National University of Sciences and Technology (NUST)
School of Electrical Engineering and Computer Science

Department of Computing

Digital Image Processing

Class: BSCS-9B

Research Project

Topic

Airplane Detection using Remote Sensing Imaging

Submitted By: Fatima Seemab and Mahum Samar

CMS ID: 291310,290647

BSCS 9B

National University of Sciences and Technology (NUST)
School of Electrical Engineering and Computer Science

Write brief abstract about the given topic.

Remote sensing is used to capture the images of the objects without making any physical contact with the object. The image acquisition is performed using special cameras. The object detection in remote sensing is of very high importance. The convolution neural networks (CNN) are used for object detection. This is a survey paper of researches done for the purpose of aircraft detection using digital image processing techniques with deep neural networks. Both the one stage and two stage detection of CNN are used for research of aircraft detection and recognition. One stage can be used for UAN (unmanned aerial vehicle) images and two stage can be used for the target detection for satellite images. The airplanes can be detected from the high-resolution remote sensing videos by using its frames and the occluded planes can be detected using the individual parts of the planes. The CNN also has the techniques for detecting the small sized targets.

Write detailed introduction about your survey paper.

Aircraft recognition of remote sensing is very important in both the civil and military application. Airplane recognition is object detection/target detection. As the airplane recognition is a task of great importance, besides this it is also faced by numerous challenges including pose variation, complex backgrounds, diversified target scales, and special perspectives, occlusion, viewpoint, resolution, illumination, changes, and shadowing effect, etc. The computational cost also greatly increased in case of high quality and large-scale images.

The research has evolved with the usage of conventional neural networks and deep learning techniques in airplane recognition. With the CNN use, the speed and performance of the

relevant task is greatly increased. There are mainly three types of methods used for object detection in remote sensing images.

Classic detectors consist of a sliding window where a classifier is applied on each grid for all the images.
Two stage detection algorithm: generating region proposals purpose is to filter out all the negative locations in an image and classifying it. They are region based convolutional neural network (RCNN)
One Stage detection algorithm: This type does not require the region proposal. Three methods used for this purpose are SSD, YOLO and RetinaNet.

In this paper the survey airplane detection using remote sensing is conducted. This paper describes about the core concept, pros and cons and limitations of the methods used in the airplane detection. The research is conducted using the two stage non-maximum suppression filtering for detection of airplanes in the remote sensing images. Other uses the spatial and temporal features of the remote sensing videos to detect the aircraft. One of the papers have detected airplane in remote sensing images using an effective method of ACF (aggregated channel features) while other research is conducted using local attention network of deep neural networks for detection of occlude aircraft due to any reason.
In one paper, the Online exemplar-based fully convolutional network for aircraft detection in remote sensing images is used. One way to detect the airplane is to detect the landmarks and apply template matching to it. Airplane can also be incorporated by target detection includes data preprocessing and head detection. The research has also been done on two staged detection using RLN as region proposal and faster RCNN for object detection. For very high-resolution remote sensing images PRROI (precise region of interest) is used along with the precise mask RCNN to avoid loss of precision. For unmanned aerial vehicle (UAV) a single staged detection algorithm YOLO has been used.

Provide the survey/comparisons related to topic.

1. Two-stage Airplane Detection with NMS Filtering in Remote Sensing Images

Many companies and researchers are interested in object detection based on deep learning. Two stage object detection better than one stage detection. The proposed method has detection based on Faster R-CNN. It is 2 stage detection using box detection method and feature pyramid networks for recognizing objects on different scale. Non-maximum suppression is used for postprocessing to avoid duplicate boxes due to multiple classes of airplanes and choose most confident class as final box result and delete the duplicate one. The 24 different classes detected thus improving recall and average precision metrics. The parts of F R-CNN are backbone that identifies feature maps, region proposal networks for region proposals using backbone output, RoI Pooling for different sizes input into same output size and the task branches used for detection and classification. Feature pyramid network Used for different scales detection. RoIAlign instead of RoI pooling to extract small feature map for accurate pixel alignment and feature extraction. Cross-entropy loss used to compute the loss function. It uses smooth L1 for regression task. Non-maximum suppression to correct class. The filtering operation is strength of NMS. The box with highest score are kept. 1600 images in 24 classes are used as a dataset. For training, ResNet-50 pretrained model as backbone. The recall and AP result are achieved through NMS.

