Journal of Advanced Zoology
ISSN: 0253-7214
Volume 45 Issue S-4 Year 2024 Page 186-196

Object Detection In Video Streaming Using Machine Learning And Cnn Techniques

Gouri Amol Vaidya[1*]

[1*]*Science Department, Indira College of Commerce and Science.*

***Corresponding Author:** Gouri Amol Vaidya*
**Science Department, Indira College of Commerce and Science.*

Abstract

Object detection is affecting a lot of areas where images play a crucial role. It therefore becomes necessary to address this issue by building effective image detection in video streaming systems, that can detect such traces of objects that may cause harm for flight landing and take-off, either by identifying the object type, detecting image size or evaluating under different environmental conditions . There are a number of techniques like Background subtraction, similarity matching, convolutional neural networks, end-to-end feed forward neural network to detect objects in video streaming but with limitations of correctly identifying object in live video steam under illumination changes, non-stationary backgrounds and similar looking background pixels and foreground pixels form a complex background. Hence it is now required to develop system that requires less training and no human intervention for object detection in live video streaming. For this study we have taken the application of runways where providing minute object detection is necessary for flight safety.

Keywords: Object Detection, Machine Leaning, CNN techniques.

I. INTRODUCTION

Object detection is a division of Computer Vision, in which objects that a human eye can observe in the images of video stream can be spotted, localized, and identified by computers. An image is a single frame that represents a static instance of a naturally occurring event. On the other hand, a video comprises of a series of instances of static images displayed in one second, giving the effect of observing a complete naturally occurring event.

1.1 What is Object recognition?

Object recognition involves object classification, detection and localization. Image classification is about labeling the identified object and localization is about drawing a rectangular box around the object. Image classification and localization together termed as image recognition. [1]

As such, we can distinguish between these three computer vision tasks:

1. Image Classification: Predict the type or class of an object in an image.

- Input: An image with a single object, such as a photograph.
- Output: A class label (e.g. one or more integers that are mapped to class labels).

2. Object Localization: Locate the presence of objects in an image and indicate their location with a bounding box.
- Input: An image with one or more objects, such as a photograph.
- Output: One or more bounding boxes (e.g. defined by a point, width, and height).
3. Object Detection: Locate the presence of objects with a bounding box and types or classes of the located objects in an image.
- Input: An image with one or more objects, such as a photograph.
- Output: One or more bounding boxes (e.g. defined by a point, width, and height), and a class label for each bounding box.

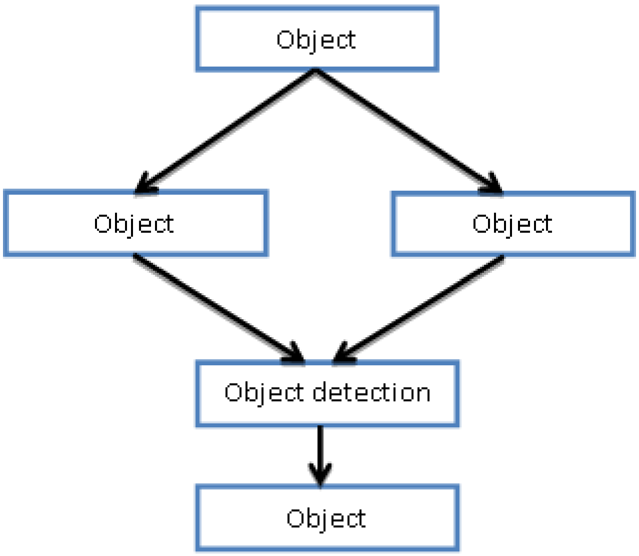

Figure 1: Object Recognition Computer Vision Tasks: an overview

1.2 Applications of object detection in video streaming

The applications of object detection in video streaming occur in various fields like sports, security and defense, traffic surveillance, runway safety, video communication in movie apps, robot vision, video games, automated vehicle systems, machine inspection, analyzing emotions and human behaviors, patient monitoring, train safety, etc. It is therefore necessary to address the issues of object detection, in the areas but not limited to given above, in various environmental conditions like different lighting conditions, various weather conditions, image blurring due to camera vibration, low visibility due to foggy conditions, non-stationary backgrounds, varying illumination, frames with camouflage etc.

