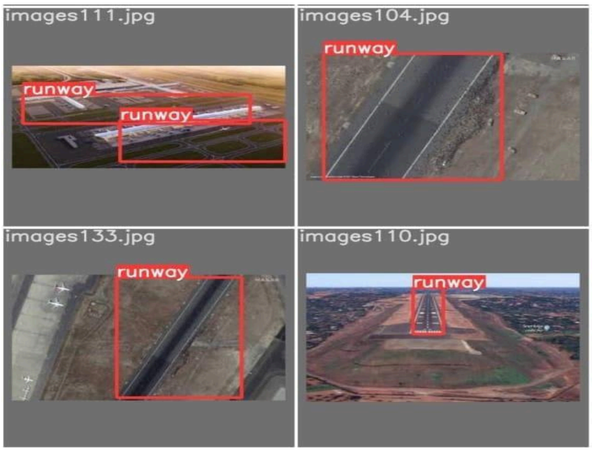


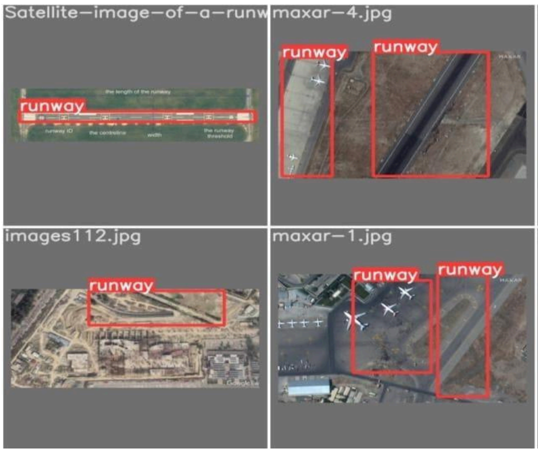
Fig.7. Network View Of Routes

- Naive Bayes Classifier is applied on the training dataset and the trained model is taken predicting Labels given the test dataset. The model worked pretty well for the data by giving binary solutions like yes or no which typically means whether to prefer that route or not for every data point considering weather factors and table 1.1 shows the accuracy measure for the dataset

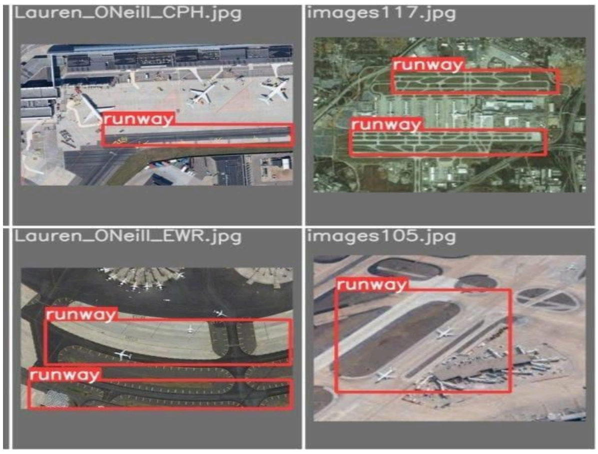
- Support Vector Machine (SVM) is used on training dataset with the same intent as a naive classifier and the well-trained model is used for performing prediction for the test data points and the table 1.1. shows the accuracy measure
- Gaussian Mixture Model (GMM) is a clustering approach applied on a training dataset to get a cluster of the same airlines and applied regression algorithm to get possible routes for the same source and destination by taking possible hotpots/stops into account and the model is made to suggest the best route by considering weather factors



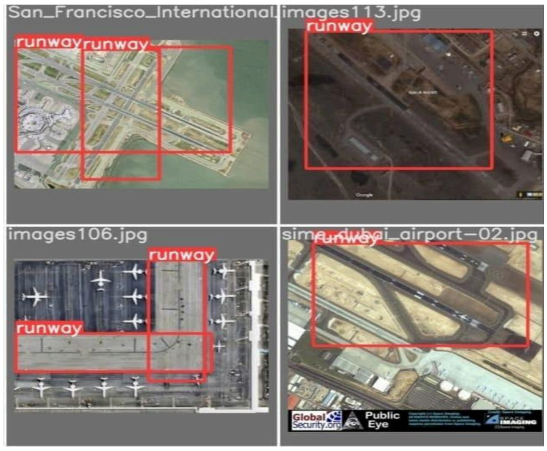

(a)




(b)




(c)




(d)