2. Airplane Detection in Optical Remote Sensing Video Using Spatial and Temporal Features

The object detection for airplanes is conducted from optical remote sensing videos. The spatial features are obtained from structured forests edge detection and the temporal features from neighboring frames are obtained through improved frames differencing method. The detection is achieved using deep neural networks. The challenges include that the video size is large

while airplane is small so there is low contrast in background and target airplane. The spatial resolution less than 1m that results in poor motion frames. The contributions made by the proposed method are structured forest edges highlight contrast in airplane and background. The differences of the adjacent frames strengthen motion features of airplane in the video. In the fusion network, spatial and temporal features combined using deep neural network. In structured forest edges, the video is divided into frames. Since the edges are tight so non maxima suppression is applied. The grouping strategy results in colorful and discrete features. The edges are detected using this algorithm. For three differencing method in improved frames differencing, we use 3 adjacent frames and obtain 2 difference images and perform AND operation on them. The image enhancement is performed on spatial domain and the threshold segmentation and used along with morphological expansion. The expansion operation is used to suppress void phenomenon. The videos used are in AVI format. All the videos have been through pre-processing with geometric correction, radiometric correction and image stabilization. The dataset has the fixed wing airplanes and cloud occlusion is removed.

3. An Effective Method Based on ACF for Aircraft Detection in Remote Sensing Images

The multiscale sliding window framework is used based on aggregate channel features. The well-designed features are used that contain rich information for detecting the target airplanes. The fast feature pyramids algorithm is implemented for multiscale aircraft detection. The training is performed using Cascade AdaBoost with bootstrapping. The 2 step non maxima suppression algorithm is used for deleting the duplicate detection. In proposed method, bottleneck is the feature representation because computational cost increases. It used aggregated channel features (ACF). This has improved feature representation and as it is simple so computational price do not increase much. The method uses ACF and novel method for feature pyramid and deep exploration for adjusting the features accordingly which helps in detecting various types of airplanes automatically. The bootstrapping passes are used to train

detector and for merging 2 step non maximum suppression is utilized for separating airplanes with their own shadows. The registered map of the original image is called the channel which are translationally invariant. The ACF uses three channels. The channels of the training images are computed. The subsampling is performed by calculating the sum of every block in aggregated channel. Then a 1d vector is formed from aggregate channel. The proposed method uses the novel fast pyramid method, to decrease computational cost for detecting multiscale target airplanes in the images. AdaBoost algorithm selects small number of features that are important for detection as the input features can be very large in number. The Cascade AdaBoost is used to train the neural network. For the hard examples the bootstrapping is used to eliminate the negatives. Multiscale sliding window used for detecting fast pyramid construction and sliding window detection. pre-smoothing and post-smoothing are applied for faster detection. The overlapping detections of the target airplane can be deleted by using the non-maximum suppression. It used 2 step NMS. In 1st step greedy NMS i.e., greedy clustering by fixed threshold and 2nd step using the scale ratio and score ratio to remove overlapping.

4. Local Attention Networks for Occluded Airplane Detection in Remote Sensing Images

The detection of occluded airplanes is a huge challenge. The local attention network is proposed in this paper. Each visual part is represented by one channel. The divide and conquer rule is used to detect airplanes. The 4 visual parts of the airplanes are used i.e., head, left/ right wing, body and the tail. The additional attention branch is introduced so that individual features can be detected even for the occluded target. The circle is drawn around the geometric center for the visual part. The large part of the earth is covered with clouds every day or airplanes are occluded by the hanger's roof. The hand-crafted methods have been used in past and deep learning methods. The proposed methods used the detection by parts idea. The deep convolution networks are used for this method. The attention mechanism is used for investigating different feature location spatial correlation. The additional attention branch is

used for individual parts highlighting and suppressing the background in the images. The features are extracted using standard feature pyramid network and attention branch used to enhance the extracted features. The 3-detection head each having 3 branches are used i.e. local attention branch, detection branch and the classification branch. The output are the bounding boxes which are post processed by the non-maximum suppression to get the final output. The proposed method is designed on the detection system "RetinaNet". The loss of features detection is reduced by calculating attention map prediction loss, bounding box regression loss and anchor category classification loss.