Object detection can trace minute events to solve problems in various fields mentioned above like tracking a ball during a football match, tracking movement of a cricket bat, tracking a person in a video, detect moving objects, such as cars, ships, drones and missiles, to reduce illicit trading, drug trafficking, irregular border crossing, trafficking in human beings, smuggling, license plate recognition, vehicle number finding, traffic intensity determination, vehicle speed calculation, band violation, vehicle classification, sign language recognition, vehicle counting, pedestrian detection, comma patient movement tracking, identification of dent on railway track or a broken track itself, movement tracing on railway track, etc.

1.3 Scope of the document

This document introduces object detection and the areas that remain affected by it. Further, a survey of the history of object detection, the need for object detection in video streaming, state-of-the art techniques in object detection in video streaming and upcoming challenges in this area of object detection in video streaming are documented. The scope of this research is further mentioned which includes understanding and developing techniques using theories from machine learning and convolutional neural networks (CNN) in runways video image object detection by analyzing the background and foreground w.r.t various environmental factors. The objective of the study, the research questions to be addressed, the methodology to be adopted for research and the tentative schedule of the proposed work are further mentioned. In conclusion, the expected outcome and contribution of this research work are highlighted.

II. REVIEW OF LITERATURE’

2.1 Object detection tasks and algorithms used for its implementation

Table 1: Object Detection Algorithms

Task	Algorithm
for detection of moving object from a video sequence	1. Background subtraction (FCDH based [1] background subtraction mechanism) • optical flow- sensitive to noise and very inefficient for traffic applications. • temporal difference-approproate for when vehicles are motionless or exhibit limited mobility • background subtraction- best for detection of moving vehicles in video sequences [6]
Background modeling and foreground detection	similarity matching [2]
frame differencing	frame differencing algorithm[2]
detection of objects in images	convolutional neural networks, YOLO [3]
classification or detection on single images	deep-learning methods
static detection	end-to-end feed forward neural network
multi-frame analysis	recurrent Long Short-Term Memory (LSTM) network
Movement tracking	using optimized Kalman filter.

2.2 Techniques in object detection for video processing

- Region-Based Convolutional Neural Networks, or R-CNNs, are a family of techniques (R-CNN, Fast R-CNN and Faster R-CNN model) for addressing object localization and recognition tasks, designed for model performance. [1]
- **YOLO:** You Only Look Once, or YOLO: it is a second family of techniques for object recognition designed for speed and real-time use. YOLO-You only look once. Let's understand the objective with an example; when a car is driving itself means in self-driving car and if it wants to know if there is any other vehicle or pedestrian on the road. Thus it can be used to automatically detect the object (either moving or stationary) in real time. Another application is while playing Amazon prime video it shows the celebrity info on the vertical strip besides by identifying the actors on the screen.

The algorithmic output for the implementation of different tasks in object detection are found to be as below [2] -

- Image classification: Algorithms produce a list of object categories present in the image.
- Single-object localization: This term refers to localization of objects of one type within an image.

Figure 2-(a) Input Image

Figure 2-(b)

Fig. 2- The Faster RCNN object detector is easily able to detect a single object of type horse in the image.

