Integrated Science

Journal of Integrated
SCIENCE & TECHNOLOGY

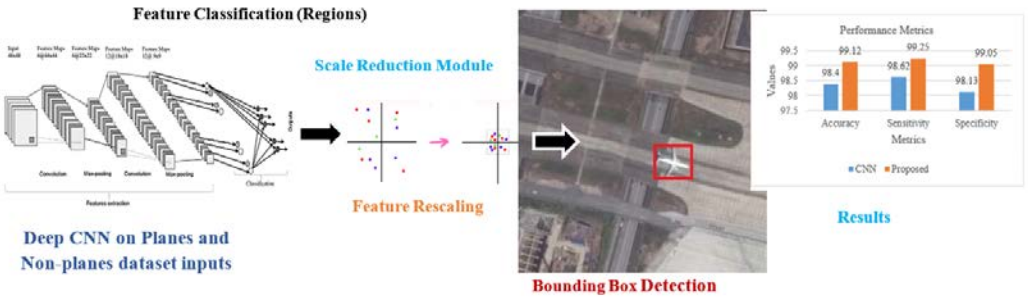
RRBCNN: Aircraft detection and classification using Bounding Box Regressor based on Scale Reduction module

Bhavani Sankar Panda[1*], Kakita Murali Gopal[1], Rabinarayan Satpathy[2]

[1]*GIET University, Gunupur, Odisha, India.* [2]*SRI SRI University, Bhubaneswar, Odisha. India.*

Received on: 27-Aug-2022, Accepted and Published on: 11-Nov-2022

ABSTRACT



Research on aircraft detection has recently resulted in significant military and civil applications. Concurrently, research relying on DL (Deep Learning) and image processing has gained wide interest in object tracking. However, variations in the kind of aircraft, complex background, and pose have made it complex to detect aircraft efficiently. Thus, an effective algorithm is needed to solve this difficulty. For this purpose, the study considers the plane-sent dataset. It proposes RRBCNN (Region Regression Based Convolutional Neural Network) for classifying stationary aircraft into "planes" and "not planes" with high accuracy. The study also introduces SRM (Scale Reduction Module), which RRBCNN utilizes for minimizing the feature-map scaling. It also eliminates information loss to enhance the training rate. This assists in improving aircraft detection by affording suitable bounding boxes. The performance of this system is analyzed in terms of AC (Accuracy), FPR (False Positive Rate), MR (Missing Rate), ER (Error Ratio), precision, average processing time (T), recall, specificity, sensitivity, and FN (False Negative) rate. Three conventional types of research are considered for comparative analysis to evaluate the efficiency of the current work. Finally, the efficiency of the introduced methodology is proved through analysis.

Keywords: Stationary Aircraft Detection, Image Processing, Deep Learning, Region Regression based CNN, Scale Reduction Module

INTRODUCTION

With the technological development corresponding to remote sensing and the improvement of image resolution, the automatic detection of aircraft in high-resolution images is significant in military applications. It has also become a hotspot in the aviation field. Detecting aircraft is an important area in analyzing remote sensing images. Specifically, identifying stationary aircraft might be of deliberated importance as it affords a successful and comprehensive process for decision-making to the rapidly progressing military operations. Moreover, satellite images highly rely on the research corresponding to the geometric structures of aircraft while detecting planes by image processing methods. Evident aircraft configuration plays a significant part in detecting satellite pictures.[1] However, various disturbing artifacts can exist around the aircraft. Therefore, it is important to carry out pre-processing measures to eliminate specific noise and explain aircraft characteristics for efficient detection. Several aircraft recognition and identification experimentations are explored by different researchers.[2]

Accordingly, an efficient method for detecting airplanes in remote sensing images relying on the non-maximal and multi-layer feature fusion algorithm has been recommended. Following the general lower leveled remote sensing and natural image features, R-CNN (Region-based Convolutional Neural Networks) have been selected for transfer learning through limited data.[3] Subsequently, L2-kind normalization, dimensionality reduction, feature connection, and scaling have been suggested for efficient low-level and high-level feature fusion.[4] Lastly, a non-maximal suppression technique relying on the soft decision operation has been endorsed to solve overlap issues of the detection boxes. Experimentations explored that the recommended system could efficiently enhance

*Corresponding Author: Bhavani Sankar Panda
GIET University, Gunupur, Odisha, India
Email: bs.panda@giet.edu

Cite as: *J. Integr. Sci. Technol., 2023, 11(1), 412.*
URN:NBN:sciencein.jist.2023.v11.412

 ISSN: 2321-4635
http://pubs.thesciencein.org/jist

the representation capacity of small and weak objects and accurate and quick detection of airplane objects.[5]

Similarly, meta-analysis has been conducted by considering the outcomes of conventional research to document the development and employment of supervised land cover image classification. This process encompasses several factors like sensors, targeted classes, geographical areas, segmentation algorithms, uncertain variables, land cover kinds, supervised classifiers, size of training sample, and accuracy analysis techniques. Scientific progress in supervised classification has explored high remote sensing image resolution like UAV and has been profitable in attaining better accuracy.[6] However, there exist various exceptional cases. For instance, Pleiades images have been mainly employed in urban regions, leading to anomalous minimum accuracy for classification. Thus, employing widely employable remote sensing images or the land cover kinds is vital for further verification of object-based image classification methodology.[7,8]

In addition, an agile CNN framework termed Sat CNN has been recommended for remote sensing images having a high resolution for classification. Following recent enhancements to CNN architecture, convolutional layers with minimum kernels have been utilized to construct efficient CNN architecture.[9,10] Experimentations on the SAT datasets proved that it could efficiently and quickly learn features for handling intra-class diversity with minimum convolutional kernels. Moreover, deeper convolutional layers permit spatial relationship spontaneous modeling.[11] Likewise, aircraft recognition technique relying on SVM[12] and two-layered SAM (Saliency Analysis Model)[13] has been introduced for the high resolution and wide area remote sensing images. Initially, FLS (First Layer Saliency) has been incorporated with spatial frequency and VSA (Visual Saliency Analysis)[14] method relying on color space for minimizing the background interference. Subsequently, feature descriptors have been used that rely on Hu and SIFT moment for accurately describing the aircraft features. Lastly, the features of these objects have been fed into SVM, and aircraft have been recognized. Empirical outcomes represent the reliability and efficacy of the suggested system.[15] To enhance the system, FCNN (Fully CNN) has been suggested. Outcomes explored minimum memory requirement and test time than a traditional framework. The precision rate has also been satisfactory, and the endorsed approach has been found valuable in effectively detecting objects in remote sensing images. In the future, the detection ability of the network has to be enhanced.[16-18]

Though conventional works attempted to perform efficient aircraft detection, various challenges still existed due to color, the uncertainty of a few classes, density, and shape variations.[19,20] Hence, for efficient detection, the present study proposes RRBCNN (Region Regression-based CNN) that works based on a topmost executing detection model named faster R-CNN. This is pre-trained and fine-tuned with the chosen dataset. Furthermore, a data invariant SRM (Scale Reduction Module) is also introduced to resize the hierarchical feature map. This also enhances the training efficiency of adaptation that minimizes the feature maps' scale. Hierarchical features obtained from these feature maps explore the positioning data of the object in the overall image. In addition, extracted features in FC (Fully Connected) layers can efficiently characterize preferred information. Thus, this study mainly aims to frame RRBCNN that is fine-tuned for accurately detecting stationary aircraft.

