Detection Of Small Objects Based on Deep Learning

SONALI GUPTA[1], RUCHI GARG[2]
[1] *Assistant Professor, J.C. Bose University of Science and Technology YMCA Faridabad*
[2] *J.C. Bose University of Science and Technology YMCA Faridabad*

Abstract— The pervasive and wide applications like scene understanding, video observation, advanced mechanics, and self-driving frameworks set off tremendous exploration in the area of computer vision in the latest ten years. Being the center of this multitude of uses, visual acknowledgment frameworks which incorporates picture order, confinement and location have accomplished extraordinary exploration energy. Because of critical improvement in neural organizations particularly profound learning, these visual acknowledgment frameworks have accomplished amazing execution. Object identification is one of these areas seeing incredible accomplishment in computer vision. This paper demystifies the job of profound learning strategies in view of convolutional neural organization for object recognition. Profound learning structures and administrations accessible for object identification are additionally articulated. Profound learning strategies for cutting edge object location frameworks are evaluated in this paper.

Indexed Terms-- Intelligent driving, Object detection, CNN, Small objects

I. INTRODUCTION

In late years, profound learning innovation has been generally Applied in the fields of the machine vision, for example, the identification of objects, the recognition of the face, the automated driving and helped driving, with the rapid advance of the innovation of consciousness made by man. The location of the item is to use the techniques for the management of the image and the example confirmation to recognize the item in the image. It is a moving fact to use PC innovation to recognize the recognition of objects. The complex foundation, the impediment of objects, the low goal of the low image, a small scale of articles and the change of disposition influence the impact of the discovery of elements.

Conventional articles recognition calculations are generally based on the sliding window model and the need to free themselves includes incredibly. There is a safe subjectivity in the extraction of false components, which has the unrest, intricate of estimation and defenseless practice. During the challenge of the location of the Image net object from 2010 to 2010, individuals understood that the progress of the identification of the article was extremely slow. Using usual strategies to build the location frame of the built-in item requires a lot of work, but only a bit of advantage. Consequently, it is important to propose a higher technique to work on the presentation of the article discovery. So just in time as 1998, [1] proposed the structure of the convolutional Neural Organization Lenet5 for the recognition of numbers, and the neuronal organization grew gradually in the ten years that accompany it. In 2012, [2] he proposed the structure of Alex net, who won the 2012 Image net test at one time, and the impact of recognition was a long way in customary calculus, along these lines.

Establish an impact on significant learning computers. With the persevering improvement of the execution of PC equipment, the estimation of the acknowledgment of the article subject to significant learning has been transformed into the standard calculation due to its fundamental and convincing association structure. In request circumstances, such as machined air conditioning vehicles, the items of the image obtained by the crashing rocket are small, with information on little and tenuous components. The small revelation article is transformed into an inconvenient problem to handle the detection of objectives

In 2013, [3] proposed the calculation of R-CNN and it effectively applied it to the protest location. From that point, the location of the Deep Learning-dependent object has entered the phase of an unstable turning of events, providing innumerable discovery techniques of deep learning objects, such as R-CNN, YOLO,

SSD, et al. These articles discovery calculations can be divided into two classifications according to various pensions: Calculations of location of objects that depend on order and the calculations of recognition of elements that depend on the relapse.

A. Two-stage algorithms

The location calculations of the article that depend on the group are divided into two phases, so they are also called two-stage calculations. To begin, the calculation must eliminate the regional configuration of the competition, then, at that time, the class guarantee and the area change of the UP region and eating are completed. Finally, give in the consequence of the discovery of articles. Agents' calculations are the calculations of the R-CNN series that depend on the extraction of the area, such as R-CNN, SPP-Net, R-CNN