Figure 3-(a) Input Image

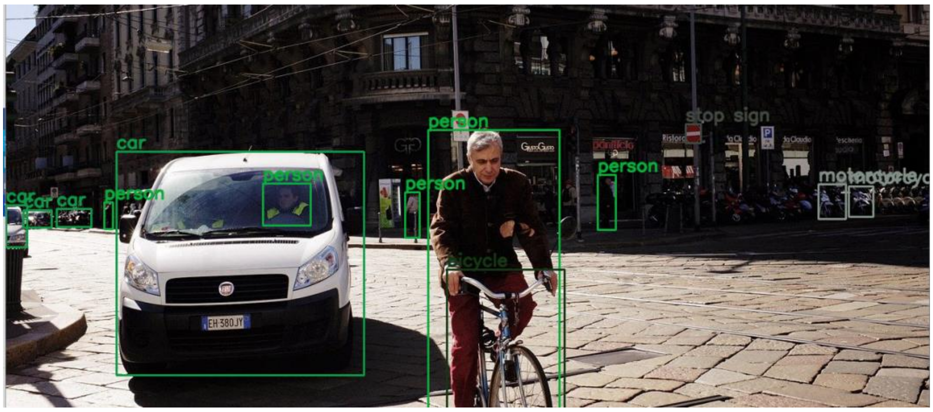

Figure 3-(b)

Figure 4. Faster RCNN network is able to detect many objects in this street image.

2.3 Literature Survey

Research papers are reviewed from the year 2001 till 2020 and analyzed w.r.t the parameters of Environmental conditions, Features extracted, Algorithms, Future scope and dataset used.

A general-purpose method [3] for segmentation of moving visual objects (MVOs) based on an object-level classification in MVUs, ghosts and shadows ghosts in video streams by exploiting color and motion information was proposed by R. Cucchiara et. Al.. Background subtraction technique is used for detecting moving objects in videos from stationary cameras as well as. The techniques are unconstrained to indoor and outdoor video scenes for environments conditions like light condition variations due to clouds in outdoor scenes or turning off lights in indoor scenes. Also another factor considered is that of the objects that modify their status from stopped to moving or vice versa. For object segmentation and evaluation, the motion, trajectory and speed are considered for vehicles, pedestrians, bicycles, and so on. Algorithm evaluation is done by applying MVO detection method for traffic monitoring of urban areas and highways, surveillance of parking zones, and indoor people detection and tracking. The proposed method is considered as general-purpose for applications that differ in density of objects in the scene, object size and number of frames of object presence. The aim of flight safety during aircraft landing approach to an airport is proposed by [4] Christian Stephana et.al. to perform automatic extraction of runway structures in infrared remote sensing image sequences. The methodology considers adverse weather conditions to provide solution by the combination of sensor vision and synthetic vision systems (Enhanced and Synthetic Vision System- ESVS). The proposed algorithm is used to detect runway structures and obstacles over a sequence of infrared images during aircraft landing. No additional database, INS or GPS is used at this processing stage but the algorithm generates several runway and obstacle hypotheses and the final decision in ESVS is taken in the further processing stage: fusion of radar and IR data hypotheses and synthetic vision data.
The work presented by S. Hantscher, H. Esse et. al. [5] facilitates and enhances the effectiveness of aircraft landing strip observation in an approach for FOD by multimodal netted 2D/3D sensors. It is a fully automatically controlled system that allows identification of debris while take-off and landing and avoids sources of error induced by human observers. The techniques consider all weather conditions. The paper focuses on the techniques used for the runway surveillance, especially the radar part for stand-off object detection of debris and a real-time 2D and 3D time-of-flight (ToF) camera system for close-up surveillance. The proposed system gives an alarm in the case of a dangerous item on the runway. The regular background subtraction method to detect object in a moving image frame is that of dynamic background and ghost effect as per Sujoy Madhab Roy and Ashish Ghosh suggested in a work to develop real-time record sensitive background classifier (RSBC) [6]. These problems are resolved in proposed method which makes use of combination of two algorithms, the Adaptive Background Model (ABM) and the Neighborhood Background Model (NBM).
Augmented Reality Surveillance System for Road Traffic proposed by Alexander Filonenko et. Al. aims at the road traffic monitoring system developed based on augmented reality. The region of interest is set as boundaries of the roads found by the lane detection algorithm based on vanishing point detection. To detect the moving cars on the road the very first step is building a background model. The scoreboard algorithm which is a tradeoff between accuracy and computational cost is used in this work. The mapping between camera and real world coordinates was introduced. The number of objects, their color, ways of getting real sizes of the objects, whether cars are present or not were the parameters used for classification of objects. Classification Algorithm used was the connected components labeling algorithm. Dataset used is that of Google Maps. The proposed algorithm does the job of detecting moving vehicles in low and high bit-rate video streams.