The major contributions of this study are listed below.

- To detect the stationary aircraft with maximum accuracy using the proposed fine-tuned RRBCNN (Region Regression Based Convolutional Neural Network).
- To introduce SRM (Scale Reduction Module) to minimize the feature-map scaling with no information loss for enhancing the performance of RRBCNN to yield higher accuracy in detecting stationary aircraft.
- To analyze the performance of the proposed system through comparative analysis concerning various performance metrics for evaluating its efficiency.

REVIEW OF EXISTING WORK

Traditional systems have utilized various methods to detect aircraft. Accordingly, a comparative analysis has been undertaken considering the conventional CNN (Convolutional Neural Network) based model for object detection, YOLO-v3 (You Look Only Once-v3),[21] Faster R-CNN (Faster Region-based CNN)[22] and SSD (Single Shot Multi-box Detector)[23] to handle limited labeled data for automatic detection of airplanes in the satellite images. In addition, data augmentation processes like cropping, rescaling, and rotation were employed on test images for the artificial enhancement of the training data. Besides, NMS (Non-Maximum Suppression)[24] has been endorsed at the end of YOLO-v3 and SSD to discard various detection incidences near individual objects detected in overlapping regions. Trained networks have been employed for VHR (Very High Resolution) test images encompass airports and their surroundings for assessing their performance. Outcomes corresponding to the analysis of test region accuracy confirmed the efficiency of Faster R-CNN in terms of precision, accuracy, and F1-score. Following this, YOLO-v3 showed slightly minimum performance by affording a trade-off between speed and accuracy. Lastly, SSD explored minimum performance in detection. However, it was effective in localizing the object. Results have also been assessed concerning detection accuracy and object size, which confirmed that medium and large-size airplanes had been detected with maximum accuracy.[25] Conventional works have used different techniques. Correspondingly, the edge box algorithm has been used to detect an object that supports edge information for object detection and has also been robust to variation in the object size. CNN has been suggested for automatically classifying images that efficiently learn the optimal features from huge data. Besides, CNN has been uniform to minor shifts and rotations in the target object. Motivating empirical outcomes have been attained on a large dataset. Better recall and precision rate of the system explored its efficiency.[26] In addition, a scheme to detect aircraft relying on CNN and corner clustering has been suggested. It has been partitioned into two major parts: classification and regional proposal. Initially, candidate regions get generated through mean shift clustering algorithms. Subsequently, CNN has been utilized to extract features, and candidate regions that probably comprise aircraft have been classified. In comparison to traditional

methodologies like SS (Selective Search) + CNN, edge box +CNN and HOG (Histogram of Oriented Gradient) + SVM (Support Vector Machine), recommended approach possesses maximum efficiency and accuracy as it could automatically learn relevant features from large data and afford better outcomes.[27,28]

Likewise, an aircraft detection method in the remote sensing images has been endorsed that relies on deep ResNet (Residual Network) and SV coding. Initially, a variety of ResNet having minimum layers have been outlined for enhancing the feature map's resolution, and the multi-level convolutional features have been incorporated into the information feature depiction for the region proposal. Concurrently, HOG has been extracted with SV coding that supports CNN to perform effective feature extraction for efficient classification. Analysis has been undertaken on the remote sensing dataset. Empirical results explored satisfactory outcomes even on complex backgrounds.[29] Similarly, an approach has been suggested for vision-based long-range aircraft detection. Deep CNN has been trained to learn visual features of aircraft by mid-air flight data and head-on collision encounters amongst two-winged aircraft. Then, a method has been recommended that integrated the learned features with the handcrafted features utilized by conventional study. Lastly, performance analysis has been undertaken on real flight data from UAV (Unmanned Aerial Vehicle), which is enhanced better than traditional systems with no extra false alarms [30]. In addition, a framework based on CNN and reinforcement learning has been endorsed. Aircraft in the remote sensing images could be robustly and accurately positioned with the assistance of a search method that the candidate region has dynamically minimized to accurate aircraft location. The detection framework solves the complexities faced by traditional reinforcement learning techniques to detect specified objects. To enhance the system further, restricted edge boxes have been adopted that could produce candidate boxes of high quality by prior knowledge of the aircraft. Subsequently, smart detection agent has been trained by apprenticeship and reinforcement learning. The detection agent correctly positions aircraft within various actions. This also performs better in comparison to the greed approach. Finally, the CNN model that detects the probability produced in detecting aircraft has been recommended. Experimentations represented the efficiency and accuracy of the endorsed outline.[9] An efficient framework for aircraft detection relying on CNN has been employed to detect multiple targets in complicated scenes. CEdge Boxes (Constrained EdgeBoxes) have been designed to produce many target candidates precisely and quickly. Then, to solve the demerits of utilizing traditional classifiers and handcrafted features, an altered GoogleNet integrated with F-RCNN has been explored for extracting valuable features to perform effective detection. Furthermore, an ensemble and multi-model technique have been suggested to minimize FAR (False Alarm Rate) induced by unbalanced complex background and target distribution. Extensive experimentations undertaken on the dataset attained by QuickBird exhibited the superiority of the suggested system to detect aircraft better.[31]

An efficient framework based on landmark detection has been suggested to improvise system performance. It possesses a robust ability and requires only minimum labeled data. Alteration of the vanilla network has been applied for landmark detection concurrently by regression that encodes geometric restrictions. Loss function has also been recommended to create a fair network for varied aircraft. Integrated with careful design of post-processing and pre-processing methods, accurate landmark positions could be attained. Following this, the template matching method has been used for target recognition. The recommended framework could handle various positions, backgrounds, and aircraft poses efficiently and effectively, confirmed by experimentations.[32]

Moreover, DON (Deep patch Orientation Network) has been suggested to track the multi-ground target. This is general, and it could learn the target's orientation by relying on structural information in training samples. These methods influence the performance of detection outline into two major aspects: the enhancement of target detection through the patch-based model to localize the target in the detecting element. It also improves each track's motion characteristics by including orienting information as integrating into the tracking element. Following the DON structure, YOLO and FrRCNN (Faster Region CNN0 with SORT (Simple Online and Real-time Tracking) have been used as the case study. Analysis revealed that the overall prediction rate has enhanced.