Fig.8. Detected Runways

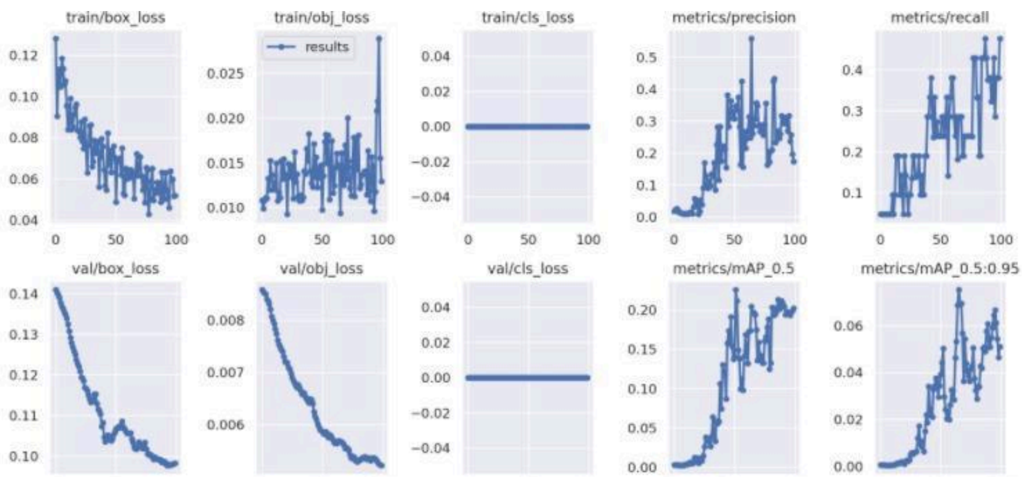



Fig 9. Training and Validation Loss

TABLE 1.EXPERIMENTAL RESULTS

| Model | Accuracy |
|---|---|
| Naive Bayes Classifier | 0.87 |
| Support Vector Machine | 0.91 |

## IV. CONCLUSION

This paper puts forward a new model which helps pilots predict the best possible routes from a point in the path to their destination in the absence of ATC, it also helps them trace out the runway when during bad weather. The model takes Wind Gust, Wind Speed, Humidity, Temperature, Pressure, fuel capacity as input and predicts the best routes possible. To trace the runway, the model uses the YOLOV5s algorithm and detects the runway in which the aircraft is supposed to land. The proposed methodology helps ATC management and flight pilot to give pre-tactical traffic information in the form of graph visualizations, which will give information about pattern cost for every possible path along with weather updates. The model is made to interpret the best pattern route considering commercial, operational, and financial capacity in pre tactic traffic plan. For future work, advanced solutions during take-offs, to ensure safe landings prediction of runway lanes using Kalman filter can be done

## V.REFERENCES


*[1] Guo, X., Denman, S., Fookes, C., Mejias, L., & Sridharan, S. (2014). Automatic UAV Forced Landing Site Detection Using Machine Learning. 2014 International Conference on Digital Image Computing: Techniques and Applications (DICTA).*

*[2] Richards, R. A. (n.d.). Application of multiple artificial intelligence techniques for an aircraft carrier landing decision support tool. 2002 IEEE World Congress on Computational Intelligence. 2002 IEEE International Conference on Fuzzy Systems. FUZZ-IEEE'02. Proceedings (Cat. No.02CH37291).*

*[3] G. Yang et al., "Face Mask Recognition System with YOLOV5 Based on Image Recognition," 2020 IEEE 6th International Conference on Computer and Communications (ICCC), 2020, pp. 1398-1404, doi: 10.1109/ICCC51575.2020.9345042.*

*[4] Sriratanawilai, S., & Erjongmanee, S. (2018). Route prediction in air travel network using socio-economic factors and learning models. 2018 5th International Conference on Business and Industrial Research (ICBIR). doi:10.1109/icbir.2018.8391174*

*[5] Lee, Der-Hau & Liu, Jinn-Liang. (2021). End-to-End Deep Learning of Lane Detection and Path Prediction for Real-Time Autonomous Driving.*

*[6] Hellström, T. and Ringdahl, O. (2006) 'Follow the Past: a path-tracking algorithm for autonomous vehicles', Int. J. Vehicle Autonomous Systems, Vol. 4, Nos. 2-4,pp.216–224*

*[7] Parsa, Amir Bahador & Movahedi, Ali & Taghipour, Homa & Derrible, Sybil & Mohammadian, Abolfazl. (2019). Toward safer highways, application of XGBoost and SHAP for real-time accident detection and feature analysis. Accident; analysis and prevention. 136. 105405. 10.1016/j.aap.2019.105405.*

*[8] Xie, Z., & Zhong, Z. W. (2016). Aircraft Path Planning under Adverse Weather Conditions. MATEC Web of Conferences, 77, 15001.*

*[9] Marcos, Rodrigo & García-Cantú, Oliva & Herranz, Ricardo. (2018). A Machine Learning Approach to Air Traffic Route Choice Modelling.*

*[10] Aytekin, Ö., Zongur, U., & Halici, U. (2013). Texture-Based Airport Runway Detection. IEEE Geoscience and Remote Sensing Letters, 10(3), 471–475.*

*[11] Moore, A. J., Schubert, M., Dolph, C., & Woodell, G. (2016). Machine Vision Identification of Airport Runways with Visible and Infrared Videos. Journal of Aerospace Information Systems, 13(7), 266–277.*

*[12] Liu, C., Cheng, I., & Basu, A. (2018). Real-Time Runway Detection for Infrared Aerial Image Using Synthetic Vision and an ROI Based Level Set Method. Remote Sensing, 10(10), 1544.*

*[13] J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.*

*[14] I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.*

*[15] K. Elissa, "Title of paper if known," unpublished.*

*[16] R. Nicole, "Title of paper with only first word capitalized," J. Name Stand. Abbrev., in press.*

*[17] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].*

*[18] M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.*