5. Online Exemplar-Based Fully Convolutional Network for Aircraft Detection in Remote Sensing Images

The intraclass variation of airplanes make it difficult to detect them. The paper proposes end-to-end fully convolutional network (FCN) and to detect the intraclass airplane the online exemplar mining is embedded into this method. This reduces the cost of training the network for intraclass airplane detection. Firstly, relationship is built based on basic exemplar selected for metadata of scales and aspect ratio of labeled bounding boxes forming a group. The scale of the bounding boxes is adjusted based on intersection over union (IoU) between augmented bounding box and ground-truth box. Then updating the relationship using the space of high-level features using the clusters formed using the clustering algorithm and then these aircraft examples will be used for training the exemplar detectors. For efficiency, the proposed method developed circle response map (CRM) to construct FCN consisting of convolution layer and circle pooling layer. The high scores indicate the presence of aircraft target. The online exemplar-based fully convolutional network is proposed.

6. Aircraft Recognition Based on Landmark Detection in Remote Sensing Images

Different methods have been used for airplane detection in remote sensing. Classifying the images to detect the target has two main drawbacks. Firstly, it does not work on a new type of aircraft. Works only on the types present in dataset. Secondly requires large number of labelled training samples. Because of these disadvantages, a method was proposed to **detect airplanes using their landmarks detection** and template matching as the technique. For landmark detection we do not use the facial landmark detection technique because the RS images are taken only from top, and aircrafts can rotate in any direction.

A regression network is used for landmark detection. In this method whole image and all landmarks are processed simultaneously. Eight landmarks are used to specify the aircraft. Labelling of aircraft landmark is shown in figure 1.

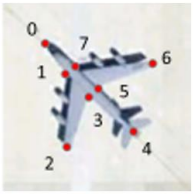
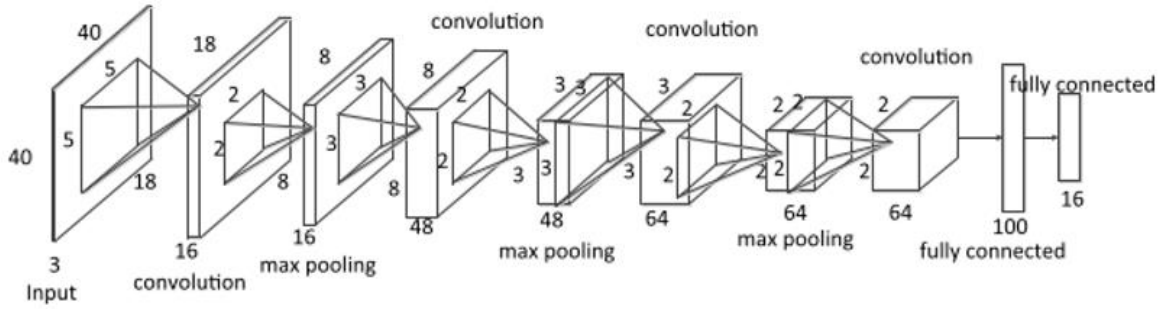


Fig.1

The image dataset is improved by reflecting and rotating the images. The images in the data set are sampled to 40 x 40.For aircraft rotation, mirror images are included in the data set by rotating the image 5 degree from 0 to 370 degree. This data processing increases the performance.

This techniques' CNN is six layered base model based on vanilla CNN(method for facial landmark detection).It has 4 convolutional layers and 2 fully connected layers followed by max pooling. Loss Method detected for the proposed CNN also includes the position of wingtips



For airplane recognition in this technique, we used the 8 landmarks as a template. To calculate a template vector first the image is rotated so that its body direction is upward and then calculate those eight landmarks in the coordinate system. For recognizing the object Euclidean distance is calculated between the target and template.

The experiments results have proved that this method can handle different types and pose of aircrafts

7. AIRPLANE DETECTION AND RECOGNITION INCORPORATING TARGET COMPONENT DETECTION

Airplane detection is target detection. Target detection involves feature extraction and classifying targets. Target detection performance and speed is greatly increased with the advancement of CNN and deep learning. But the Target in remote sensing have some challenges because of complex backgrounds, diversified target scales, and special perspectives.

For target detection, previous researchers used several techniques including candidate object region and CNN, Region Local Network, hierarchical detection framework to detect the airport first and then perform aircraft target detection. But the missing part in these researches is accuracy of recognition between similar airplanes

These remote images can be oriented and arbitrary oriented bounding boxes are used for labelling. Research has improved the performance of object detection in OBB (arbitrary oriented boundary box).This researches method used the advantage of oriented target by OBB, but the head is uncertain in 4 direction. This process also used object head detection method.

The method used in this research is explained from this diagram.