Moreover, the IDsw (Identity switches) have minimized by about 67% by not impacting the tracking component's computational complexity. Thus, the endorsed methodology has effectively tracked the ground targets.[33] A specific algorithm for airplane detection has been suggested, namely R-FCN (Region-based Fully Convolutional Networks) and KF (Kalman Filter). This framework has been utilized to afford a plane's location message to track a model. To minimize detection time, a certain area has been cropped in an individual frame relying on the position of the bounding box of the earlier frame. The detection region's scale has been altered based on the target size. R-FCN has been considered an observation model integrated with KF to modify trace prediction. When the identification of subsequent frames varies to enhance the detection rate, the bounding box arriving at the forthcoming frame has been adjusted after and before outcomes. This methodology has been confirmed efficient in experimentation.[3,34]

Significant issues identified through the analysis of conventional research are listed below.

- The definitive study suggested CNN for aircraft detection. However, this CNN performs detection only if patch input has been the target aircraft. However, it has not altered its respective positions found by the clustering center. Therefore, border regression must be analyzed to adjust the border following the aircraft size and enhance the accuracy.[27]
- Missed detections have to be reduced by using varied ground target description levels by deep layers of the network corresponding to detector elements to enhance aircraft tracking performance.[33]

With persistent exploration and improvement of DL (Deep Learning), few frameworks have evolved. Hence, the subsequent stage must improve conventional network techniques. In addition, training time and model parameters increase with the enhancement of network layers. Therefore, how network structure can be optimized for balancing the contradiction between efficiency and

performance has been the key challenge that must be considered in the future.[5]

METHODOLOGY

The research aims to introduce a learning classification methodology that detects stationary ground aircraft. Though conventional works attempted to detect aircraft, they lacked accuracy. To solve this, the present work proposes RRBCNN (Region Regression Based Convolutional Neural Network) and SRM (Scale Reduction Module); the overall block diagram of the proposed work is given in figure 1. RRBCNN takes bounding boxes throughout the target objects with their class labels. This encompasses the region-based network; it acts as an attention network to detect the region. Image is taken as the input. Then, it is processed based on RCNN to classify the regions efficiently by saving time by sharing convolutional layers. The model is pre-trained with the chosen dataset. Subsequently, it is fine-tuned with respective training data. Random splitting is accomplished with validation and training to tune the learning rate and perform effective prediction. SRM is also designed for downscaling feature maps, which enhances the training rate.

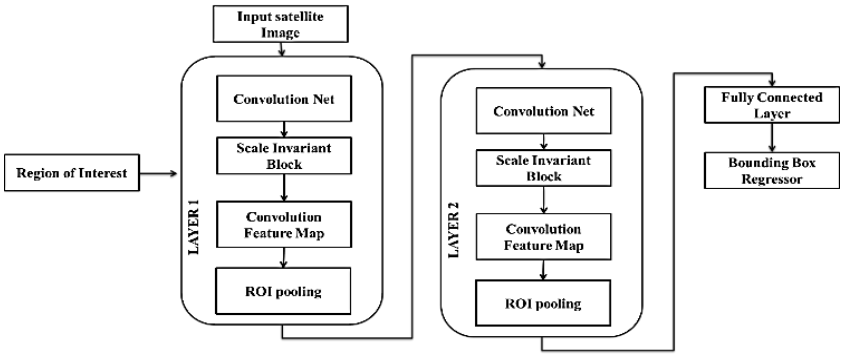


Figure 1. Overall view of the proposed RRBCNN with SRM

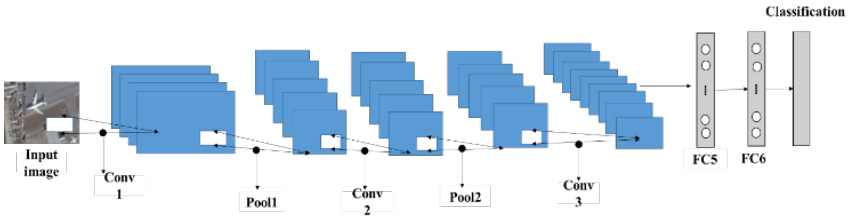


Figure 2. CNN architecture[35]

CNN (Convolutional Neural Network)

CNN is the class of DL-NN (Deep Learning-Neural Networks).[36] It is usually an efficient algorithm, especially in the image-processing domain. Currently, these algorithms are best for automated image processing. These also possess specialized applications in video and image recognition. It is also mainly used in image analysis, object detection, segmentation, and image recognition. Three kinds of layers in CNNs exist, as listed below,

Convolutional layer: Each input neuron in a general NN is associated with a subsequent hidden layer. Moreover, only a small region of neurons of the input layer connects with the hidden layer neuron.

Pooling layer: it is utilized for dimensionality reduction of the feature map. Multiple pooling and activation layers exist within CNN's hidden layer. Generally, this layer has been called a sub-sampling layer. It transfers the feature maps (input) to their corresponding output with minimum sample numbers.

FC (Fully Connected) layer: these layers exist from the last few network layers. Input to the FC layer is output from the flattened convolutional or pooling layer. Then, this is fed into the FC layer. Features from numerous pooling and convolution layers get processed with traditional CNN structures. In this context, ANN (Artificial Neural Network) is linked to individual weights of earlier layers. A schematic depiction of CNN layers in their overall structure is given in figure 2, where the input image that gets fed into various CNN layers performs learning to perform classification.

The overall algorithm of typical CNN is given in algorithm-I. Next, independent variable sequences are taken as input, and their standard deviation are divided. Then, the map size is computed by step 2. Following this, CNN has been set with the input layers and its sub-sampling layers as per step 3. Subsequently, training samples are executed and given in step 4. Finally, test samples have been executed for testing.

The algorithm I: CNN

Input: Independent variable sequence $a_i^j = (a_1^j + a_2^j + \cdots + a_i^j)$

Output: CNN construct $l_{i+s}^j = \left(l_{i+1}^j + l_{i+2}^j + \cdots + l_{i+s}^j\right)$

Step-1: Initially, read the input data and find its standard deviation. Finally, divide the data by the computed standard deviation.

Step 2: Compute map size.

$$\text{Mapsize} = \text{cons}\big(\log 3(\text{data\_size})\big)$$

-1 is the overall variables number.