(1) R-CNN

In 2014, [3] A calculation of the location of the R-CNN article is proposed depending on the extraction of the area above and the body. In contrast and the usual technique, the calculation of R-CNN has worked incredibly in its presentation, and has become the establishment of the calculation of the discovery of the local-based extraction object. Figure 1 for the scheme of the calculation transmission of R-CNN. The calculation of R-CNN changes the problem of the recognition of elements in the problem of the district order by separating the premises from the competition, which enjoys the benefits of simple technique and solid expansibility. I know that, as you can, the calculation of R-CNN actually has problems: (1) The organization must produce around 2000 districts of applicants and concentrate the elements of each UP and eating area. In this interaction, R-CNN is entangled independently, and there is a ton of reajusions activities, which confine the exposure of the calculation; (2) In the calculation of R-CNN, the regional configuration of the competition must be distorted and climbed, and the contribution of the CNN network has been used; (3) R-CNN that uses conventional strategies to create an upper and eating region, restrict incredibly the speed of the algorithm.

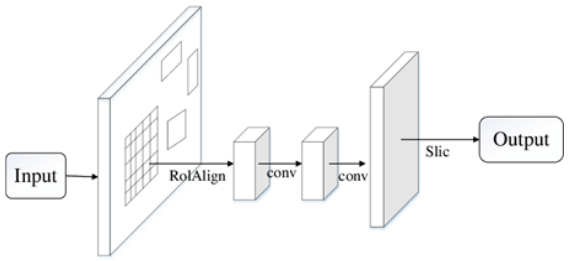

Fig.1 RCNN Flow Chart

(2) SPP-Net

In To address the restriction of the district's size of the competition in the calculation of R-CNN, [4] he planned a layer of spatial pyramidal pool (the structure is shown in Figure 2) in 2014. They train the convolutional neural organization for Manage the areas of subjective size competitors and delete. The impediment that the convolving neural organization can recognize the contribution of fixed size. SPP-NET also works on the monotonous activity of the R-CNN network in the interaction of extraction of elements of the applicant tables, working on the effectiveness of the calculation. SPP-NET understands the contribution of multiple scales from the CNN organization, and only makes the convolution of the first image once, which saves a ton of calculation time. However, there are still some topics, for example, the contradictory size of the preparation image, the huge field of vision in the space of interest of the Competition Case used, and the lack of use of directional.

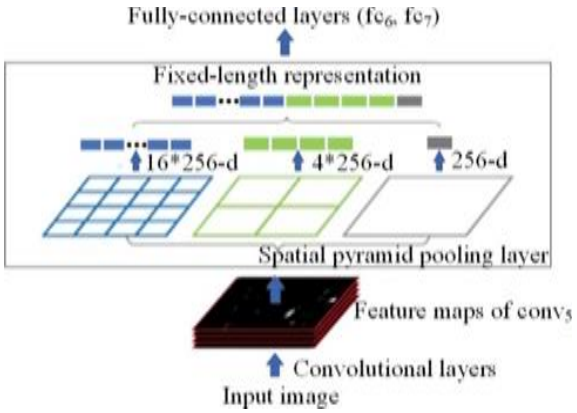

Fig.2 : SPP-Net Space Pyramid Structure

(3) Fast R-CNN

In 2015, [5] He proposed the most developed FAST R-NNN-dependent calculation of SPP-Net. Reland R-CNN moves the fundamental organization to the R-

CNN Sub Organization and offers the activity of the convolution after the general activity of the image of the information is completed. First, the selective search was used to acquire around 2000 districts of interest in the photos. Then, at that time, the whole image is sent to CNN to eliminate the most outstanding aspects. The last layer will be interested in the Territorial CNN guide in the attributes of the convolution letter, and use the loss of Softmax and L1 loss to group the probability and line of line relapse. R-CNN Quick understands the beginning of finishing the preparation mode, but actually uses a particular hunting calculation, which expands the calculation time.