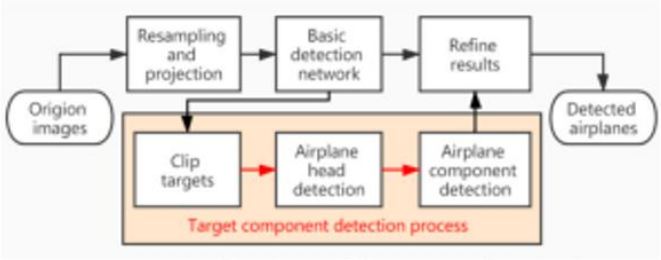


Fig. 2: The flowchart of the proposed method

Missing target (Distortion) might occur because of different project coordinates in remote sensing. The method UTM (Universal Transverse Mercator) used in projection conversion and then down sample the images to 0.81 meters. CNN method used for basic detection network is R 3 Det [9] with RetinaNet. Its refinement module is a solution for feature misalignment makes the bounding box more accurate. Objects have direction information. We start from the left of the nose and take the slices and rotate in 4 orientations 0,90,180 and 270 which makes the training dataset representation of 4 orientations. The reason for the issue of recognition between similar airplane categories is that the info which differs them is located on the local engine components. In nose up airplane slices, we take slices of tail engine components and design CNN to train and classify the component slices.

8. FAST AIRCRAFT DETECTION BASED ON REGION LOCATING NETWORK IN LARGE-SCALE REMOTE SENSING IMAGES

The two different techniques used in above methodology is first to recognize the type of airplane and second is specified to recognize the similar like airplanes types but till now the large scale images cannot be used for object detection but as the RS images resolution becoming higher, there is a need to have a mechanism which will deal with high scale remote sensing images. Object detection is time consuming for large scale images and algos design for simple images cannot be used for special objects like aircraft that's why RLN is used to improve after RCNN framework. RLN basically identifies the areas where these objects can be found for example runway. Common sliding window and cropping method is not enough to compete for this work because of long running time and redundancy process.

The method introduced in this research is RLN based aircraft detection in large scale images. Image resizing is a common technique used for saving cost of time. But after resizing, object detection is not possible because of small size of objects. To deal this we calculate the ratio of annotated object to image. If ratio is relatively small, object is difficulty to be detected after resizing. Second Tech is to use the image segmentation method, but it has its own problems. So a solution is to first look at whole image and find the region where aircraft can be detected To implement it we made the computer to recognize the region using an RLN named region locating network and then in that region aircraft is detected using faster R-CNN.
Two main techniques for object detection are end to end and region based.
In range-based proposals from images and then classify them. CNN methods used includes R-CNN [4], Fast R-CNN [5] and SPP-net [6]
In end to end, divides image in 7*7 parts and each parts look for object. SSD uses different separation numbers like 4*4 and 8*8.Bigger boxes detect big objects while small ones are used for smaller size object detection.
This research is based on following three steps
Step 1, locate the region of aircrafts using Region Locating Network (RLN).
Step 2, locate aircrafts with Region Proposal Network (RPN).
Step 3, classify the aircrafts with Fast R-CNN and final process.

Data training can only detect the location of an object not the region, so we have to perform data preprocessing. Clustering algorithms are used for this purpose. In this technique hierarchical based methods are used. We made an independent cluster for each airplane and then performs multiple iterations to cluster them together if the distance between two are less than a threshold put it in the same cluster. Two thresholds are set for size. Some aircrafts far from the cluster are not added because of negative effect.

RLN is fully conventional network. Is structure is same as RPN which has 16 layers. We use the preprocessed training data as RLN. We use IOU (intersection over union) whose formula is So/(S box1+S box2-So) where box 1 and box 2 are two area proposals. If value is in between 0.2-0.7 its a positive input else not.

RCNN includes RPN and object detection network. CNN takes the output of RPN and classifies it for the proposal. We use 4 convolutional layers in the process.

9. OBJECT DETECTION AND INSTANCE SEGMENTATION IN REMOTE SENSING IMAGERY BASED ON PRECISE MASK R-CNN

A lot of advancement in remote sensing but there are still issues because of occlusion, viewpoint, resolution, illumination, changes, and shadowing effect, etc. Further increased computational cost in case of high quality of images, creates difficulties for object detection. They have poor performance in high resolution images because of scale diversity, visual specificity, small target problem, multi-directional problem.

A Precise Mask RCNN is used for object detection and instance segmentation. Targets of RS are usually small and there is a loss of precision after CNN's pooling layer. So in this method PrRoI pooling layer is used because they have that has a continuous gradient on bounding box coordinates.

Mask RCNN is a two stage procedure and RCNN are used with the PROI pooling.