Then, its loss function is given by,

$$\text{Loss: } A^b = \frac{1}{2}\sum_{i=1}^{r}\left(u_i^b - v_i^b\right)^2 = \frac{1}{2}\left\|u^b - v^b\right\|_2^2$$

Step 3: Set CNN with sub-sampling and input layers,

$$\delta^D = (X^{D+1})^R \delta^{D+1} f'(g^D)$$

Step-4: Execute training samples given by,

$$\text{Ncc} = \text{cnntrain}(\text{cnn}, \text{Ra}, \text{Rv}, \text{choose});$$

$$\begin{cases} \dfrac{\partial A}{\partial X^D} = a^{D-1}(\delta^D)^R \\ \Delta X^D = \eta \dfrac{\partial A}{\partial X^D} \end{cases}$$

Step-5: Execute test samples, initiate the test,

$$Z_t^D = f\left(\sum_{k \epsilon J_t} a_k^{D-1} i_{jt}^D + d_t^D\right)$$

RRBCNN (Region Regression-Based Convolutional Neural Network)

Initially, RRBCNN extracts numerous regional proposals from the input image, class labeling, and bounding boxes. Subsequently, CNN is utilized for performing forward propagation on individual region proposals for feature extraction. Then, features of the regional proposal are utilized for class prediction, thereby exploring the bounding box. This proposed algorithm possesses merit over a generation of the random proposal as it restricts the proposals with high recall. The architecture of RRBCNN and its internal processes are shown in figure.3, where the source image and target domain are fed into CNN blocks. The introduced algorithm employs feature alignment components on proposal and domain features. For the hierarchical alignment module of the domain feature, various sub-modules of the adversarial classifier are executed on block 3, block four and block 5. To minimize the size of feature maps, SRM is used. This is then fed into RPN (Regional Proposal Network), which comprises classification and regression prediction. Then, it is given to ROI pooling and then to FC layers. Finally, classification scores and the bounding box outcomes with respective features for domain classifiers are attained.

Thus, RRBCNN performs several steps for predicting aircraft by bounding boxes. Accordingly, this proposed algorithm is intended to compute localization confidence with the location that will assist NMS (Non-Maximum Suppression). This network identifies Gaussian distribution and bounding box positioning and is given by equation 1.

$$Q_\Theta(a) = 1/2\pi\sigma^2 e^{-\frac{(a-a_y)^2}{2\sigma^2}} \quad (1)$$

Where a_y represents the estimation of bounding box positioning, standard deviation (σ) computes uncertainty in estimation afforded by FC layers on fast RCNN and the final layer. In addition, the loss function is utilized rather than ReLU for avoiding $\sigma = 0$. When $\sigma = 0$, the network has extreme confidence regarding computed bounding box positioning. This can also be expressed as Gaussian distribution, a Dirac-delta function given by equation 2.

$$Q_E(a) = \delta(a - a_i) \quad (2)$$

Where a_g indicates ground-truth bounding box positioning.

The aim of localizing objects in this condition is to compute Θ, and is given by equation.3.

$$\Theta = \arg\min E_{KL}(Q_E(a)||P_\Theta(a)) \quad (3)$$

Loss function (A_r) is used for the bounding box regression and is computed by equation. Four and equation 5.

$$A_r = E_{KL}(Q_E(a)||Q_\Theta(a)) \quad (4)$$

$$\int Q_E(a)\log\frac{Q_E(a)}{Q_\Theta(a)}dx$$

$$= -\int Q_E(a)\log Q_\Theta(a)dx + \int Q_E(a)\log Q_E(a)dx$$

$$= -\int Q_E(a)\log Q_\Theta(a)dx + \int B(Q_E(a))$$

$$= -\log Q_\Theta(a_i) + B(Q_E(a))$$

$$= \frac{(a_i-a_y)^2}{2\sigma^2} + \frac{1}{2}\log(\sigma^2) + \frac{1}{2}\log(2\pi) + B(Q_E(a)) \quad (5)$$

When location (a_y) is inaccurately estimated, the variance (σ^2) is expected to be small so that loss will also be minimum. $\frac{1}{2}\log(2\pi) + B(Q_E(a))$ does not disturb decision-making as it does not rely on the estimated parameter (Θ). Hence, it is given by equation.6.

$$E_{KL}(Q_E(a)||P_\Theta(a)) \propto \frac{(a_i-a_y)^2}{2\sigma^2} + \frac{1}{2}\log(\sigma^2) \quad (6)$$

When $\sigma = 1$, loss reduces to standard Euclidean loss and is given by equation.7.

$$E_{KL}(Q_E(a)||P_\Theta(a)) \propto \frac{(a_i-a_y)^2}{2} \quad (7)$$

Loss varies concerning location confidence and location estimation and is given by equation.8 and equation.9.

$$\frac{d}{d\sigma}E_{KL}(Q_E(a)||P_\Theta(a)) = \frac{(a_y-a_i)^2}{\sigma^{-3}} - \frac{1}{\sigma} \quad (8)$$

$$\frac{d}{da_y}E_{KL}(Q_E(a)||P_\Theta(a)) = \frac{(a_y-a_i)^2}{\sigma^2} \quad (9)$$

Conversely, as σ exists in denominators, gradient occasionally could explore at the initiation of the training. To eliminate this issue, the network detects $\alpha = 1/\sigma 2$ rather than σ. Following the batch normalization, a small constant $\epsilon = 0.0001$ is integrated into the long-term to prevent it from becoming negative infinite and is given by equation 10.

$$E_{KL}(Q_E(a)||P_\Theta(a)) \propto \frac{\alpha}{2}(a_i - a_y)^2 - \frac{1}{2}\log(\alpha + \epsilon)$$

$$= \frac{d}{da_y}E_{KL}(Q_E(a)||P_\Theta(a)) = \alpha(a_y - a_i)$$

$$= \frac{d}{d\alpha}E_{KL}(Q_E(a)||P_\Theta(a)) = \frac{(a_y-a_i)^2}{2} - \frac{1}{2(\alpha+\epsilon)} \quad (10)$$

For $|a_i - a_y| > 1$, loss identical to smooth A_1loss stated in fast R-CNN and is given by equation.11.

$$A_r = \alpha\left(|a_i - a_y| - \frac{1}{2}\right) - \frac{1}{2}\log(\alpha+\epsilon) \quad (11)$$

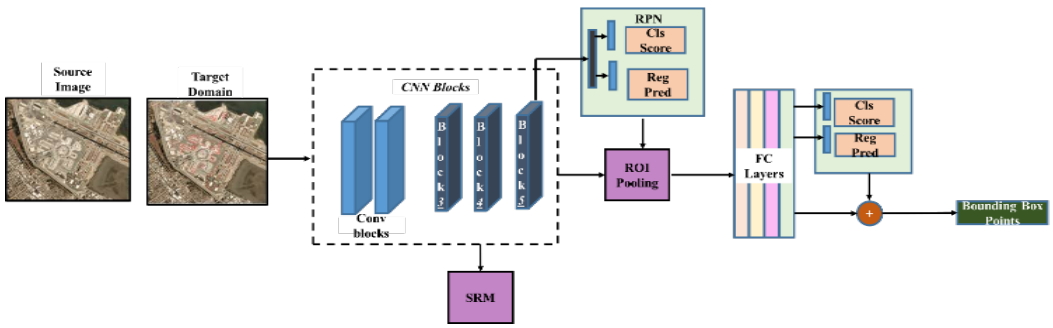


Figure 3. Architecture of RRBCNN and its internal process

To use loss, parameterizations are adopted corresponding to $a1, b1, a2, and\ b2$ coordinates rather than $a, b, m, and\ d$ coordinates utilized by R-CNN and are given by equation 12.