The future scope includes the job to
1. Show on the map not only cars but other participants of the city life such as pedestrians and bikes
2. Estimate velocity of Every detected moving object
3. Consider ground plane as non- planar surface
4. get boundaries of buildings, parks, and other objects (Using map services) and show to the user

Satish Kumar V et. Al. focus on detection of objects on runway from a safe distance so that the pilot can take decision about safe landing using electro-optical and infrared sensors. This paper presents techniques to detect runway and runway incursions using electro-optical color camera and medium wave infrared sensor on-board the aircraft during approach for landing. The performance of the proposed techniques are evaluated in flight simulators with simulated images of electro-optical and infrared sensors on-board the aircraft during approach for landing at a distance of 3 nautical miles from runway threshold. The environmental conditions considered are day/night and low visibility CAT 2 foggy conditions. Also effectiveness of the techniques with statistics of runway detection, miss detection and false alarm for different case studies have been provided and discussed.

Dhakate Pankaj in his paper entitled “Real-Time surveillance for critical activity detection in ICUs [9]” proposed a novel approach of application of object detection in the medical field. The goal of the work is to detect minute movement of coma patient. The methodology follows the human eyes visualization concept by using pair of identical two-dimensional cameras to generate stereoscopic video. The proposed method in this paper will activate a warning system, highlight the changes, and capture the live streaming video when minute movement of coma patient is detected, also it keeps track on mental stress of patients. Detection tracking of minute movement has been made possible using Kalman filter. Project implemented using this paper trim down memory requirements for activity storage. Future Scope is about minimizing the response time for tracking the object in real-time and expands application areas. Also simulation based stress analysis system can be designed to achieve better accuracy. To detect and measure effect of stress on coma patient two more factors can be considered – 1) Gender and 2) Age.

Shiva Kamkar and Reza Safabakhsh proposed about vehicle detection, counting and classification under different environmental conditions [10]. Vehicles are classified using random forest algorithm which is trained to classify vehicles into small (e.g. car), medium (e.g. van) and large (e.g. bus and truck) is done. Vehicles are selected using an active basis model and their verification is done according to their reflection symmetry. Features extracted are vehicle length and the correlation computed from the grey-level co-occurrence matrix of the vehicle image within its bounding box. Datasets used are Seven video streams which contain common highway challenges. Environment parameters considered are Different lighting conditions, various weather conditions, camera vibration and image blurring. As a part of Algorithm Evaluation, the efficiency of the proposed method is evaluated for use in traffic monitoring systems during the day (in the presence of shadows), night and all seasons of the year.

Noëlle M. Fischer et. Al. combine an end-to-end feedforward neural network for static detection with a recurrent Long Short-Term Memory (LSTM) network for multi-frame analysis to detect moving vehicles in video [11]. While presenting the practical aspects, the authors have selected the parameters like optimizer and batch size. The end-to-end network is able to localize and recognize the vehicles in video from traffic cameras. We show an efficient way to collect relevant in-domain data for training with minimal manual labor. Our results show that the combination with LSTM improves performance for the detection of moving vehicles.