(2) Faster R-CNN

In request to tackle the issue of Fast R-CNN calculation, [6] He proposed the rapid R-CNN calculation in 2015. R-CNN uses RPN (regional proposal networks) to supplant the calculation of particular hunting. It is isolated in two sections: the initial segment is to separate the profound organization convolutionary from competitors' premises; The back is to use the Fast R-CNN Finder depending on the extraction of area. RPN Network is mostly used to create the Item Region of the article. First, a progression of anchor boxes are created, which are cut and separated, and the classifier is used to recognize if they have a place with the nearest object or the foundation. In the intermediary, the limit relapse box adjusts the anchoring box to frame a more accurate proposed region. Figure 3 is the RPN network structure table.

The fastest R-CNN calculation entered the complete image for the first time and entangled it to obtain the component map. Then, at that time, the element diagram is entered into the RPN organization to obtain the data from the applicant's box. Then, at that moment, it is ordered by the qualities of the extracted and eating districts. Finally, use the return to change the situation of the top and science box. Figure 4 shows the construction of the faster R-CNN system. calculation.

B. One-Stage Algorithms

Article The calculation of the location that depends on the relapse is known as a stage calculation, will protest with the discovery process is improved in a solidarity with the start to finish the problems of relapse. It is not the same as the calculation of two stages that depends on the extraction of the location, the calculation of a stage can understand highlighting the exchange through a solitary preparation, and its discovery speed has been extraordinarily improved. The regular calculations of location of objects of a stage incorporate YOLO, SSD and different calculations.

(1) Yolo

In 2016, [7] proposed another calculation of the article location, YOLO. Yolo's calculation sees the recognition of objects as a problem of spatial relapse. The neuronal organization can obtain the expectation of the limit box and the likelihood of class from the total image by an activity. Yolo's calculation initially changes the size of the image, then, at that time, enter the image in a solitary organization with convolution and cycle the results of the identification according to the certainty of the model. The general design and the discovery interaction of Yolo's calculation are moderately simple, so the recognition speed is fast. The calculation of Yolo can make the world expectation dependent on the semantic data of the image, and the location rate of the Foundation is low. However, Yolo's calculation has the topics of low substitution and the low revision rate, and the impact of defenseless recognition for objects with exceptionally brief distances and small areas. In addition, the speculation capacity of Yolo's calculation is moderately. weak.