$$\left.\begin{aligned} u_{a1} &= \frac{a1-a1_c}{m_c}, u_{a2} = \frac{a2-a2_c}{m_c} \\ u_{b1} &= \frac{b1-b1_c}{d_c}, u_{b2} = \frac{b2-b2_c}{m_c} \\ u^*_{a1} &= \frac{a1^*-a1_c}{m_c}, u^*_{a2} = \frac{a2^*-a2_c}{m_c} \\ u^*_{b1} &= \frac{b1^*-b1_c}{d_c}, u^*_{b2} = \frac{b2^*-b2_c}{m_c} \end{aligned}\right\} \quad (12)$$

Where $u_{a1}, u_{b1}, u_{a2}, u_{b2}$ indicate predicted locations (that is, a_y in equation.1), $u^*_{a1}, u^*_{b1}, u^*_{a2}, u^*_{b2}$ represent ground-truth locations (that is, a_i in equation.2). Moreover, $a1_c, a2_c, b1_c, b2_c, mc, d_c$ exist from anchor box, a_1, b_1, a_2, b_2 from the corresponding predicted box. Finally, it is noted that σ (standard deviation) is detected after bounding box normalization.

SRM (Scale Reduction Module)

For efficient training of hierarchical domain-feature alignment, SRM is used by RRBCNN, which intends to downscale feature maps with no information loss.

This module performs two main steps, as discussed below.

- 1x1 convolutional layer is executed to minimize the feature map's channel numbers in individual blocks.
- This step attains the features of the domain information to reduce the feature dimensions for effective training.
- Calibrating the features through scale reduction during the enhancement of the feature map's channel minimizes the size of the training set and improves feature dimensionality.

Furthermore, $S * S$ adjacent pixels from the feature maps are collected to generate new pixels. This step assists in easier computation with minimum parameters, thereby improvise training efficiency. The subsequent step is framed as per equation.13.

$$E^D_{(x,y,z)} = E^A_{x\times b+z\%b^2\%b, y\times b+\left[z\%\frac{b^2}{b}\right], [z\%b^2]} \quad (13)$$

Where E^A represents feature maps before the second element, moreover, (x, y, z) indicates the element on the c^{th} feature map positioned at(x, y), count from0. The E^D indicates scale-minimized feature maps, and b represents the sampling factor that denotes the feature map's adjacent (bxb) pixels. This is merged into a single feature. SRM has parameters only in the initial element, and the parameters get minimized while efficiency in training gets enhanced.

RESULTS AND DISCUSSION

The results obtained from the execution of the introduced system are discussed in this section with dataset description, performance metrics, experimental results, and comparative analysis.

Dataset Description

The study considered the planes-net dataset taken from https://www.kaggle.com/rhammell/planesnet This dataset intends to address the complex task associated with aircraft detection in satellite images. Automating this process could solve several challenges, like defense intelligence and airport monitoring for traffic and activity patterns. The dataset is in JSON format. Loaded objects comprise data, labels, scene ids, and location lists.

Table 1. Dataset image description with a class label

Class Label	Number of Images	Image Size / Type
Plane	8,000	20×20 / RGB
Not Plane	24,000	

Each image filename has a format: (label) (scene_id) (longitude_latitude).png

- **Label:** Value 0 or 1 indicating "no-plane" or "plane" class. In addition, the class label, number of images, and type or size of the image are given in table-1.
- **scene_id:** image-chip was retrieved from the unique ID of the planet scope visual scene. This could also be utilized for discovering and downloading the overall scene.
- **longitude_latitude:** Latitude and longitude coordinates of the image's center point with values partitioned by an underscore.

Sample images of the dataset

The considered dataset encompasses images with class labels as "planes" or "not planes," as shown in figure 4 and figure 5. In addition, a specific image is shown in figure 6 to provide a clear view of images in the dataset.

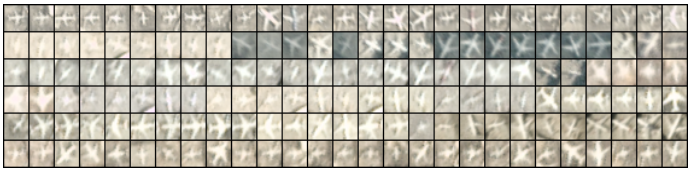


Figure 4. Planes

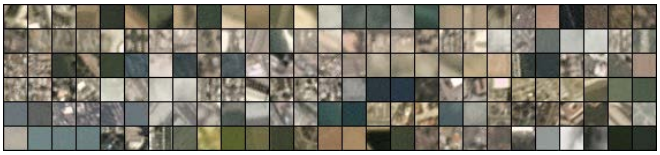


Figure 5. Not Planes

Figure 6. Image from planes, not dataset

Performance Metrics

The metrics considered for analyzing the performance of the proposed system are discussed in this section.

AC (Accuracy)

It is defined as the computation of overall correct classification and is given by equation.13.

$$\text{AC} = \frac{\text{number of aircraft\_detected}}{\text{number of aircraft}} * 100\% \quad (13)$$

FPR (False Positive Rate)

It is described as a proportion of incorrect and negative cases identified as positive in respective data. It is given by equation.14.

$$\text{FPR} = \frac{\text{number of false aircraft\_detection}}{\text{number of aircraft\_detected}} * 100\% \quad (14)$$

MR (Missing Rate)

MR is defined as the ratio of uncertainty resulting from total uncertainty or missing data and is given by equation.15.

$$\text{MR} = \frac{\text{number of aircraft\_missing}}{\text{number of aircraft}} * 100\% \quad (15)$$

ER (Error Ratio)

The proportion at which error occurs is termed ER and is given by equation.16.

$$\text{ER} = \text{MR} + \text{FPR} \quad (16)$$

Average processing time (T)

It is stated as the computation that permits the realization of the time taken for processing stepwise procedures for aircraft detection.

Precision and Recall

It is the proportion of accurate detection outcomes, whereas recall indicates the proportion of the true objects detected and is given by equation.17 and equation.18.