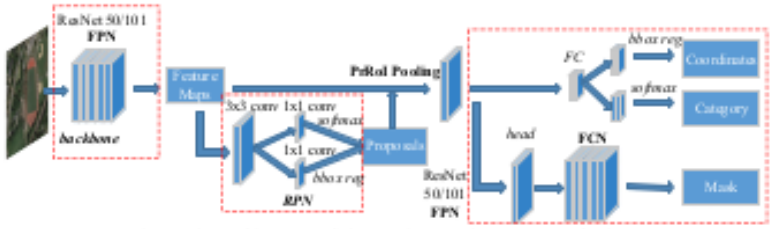


Fig.1 The architecture of the precise Mask R-CNN method.

A region proposal networked named as fpn feature pyramid network generates the bounding boxes proposals. Then regression is performed using RCNN and a mask branch for segmentation.

Because of variation in sizes of objects, researchers used FPN with RESNet of depth 50 or 101 layers and then fuse them which has improved the accuracy. The Fast R-CNN branch includes a

classification and a regression task. The main function of the branch is to provide more accurate bounding boxes for detection. The mask branch is a small FCN applied to each RoI, predicting a segmentation mask in a pixel-to-pixel manner.

ROI pooling performs the quantization on the coordinates having gradient on bounding boxes. This has been avoided by Precise RoI Pooling (PrRoI Pooling).

10. Detection of Airplanes on the Ground Using YOLO Neural Network

The UAV based technology has a great benefit in terms of cost and perform. This paper based on airplane detection in aerial images from unmanned aerial vehicles. This paper describes the result of two single stage methods YOLO and Tiny YOLO v3.YOLO is the successful method for detection from the video.

UAV process consists of three stages.

Raw data is recorded during flight.

Processing of real time data.

Autonomous decision based on data.

All stages should be performed in fraction of seconds, so we use a single stage detector based on CNN.CNN has parallel nature and runs on GPU increases the speed. In this method we do not go for the proposed area, just evaluate the whole image. Input image is represented as size of nxmx3 n and m are width and height. Various sizes images are resized to 416*416 because the output gives an odd number of cells which we can locate the single center easily. The image is divided in 13x 13 cells. Amount of bounding boxes covered by a cell is also recorded.

Although CNN has some yolo network training on it but requires some additional parameters: batch size, learning rate, decay, iteration number, and detection thresholds.

The images from UAV are different as taken from top-down view which increases the performance for the network having satellite images. Higher precision level is achieved using yolo v3 as the neural network.

Provide at least 5 implemented and technological examples of your topic.

The airplane detection in remote sensing images has a vast application in for both in military and civil.

Military:

It is used for protecting the air space on the borders. When the airplane enters the territory, it can immediately capture it if it not owned by the country. It is also used when training the fighter pilots for efficient training and avoid any accidents. We can detect the number of airplanes present in the restricted area.

Civil:

The air traffic is controlled using remote detection. The number of planes on the route can be checked and their routes can be determined.

Provide pros and cons of all the five papers.

In Two-stage Airplane Detection with NMS Filtering in Remote Sensing Images have good and eligible performance on average recall. It mapped all planes classes. The good detection of background and foreground can be achieved through this method. The precision of recognition task with the data of imbalanced classes has to be improved so Balanced Cross-Entropy Loss might be used to improve it.

In Airplane Detection in Optical Remote Sensing Video Using Spatial and Temporal Features the improved frames differencing has been helpful in detecting left to right plane moving targets and smaller targets from right to left. It has a simple implementation and low design complexity. It is highly stable. The precision, recall and F-measure is quite high and competitive in high resolution video. But the more optical remote sensing videos dataset is required to perform the in-depth research of this proposed method.

In An Effective Method Based on ACF for Aircraft Detection in Remote Sensing Images the better results even when the background is complicated, airplanes are blurred or occluded. The detection and training are fast due to the simple form of feature detection and fast pyramid method and the usage of LUV color space and gradient channels, it is illumination invariant and the bootstrapping for negative samples by inserting hard examples. The accuracy depends on the dataset and the tasks which are to achieved.

Local Attention Networks for Occluded Airplane Detection in Remote Sensing Images the airplanes can be detected even when occluded by clouds or even when a part of the airplane is visible. It has improved detection accuracy. The proposed method performs better than Faster RCNN, SSD and RFB-SSD. The detection speed is comparable to the "RetinaNet". The

addition attention branch has to be trained for all the parts of the plane which can be computationally expensive.