Mark Smearcheck et.al. propose a work of sensor-modeling framework [12] to include airborne laser range scanners, forward looking infrared (FLIR), three dimensional imagers, and visible light cameras. A monitoring system equipped with an aircraft was developed to detect, classify, and track the objects in varying weather conditions. Physical properties of an object such as size, shape, thermal signature, reflectivity, and motion were considered when evaluating the sensor most suitable for detecting a particular object. The results were used to assess the threat level associated with the objects in terms of severity and risk using statistical methods based on the sensor's measurement and detection capabilities.

Prerna Dewan and Rakesh Kumar have proposed[13] approach of using color features of the object in live video streaming for object detection. The work described in paper is making use of Background subtraction, similarity matching and frame differencing. The proposed algorithm will work firmly in complex environments such as non-stationary backgrounds, varying illumination, frames with camouflage etc. As per the author, the work executed in the paper can further be extended to detect the shadow casted by any object to make its way in many more applications for advancement in the field of computer vision.

Denis S. Andreev and Nikolay V. Lysenko [14] state that air traffic organization fails at most of the times due to poor lightening condition. Limited viability do not allow to reliably determining the height of the aircraft above the runway. Additional information or visual reference pints need to be introduced for safe landing. Apart from poor lighting conditions, the author has also considered environmental parameters like quality of runway surface as well as presence of obstacle on the surface. Obstacles can be in the form of lost containers, aircraft covering parts, vehicles etc. Three popular classification methods were chosen for performance estimation in application to the tasks of object recognition on the runway in poor visibility conditions: Support Vector Machine (Accuracy: 52.2%), K-Nearest (Accuracy: 56.5%), and Decision Trees (Accuracy: 39.1%). As a solution towards poor performance by all the three methods, author suggested Adaptive Multiscale Retinex or Convolutional Neural Networks as another classification method.

Ali Şentaş et. Al [15] have developed a system that classifies vehicles according to their type using support vector machine and convolutional neural network. They have created their own video surveillance dataset. The vehicles are classified as per their colors. Also they are classified as Bus, Microbus, Minivan, SUV, Sedan and Truck. They have used Tiny-YOLO algorithm for real-time object detection. Support vector machine (SVM) for classification of objects. In the future, instead of classical approaches of centralized computation, a distributed scalable network of collaborating computation nodes is going to be developed to process streaming real-time video data coming from traffic surveillance cameras. In this way, hierarchical topic-based publish-subscribe messaging middleware is going to be realized. Also, real-time stream processing infrastructures such as Apache Kafka, Flink, and Storm will be considered.

Ali Şentaş et. Al. did Research on Airport Runway foreign object detection algorithm [15] based on texture segmentation. The proposed algorithms and Techniques used by the author are-

1. **Gabor filters to analyze the texture of airport runway images. Gabor filter:** used to filter the fabric image to enhance the contrast of the image while suppressing the background texture information
2. **bilateral filtering:** to de-noise the image further
3. **Threshold detection:** used to achieve the detection of foreign objects on the runway by segmenting and removing small connected domain.

Wei Liang et. Al. worked for airport runway safety [16]. Features extracted are about FOD (Foreign objects and debris). Dataset used is the image of the experimental runway. It was collected on the cement road surface similar to the airport runway pavement environment. Algorithm in this paper can better suppress the interference of airport runway background and noise, and accurately segment foreign objects from the runway background.

Denis S. Andreev in his work [17] to detect moving objects for runway safety where the author has providing the crew with the information about the runway condition and obstacles presence. The purpose is to objects segmentation using background subtraction for the enhanced vision system to be used on runway. Performance of the experiment was evaluated by three background subtraction methods: K-nearest method, Gaussian mixture method and Codebook method. It was found that the Codeword method is most robust to short-term variations in the background and provides a smaller number of false detections than the K-nearest and a mixture of Gaussian methods