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive+False Positive}} \quad (17)$$

$$\text{Recall} = \frac{\text{relevant image} \cap \text{retrieved image}}{\text{relevant image}} \quad (18)$$

Sensitivity

It indicates positive segments which are suitable detected and is given by equation.19.

$$\text{Sensitivity} = \frac{\text{True Positive}}{\text{True Positive+False Negative}} \quad (19)$$

Specificity

It is stated as the quality or state of being unique and specific to groups or individuals and is given by equation.20.

$$\text{Specificity} = \frac{\text{True Negative}}{\text{True Negative+False Positive}} \quad (20)$$

FN (False Negative) rate

FN or FN error is the test outcome that wrongly denotes that a specific state does not support it.

EXPERIMENTAL RESULTS

The proposed system is experimentally implemented, and the obtained results are given in figure 7. The actual image and the detected image surrounded by a bounding box are presented, which explores the efficiency of an introduced system in detecting aircraft.

Figure 7. Experimental results

Comparative analysis

Comparative analysis is undertaken in terms of important performance metrics for evaluating the effectiveness of the introduced system. Obtained results are discussed in this section. Accordingly, DBN (Deep Belief Network), BING+CNN (Binarized Normal Gradients) +CNN, Faster R-CNN (Faster Region-based CNN), YOLOv2 and Multi-layer FF (Multi-layer Feature Fusion), and NMS algorithm have been the traditional algorithms considered for analysis. The results of this analysis are shown in table-2 and graphically represented in figure 8.

The results found that DBN explored 79.54% AC, BING+CNN showed 84.25% as AC rate, Faster R-CNN exposed 86.28% AC, YOLOv2 exhibited 90.05% AC, and Multi-layer FF and NMS explored 94.25% AC. Though conventional methods showed better and more satisfactory AC, the proposed system outperformed traditional methods by exploring 99.12% AC, as clearly presented in figure.8. On the contrary, FPR, MR, ER, and T(s) has to be minimum for an algorithm to be efficient. Correspondingly, analytical outcomes showed the efficiency of an introduced system

in terms of FPR, T(s), MR, and ER as it showed a minimum FPR of 4.78%, MR of 4.98%, ER of 10.25%, and T(s) of 0.02%. This rate is found to be less than conventional methods. It is graphically presented in figure.9. For instance, DBN showed maximum FPR at a rate of 24.13%,[5] which showed its ineffectiveness over the proposed work. Thus, high accuracy and low FPR, T(s), ER, and MR rates than existing methodologies have confirmed the efficacy of the proposed system.

Table-2. Comparative analysis concerning performance metrics [5]

Method	AC (%)	FPR (%)	MR (%)	ER (%)	T(s)
DBN	79.54	24.13	20.46	44.59	171.25
BING+CNN	84.25	18.68	15.75	34.43	6.41
Faster R-CNN	86.28	8.76	13.72	22.48	0.15
YOLOv2	90.05	6.26	9.95	16.21	0.03
Multi-layer FF and NMS	94.25	5.59	5.75	11.34	0.16
Proposed method	**99.12**	**4.78**	**4.98**	**10.25**	**0.02**

In addition, a general CNN and the proposed RRBCNN have been comparatively analyzed concerning the accuracy, specificity, FN rate, FP rate, and sensitivity. Obtained outcomes are presented in table-3. It is also graphically presented in figure 10 and figure 11.

Table 3. Comparative analysis in terms of performance metrics [35]

Model	Accuracy	Sensitivity	Specificity	FP Rate	FN Rate
CNN	98.4	98.62	98.13	0.0187	0.0138
Proposed	**99.12**	**99.25**	**99.05**	**0.0156**	**0.0115**

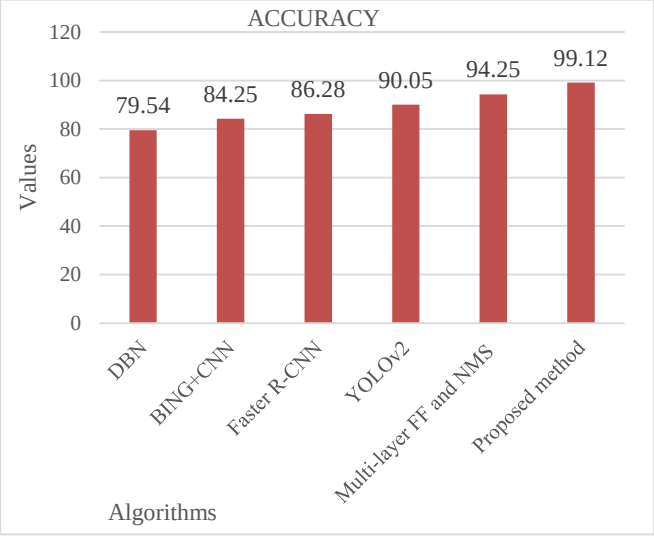


Figure 8. Analysis of the proposed and existing work[5] in terms of accuracy

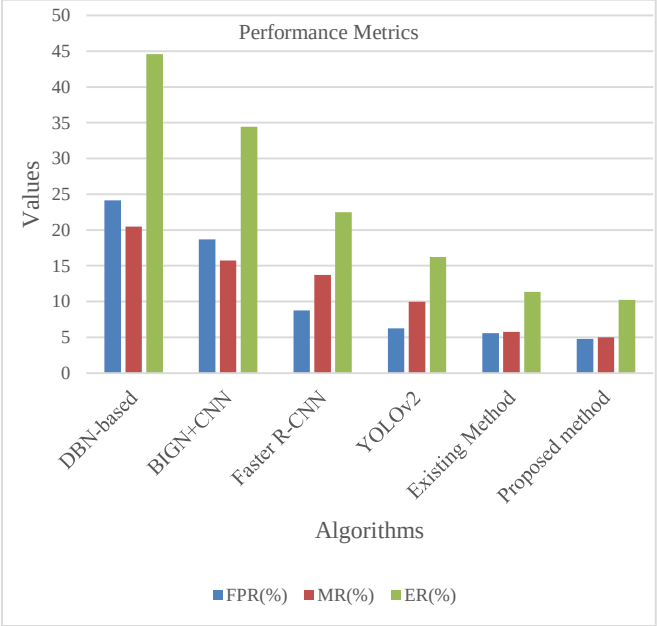


Figure 9. Analysis of the proposed and existing work [5] concerning significant metrics

From the results, it is explored that CNN showed 98.4% accuracy. In comparison, the proposed work exhibited 99.12% accuracy; the sensitivity rate of CNN was found to be 98.62%, whereas the introduced system showed 99.25%, specificity rate of CNN was 98.13%. In comparison, the proposed method explored 999.05%. Similarly, the introduced system's FP and FN rates have been efficient. In this case, FP and FN rates must be minimum to confirm the efficacy of a method. Accordingly, RRBCNN explored 0.0156 as the FP rate, less than the existing CNN that showed 0.0187% FP.