In Online Exemplar-Based Fully Convolutional Network for Aircraft Detection in Remote Sensing Images, no prior size and aspect ratio required for detection is required. For using high-level features and labelled information it is more reliable and robust. The CRM increases the speed and accuracy. As compared to the AdaBoost using HoG and AdaBoost using ACF, proposed method achieved greater accuracy and recall. But it worsens the precision because of heavy burden on the classifier. The OEM based study on weakly supervised learning is required for better results because it is expensive and unreliable for small targets to conduct manual annotation in large scale remote sensing images.

The landmark detection method has the advantages including the landmark detection technique method does not require us to have many training samples which is sometimes difficult to collect. This also reduces the effort of labelling the type also.
This method is expandable so that it can process the aircrafts recognition for those types which are not trained in the model. For a new aircraft, a landmark template type is required for detection instead of collecting a large amount of training sample.

Airplane recognition and detection incorporating target components produces better results than other CNN methods for target detection. From experiments, object head detection resnet18 is used whose validation is 100% and test is 99.89%. Component detection gets better performance than direct classification using slices. But this method has limitation that there is still chance of improvement in similar looking airplanes.

The method used for large scale images have the advantages detect mini size object make speed faster and accuracy increases. Precise Pooling technique avoids loss of precision. This research's result shows that the method is more suitable for VHR remote sensing images than the R-CNN.

The detection of airplanes on ground using YOLO has an issue when the size of image is large or not present in the data set, the network does not perform accordingly. This research is one stage edge detection.

Discuss DIP related issues in these papers and what can be improved.

In an effective method based on ACF for aircraft detection in remote sensing images, where the non-maximum suppression is applied for first step the fixed threshold must be chosen correctly for the accurate suppression.
The dataset must be correct because the image acquisition problems can cause the issued in the processing of the detection techniques.
The pre-processing techniques must be applied on the images before detection to achieve the appropriate results i.e., removing the noise, or image enhancement techniques.
The morphological operations can improve the detection of aircrafts like dilation can be used to detect the small target airplanes in the remote sensing images.
All the researches have been conducted in the spatial domain but the processing in the frequency domain is computationally fast. So, research can be conducted to apply detection techniques in the frequency domain.

Conclusion & Future trends.

Conclusion:

Since the development of convolution neural networks, it is utilized in remote sensing images for target detection like in airplanes detection. It has revolutionized the object detection by making it more precise and accurate by detecting the small targets and efficiently differentiating between foreground and background. The occluded planes can also be detected by using the individual parts of the airplane. It can also differentiate between closely located objects. It has applications in both military and civil. A vast research is being carried out to continuously improve the algorithms.

Future Trends:

The improvements in the image acquisition are being made which will ultimately improve the datasets for the remote sensing images. The satellite technology is constantly being enhanced. Remote sensing images resolution is quite high, so techniques are required to be computationally faster and efficiently perform detection. The algorithms for deep neural networks and convolutional neural networks must be improved to remove the deficiencies faced for object detection and recognition.

References:

[1] Y. Song, J. Tian, “Two-stage Airplane Detection with NMS Filtering in Remote Sensing Images”, 2020.
[2] J. Bai, W. Yu, A. Yuan, Z. Xiao, “Airplane Detection in Optical Remote Sensing Video Using Spatial and Temporal Features”, 2020.
[3] A. Zhao, K. Fu, H. Sun, X. Sun, F. Li, D. Zhang, H. Wang, “An Effective Method Based on ACF for Aircraft Detection in Remote Sensing Images”, 2017.
[4] M. Zhou, Z. Zou, Z. Shi, W. Zeng, J. Gui, “Local Attention Networks for Occluded Airplane Detection in Remote Sensing Images”, 2020.
[5] B. Cai , Z. Jiang, H. Zhang, Yuan Yao, S. Nie, “Online Exemplar-Based Fully Convolutional Network for Aircraft Detection in Remote Sensing Images”, 2018.
[6]A. Zhao et al., "Aircraft Recognition Based on Landmark Detection in Remote Sensing Images", 2017.
[7]H.Jia et al.,”airplane detection and recognition incorporating target component detection”,2021

[8]Z.Han et al.”fast aircraft detection based on region locating network in large-scale remote sensing images ”,2017
[9] H.Su et al.,”object detection and instance segmentation in remote sensing imagery based on precise mask R-CNN”,2019
[10] V.Kharchenko,I.Chyrka,”Detection of Airplanes on the Ground Using YOLO Neural Network”,2018,