Table 2: Literature Review

Sr.No.	Title	Year	Application Domain	Goal	Environmental conditions	Features extracted	Algorithms	Future scope	Dataset
1	Detecting objects, shadows and ghosts in video streams by exploiting color and motion information	2001	shopping center, US highway, Parking area, Laboratory	This work presents a general-purpose method for segmentation of moving visual objects (MVOs) based on an object-level classification in MVUs, ghosts and shadows.	light condition variations due to clouds in outdoor scenes or turning off lights in indoor scenes	Color and motion	background subtraction algorithm	Nil	No dataset
2	Automatic extraction of runway structures in infrared remote sensing image sequences	2005	Runway object detection	1. adverse weather conditions 2. Enhanced and Synthetic Vision System (ESVS) 3. an algorithm that is using only a sequence of infrared images to detect possible runway structures and obstacles on it during the aircraft landing approach.	adverse weather conditions	Features for edges, e.g. minimal area of the blob, length and elongation,	Developed using Microsoft Visual C++ 6.0 and Matrox Imaging Library.	Nil	The data from the sensor with the INS and GPS data are recorded and processed on a PC system.
3	LAOTSE, an approach for foreign object detection by multimodal netted 2D/3D sensors	2011	runway surveillance	1. searching for sources of danger imposed by debris 2. runway surveillance 3. landing strip observation effectiveness 4. techniques used for the runway surveillance 5. techniques used by radar part for stand-off object detection	normal, rainy and foggy weather conditions	dangerous (like scessors, screws) as well as non-dangerous (like piece of paper) items	change detection algorithm	Nil	radar data

Sr.No.	Title	Year	Application Domain	Goal	Environmental conditions	Features extracted	Algorithms	Future scope	Dataset
				6. stand-off object detection of debris 7. stand-off object detection of debris and a real-time 2D and 3D time-of-flight (ToF) camera system for close-up surveillance 8. giving an alarm in the case of a dangerousitem on the runway.					
4	Real-time record sensitive background classifier (RSBC)	2019		to identify moving objects from a complex video scene	challenging indoor and outdoor video scenes	Entire objects	Adaptive Background Model (ABM) and the Neighbourhood Background Model (NBM)	to add more features to handle camouflaged objects and PTZ camera motions.	CDnet 2014 dataset
5	Augmented reality surveillance system for road traffic monitoring	2014	traffic surveillance systems	The road traffic monitoring system developed based on augmented reality	detect moving vehicles in low and high bit-rate video streams	The number of objects, their color, ways of getting real sizes of the objects, whether cars are present or not	lane detection algorithm , The scoreboard algorithm, The connected components labeling algorithm	1. Show on the map not only cars but other participants of the city life such as pedestrians and bikes2. Estimate velocity of Every detected moving object3. Consider ground plane as non- planar surface4. get boundaries of buildings, parks, and other objects (Using map services) and show to the user5. them to the user as overlays for the video stream.	Google Maps

III. RESEARCH GAPS

1.to add more features to handle camouflaged objects
2.Show on the map not only cars but other participants of the city life such as pedestrians and bikes
3.Estimate velocity of every detected moving object
4.Consider ground plane as non- planar surface
5.Get boundaries of buildings, parks, and other objects (Using map services) and show to the user
6.to detect the shadow casted by any object to make its way in Runway applications for advancement

IV. OBJECTIVES OF THE STUDY

1. To study the various machine learning and CNN techniques used for video object detection.
2. To be able to use a combination of machine learning and CNN that can successfully detect object detection at given environmental conditions
3. To assess the efficiency of the proposed technique for runway safety.
4. To create avenues of research in the area of video object detection, machine learning and Convolutional neural networks.

V. CONCLUSION

The study is aimed to answer the questions like- What are the various machine learning and CNN techniques used for Object detection in video streaming? Develop machine learning CNN techniques to help provide better accuracy in object detection in a video streaming.
The scope of the research to address the above problem is restricted to the analysis phase of the video object detection process for runway video streaming. The focus of the research would be on object detection by means of machine learning techniques. For this process, image classification would be done by using either any one or a combination of machine learning and CNN techniques like nearest neighbor classifier, KNN classifier. Object localization algorithms like Yolo Alorithm would be used. In addition to this security measures to prevent counter forensics of this technique would also be tested. An outline of the methodology to achieve the same is mentioned in the later sections.