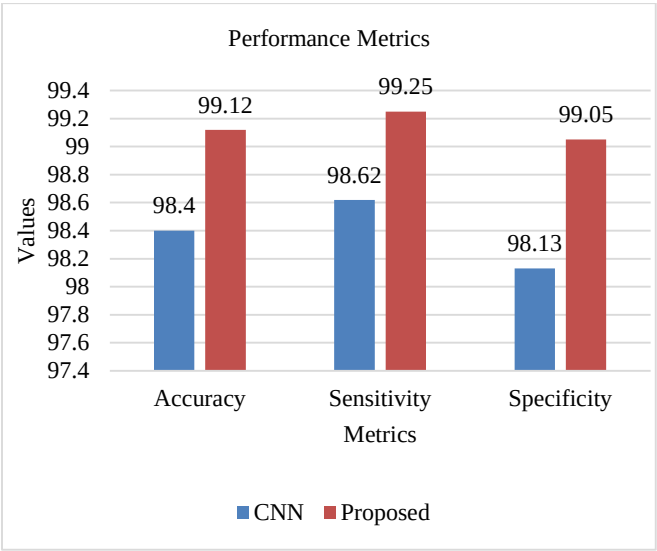


Figure 10. Analysis of the proposed and existing work concerning the accuracy, specificity, and sensitivity [35]

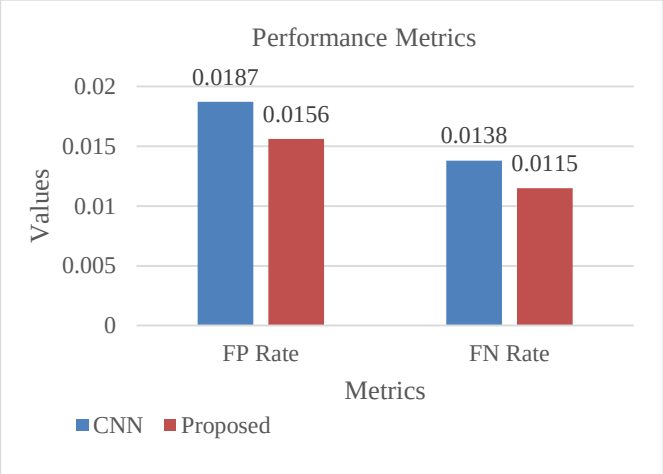


Figure 11. Analysis of the proposed and existing work concerning FN and FP rate [35]

Additionally, the analysis is carried out by considering existing methods such as the Rotation Invariant Parts Based Model, DBN, traditional variants of ResNet, Only conv4_x+HOG, Faster R-CNN, and Faster R-CNN+ResNet-50. These techniques are analyzed in terms of recall and precision. Obtained outcomes are shown in table-4 and graphically presented in figure 12.

Table 4. Comparative analysis in terms of precision and recall [29]

Method	Precision (%)	Recall (%)
Rotation Invariant Parts-Based Model	67.3	59.2
Deep Belief Networks	80.6	63.6
Faster R-CNN	87.8	82.4
Faster R-CNN+ResNet-50	88.6	83.7
Only conv4_x+HOG	89.6	84.1
An existing variant of ResNet	90.3	86.8
Proposed Method	**93.5**	**91.9**

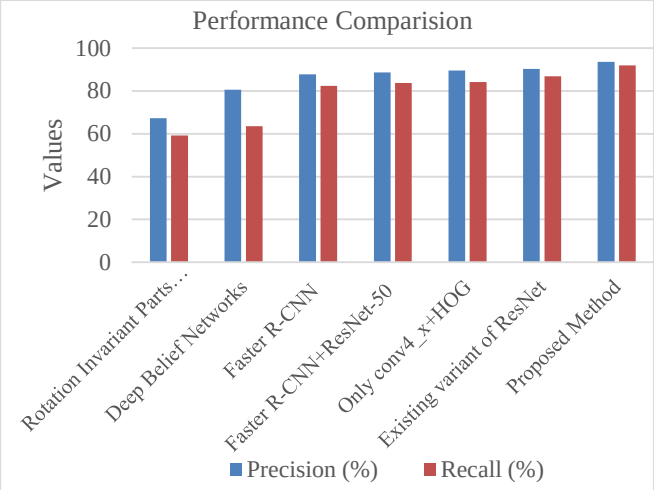


Figure 12. Comparative analysis of the proposed and existing work for recall and precision [29]

From the analytical outcomes, the existing DBN showed 80.6% precision; only conv4_x+HOG exposed 89.6% as the precision rate, while traditional variants of ResNet exhibited 90.3%. However, the introduced methodology explored high precision of about 93.5%, which is better than conventional methods. On the contrary, the recall rate of existing variants of ResNet has been high, showing 86.8%. However, the proposed system outperformed this by showing 91.9% recall.

Thus, the comparative analysis undertaken with three conventional works[5,29,35] explored the efficiency of the proposed system to existing works in terms of AC, MR, ER, FPR, T(s), precision, recall, sensitivity, FN rate, and specificity. In addition, RRBCNN introduced used SRM to downscale the feature maps that improve the training rate. This eventually enhances the prediction rate. Due to this, the efficient performance achieved by the introduced system is confirmed through comparative analysis.

CONCLUSIONS

The research mainly aimed to detect aircraft with better accuracy based on DL and image processing. For this purpose, RRBCNN and SRM were proposed. To effectively train the hierarchical domain-feature alignment, SRM was utilized by RRBCNN, which downscaled the feature maps to avoid information loss. This process has made it explore better outcomes. However, to confirm the efficiency of the proposed work, analysis was undertaken concerning AC, recall, sensitivity, precision, FPR, FN, ER, MR, T(s), and specificity. Three conventional types of research were considered for analyzing the performance of the introduced system. The outcomes revealed the efficiency of the proposed system that explored maximum AC, specificity, recall, sensitivity, and precision, while exploring minimum FPR, FN, ER, MR, and T(s). Thus, the proposed method efficiently detected aircraft with bounding boxes, showing high accuracy of 99.12%. This effective performance has made it highly suitable for aircraft detection and classification. Various other DL algorithms can also be focused on in the near future to improve accuracy.