REFERENCES

1. R. Cucchiara, C. Grana, M. Piccardi, A. Prati,"Detecting objects, shadows and ghosts in video streams by exploiting color and motion information", Proc. of 11th International Conference on Image Analysis and Processing,pp.,2001
2. Christian Stephana, Gintautas Palubinskasb and Rupert Müllerb,"Automatic extraction of runway structures in infrared remote sensing image sequences", Image and Signal Processing for Remote Sensing XI, Bruges, Belgium, ,vol. 5982,pp. 19-22,2005
3. "S. Hantscher, H. Essen, P. Warok, R. Zimmermann, M. Schröder, R. Sommer, S. Lang, M. Schikora, K. Wild, W. Koch,"LAOTSE, an approach for foreign object detection by multimodal netted 2D/3D sensors",Proceedings of ESAV'11,pp. 12 - 14,2011"
4. Sujoy Madhab Roy, Ashish Ghosh,"Real-time record sensitive background classifier (RSBC)",Elsevier Expert Systems With Applications,vol. 119,pp. 104–117,2019
5. Alexander Filonenko, Andrey Vavilin, Taeho Kim, and Kang-Hyun Jo,"Augmented reality surveillance system for road traffic monitoring",Springer International Publishing Switzerland,pp. 310–317,2014
6. Satish Kumar V., Sudesh K. Kashyap, and N. Shantha Kumar,"Detection of runway and obstacles using electro-optical and infrared sensors before landing",Defence Science Journal,vol. 64,No. 1,pp. 67-76,2014
7. "Dhakate Pankaj,"Real-Time surveillance for critical activity detection in ICUs",Conference on Computer and Communication Technologies, Advances in Intelligent Systems and Computing,vol. 1,pp. 175-186,2016"
8. Shiva Kamkar, Reza Safabakhsh,"Vehicle detection, counting and classification in various conditions",IET Intelligent Transport Systems,vol. 10,Iss. 6,pp. 406–413,2016
9. Noëlle M. Fischer, Maarten C. Kruithof, Henri Bouma,"Optimizing a neural network for detection of moving vehicles in video",Conference: Counterterrorism, Crime Fighting, Forensics, and Surveillance,vol. 10441,pp.,2017

10. Mark Smearcheck, Ananth K. Vadlamani, Maarten Uijt de Haag,"Sensor classification and obstacle detection for aircraft external Hazard monitoring",Proc. of SPIE,vol. 6957,pp. 1-12,2018
11. Prerna Dewan, Rakesh Kumar,"Detection of object in motion using improvised background subtraction algorithm",International Conference on Trends in Electronics and Informatics,pp. 651-656,2018
12. Denis S. Andreev, Nikolay V. Lysenko,"Objects recognition methods estimation in application to runway pictures taken in poor visibility conditions",IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus),pp.,2019
    Ali Şentaş · İsabek Tashiev · Fatmanur Küçükayvaz · Seda Kul · Süleyman Eken · Ahmet Sayar · Yaşar Becerikli,"Performance evaluation of support vector machine and convolutional neural network algorithms in real-time vehicle type and color classification",Springer,pp.,2020
13. Wei Liang, Zhangli Zhou, Xiangyang Chen, Xueliang Sheng, XiaoDong Ye,"Research on Airport Runway FOD Detection Algorithm Based on Texture Segmentation",IEEE 4th Information Technology,Networking,Electronic and Automation Control Conference (ITNEC 2020),pp. 2103-2106,2020
14. Denis S. Andreev,"Moving Objects Segmentation Method for Flight Vision Systems",IEEE Xplore,pp. 1356-1359,2020