REFERENCE

1. T. Wang, X. Zeng, C. Cao *et. al.* CGC-NET: Aircraft Detection in Remote Sensing Images Based on Lightweight Convolutional Neural Network. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **2022**, 152805-15.
2. L. Zhou, H. Yan, Y. Shan *et. al.* Aircraft detection for remote sensing images based on deep convolutional neural networks. *Journal of Electrical and Computer Engineering* **2021**, 2021.
3. S. Yin, H. Li, L. Teng. Airport detection based on improved faster RCNN in large scale remote sensing images. *Sensing and Imaging* **2020**, 21(1), 1-13.
4. Y. Wang, H. Li, P. Jia *et. al.* Multi-scale densenets-based aircraft detection from remote sensing images. *Sensors* **2019**, 19(23), 5270.
5. M. Zhu, Y. Xu, S. Ma *et. al.* Effective airplane detection in remote sensing images based on multilayer feature fusion and improved nonmaximal suppression algorithm. *Remote Sensing* **2019**, 11(9), 1062.
6. X.-Y. Tong, G.-S. Xia, Q. Lu *et. al.* Land-cover classification with high-resolution remote sensing images using transferable deep models. *Remote Sensing of Environment* **2020**, 237111322.
7. K. Bhosle, B. Ahirwadkar. Deep learning Convolutional Neural Network (CNN) for Cotton, Mulberry and Sugarcane Classification using Hyperspectral Remote Sensing Data. *J. Integr. Sci. Technol.* **2021**, 9 (2), 70–74.

8. G.J. Scott, M.R. England, W.A. Starms *et. al.* Training deep convolutional neural networks for land–cover classification of high-resolution imagery. *IEEE Geoscience and Remote Sensing Letters* **2017**, 14(4), 549-53.
9. Y. Li, K. Fu, H. Sun, X. Sun. An aircraft detection framework based on reinforcement learning and convolutional neural networks in remote sensing images. *Remote sensing* **2018**, 10(2), 243.
10. F. Ji, D. Ming, B. Zeng *et. al.* Aircraft detection in high spatial resolution remote sensing images combining multi-angle features driven and majority voting CNN. *Remote Sensing* **2021**, 13(11), 2207.
11. Y. Zhong, F. Fei, Y. Liu *et. al.* SatCNN: satellite image dataset classification using agile convolutional neural networks. *Remote sensing letters* **2017**, 8(2), 136-45.
12. F. Xuyun, L. Hui, S. Zhong, L. Lin. Aircraft engine fault detection based on grouped convolutional denoising autoencoders. *Chinese J. Aeronautics* **2019**, 32(2), 296-307.
13. Z. Song, H. Sui, L. Hua. A hierarchical object detection method in large-scale optical remote sensing satellite imagery using saliency detection and CNN. *Int. J. Remote Sensing* **2021**, 42(8), 2827-47.
14. L. He, C. Li. Visual saliency mechanism-based ob ject recognition with high-resolution remote- sensing images. *The Journal of Engineering* **2020**, 2020(13), 379-82.
15. L. Zhang, Y. Zhang. Airport detection and aircraft recognition based on two-layer saliency model in high spatial resolution remote-sensing images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **2016**, 10(4), 1511-24.
16. P. Ding, Y. Zhang, W.-J. Deng *et. al.* A light and faster regional convolutional neural network for object detection in optical remote sensing images. *ISPRS journal of photogrammetry and remote sensing* **2018**, 141208-18.
17. Z. Chen, T. Zhang, C. Ouyang. End-to-end airplane detection using transfer learning in remote sensing images. *Remote Sensing* **2018**, 10(1), 139.
18. Y. Xu, M. Zhu, P. Xin *et. al.* Rapid airplane detection in remote sensing images based on multilayer feature fusion in fully convolutional neural networks. *Sensors* **2018**, 18(7), 2335.
19. Z. Gang, H. Jingyu, X. Wenlei, Z. Jie. A mask R-CNN based method for inspecting cable brackets in aircraft. *Chinese Journal of Aeronautics* **2020**.
20. Z. He, L. Zhang In *Proceedings of the IEEE/CVF International Conference on Computer Vision* 2019, p 6668-77.
21. Y.-C. Lin, W.-D. Chen. Automatic aircraft detection in very-high-resolution satellite imagery using a YOLOv3-based process. *J. Appl. Remote Sensing* **2021**, 15(1), 018502.
22. B. Zeng, D. Ming, F. Ji *et. al.* Top-Down aircraft detection in large-scale scenes based on multi-source data and FEF-R-CNN. *Int. J. Remote Sensing* **2022**, 43(3), 1108-30.
23. S. Hwang, J. Lee, H. Shin *et. al.* In 2018 AIAA Information Systems-AIAA Infotech@ Aerospace **2018**, pp 2137.
24. J. Feng, J. Liu, C. Pan In *Chinese Intelligent Automation Conference*; Springer: 2019, p 482-93.
25. U. Alganci, M. Soydas, E. Sertel. Comparative research on deep learning approaches for airplane detection from very high-resolution satellite images. *Remote Sensing* **2020**, 12(3), 458.
26. M.J. Khan, A. Yousaf, N. Javed *et. al.* Automatic target detection in satellite images using deep learning. *Journal of Space Technology* **2017**, 7(1), 44-49.
27. Q. Liu, X. Xiang, Y. Wang *et. al.* Aircraft detection in remote sensing image based on corner clustering and deep learning. *Engineering Applications of Artificial Intelligence* **2020**, 87103333.
28. F. Zeng, L. Cheng, N. Li *et. al.* A hierarchical airport detection method using spatial analysis and deep learning. *Remote Sensing* **2019**, 11(19), 2204.
29. J. Yang, Y. Zhu, B. Jiang *et. al.* Aircraft detection in remote sensing images based on a deep residual network and super-vector coding. *Remote Sensing Letters* **2018**, 9(3), 228-36.
30. J. James, J.J. Ford, T.L. Molloy. Learning to detect aircraft for long-range vision-based sense-and-avoid systems. *IEEE Robotics and Automation Letters* **2018**, 3(4), 4383-90.
31. Y. Zhang, K. Fu, H. Sun *et. al.* A multi-model ensemble method based on convolutional neural networks for aircraft detection in large remote sensing images. *Remote Sensing Letters* **2018**, 9(1), 11-20.
32. A. Zhao, K. Fu, S. Wang *et. al.* Aircraft recognition based on landmark detection in remote sensing images. *IEEE Geoscience and Remote Sensing Letters* **2017**, 14(8), 1413-17.
33. A. Maher, H. Taha, B. Zhang. Realtime multi-aircraft tracking in aerial scene with deep orientation network. *J. Real-Time Image Processing* **2018**, 15(3), 495-507.
34. J. Yang, W. Zhao, Y. Han *et. al.* Aircraft tracking based on fully conventional network and Kalman filter. *IET Image Processing* **2019**, 13(8), 1259-65.
35. F. Ucar, B. Dandil, F. Ata. Aircraft detection system based on regions with convolutional neural networks. *Int. J. Intelligent Systems and Applications in Engineering* **2020**, 8(3), 147-53.
36. R.R. Karwa, S.R. Gupta. Automated hybrid Deep Neural Network model for fake news identification and classification in social networks. *J. Integr. Sci. Technol.* **2022**, 10 (2), 110–119.