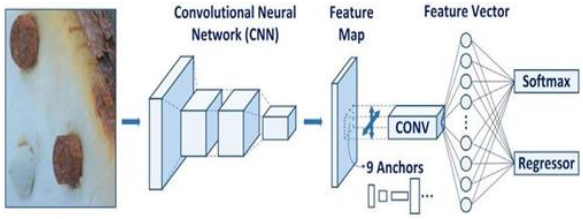

Fig.3 : RPN Network Structure

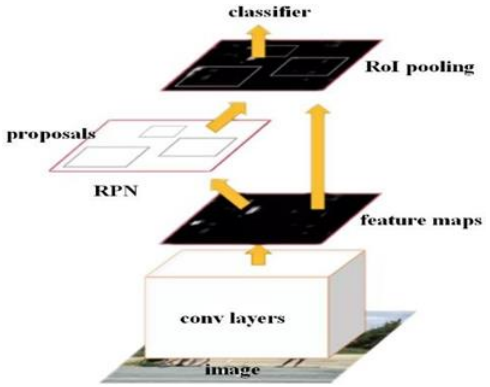

Fig.4 : Faster R-CNN Framework Structure

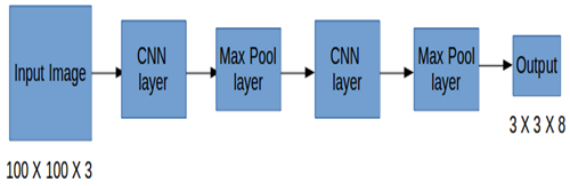

Fig.5 : Yolo Flow Chart

a. YOLO9000

In Request to address the topics of low situation and the low rhythm of Review of Yolo Calculation, [8] proposed another calculation of articles recognition YOLO9000 in 2016. YOLO9000 can distinguish 9000 items. Given that the calculation of the recognition of articles has fewer data and sum of information in the informative collection, while the characterization of the article has a lot of information and data of rich elements, YOLO9000 proposed a joint preparation strategy, which can combine the information of the Location and order information to prepare. Essential thinking is to prepare the provider of articles both in the informative index of the location and in the grouping of the information collection simultaneously, to obtain competition with the exact situation of the article using the information of the observation information collection, and to build the Quantity and work of the layout. About anguish using information on the characterization of informative collection.

b. Yolov3

Use the DarkNet53 network as an organization of component extraction. The DARKNET-53 organization obtains Resnet's thinking and adds the remaining module, which successfully addresses the question of inclination of inclination and slope that vanish in a deep organization. When the remaining module is composed of two convolutional layers and an association alternatively, there is no cluster layer or a layer of complete association throughout the Yolov3 structure. Yolov3 adds the ability to decide if an object is included in a candidate box to reduce detection errors. Using a binary classifier, each candidate box can predict multiple classifications [9].

c. SSD

In 2016, [10] proposed the SSD calculation to address the subject of the inadequate accuracy of the YOLO calculation in the object located. The SSD calculation can relate directly the area and the class of the article in an organization, so the

The identification speed is extremely fast. When performing forecasts, the SSD produces the likelihood comparison score for each element class in the default box and changes the default box to achieve a decent coincidence with the shape of the item. In addition, the organization can cause the forecast to depend on several guides of item elements with several objectives, and recognize the recognition of multi-scale objects. SSDS kills the extraction of the region and the pixel test or includes, and exemplifies all the calculation in a solitary organization. SSD enjoys the benefits of simple and effective calculation, simple preparation and coordination. I know that, as can, since the small article is related to the small region in the component diagram and the element is not evident, the impact of SSD identification in the small article is not yet great. When there is no place in the competition in the calculation of SSD, the district relapse is problematic and the Union is problematic. In addition, multi-layer component diagrams of SSD organization will be used as a free contribution from the characterization organization, and a similar article will be identified by boxes of various sizes simultaneously, which provides a Rehashed computation.

II. SMALL OBJECT DATA SET

Small object location It can be applied in the recognition of terrestrial objects by automated high vehicle, discovery by the rush hour, traffic signal identification and different scenes. It assumes a vital part in common, military, security and different fields. For the most part, the length and width of the article is

of the first size, which can be considered as a small article. However, the particular size of the small article could fluctuate a bit according to its application situation. [11] He found the size of the small article to 32 * 32 pixels in the coconut data set, and proposed an internal and external network design to recognize the objects using the district's internal and external data. [12] He distributed an informative index committed to the small recognition of objects, in which the size of the small article is 20 * 20 pixels at 512 * 512 pixels. Small articles have poorly size, low and outstanding objectives uneditable, so a single data set and a greater precision calculation for the identification of objects are required. The small article involves a moderately small region in the image, and the trademark data of the edge are not evident or in any case, missing. Due to the goal and the restricted data, the impact of discovery of small articles that depends on a deep learning is not ideal in the information collections of recognition of standard articles, and explicit informative indexes are required to the attributes of small articles. There are some information indexes of location of small items generally used.

(1) Pascal VOC

The informative collection of Pascal VOC contains dialing data for 20 types of items. It is a non-unexpected informative collection for the location of the object and the division of items. Albee The data set of Pascal VOC is anything, except a dedicated data set for the discovery of small items, is a set of legitimate general articles recognition data. Pascal VOC contains two informative collections, VOV2007 and VOV2012. VOC2007 contains 9,963 marked images, which are isolated in three sections: train, val and test. VOC2007 has 24,640 items. The Item Identification Informative Index of VOV2012 was separated in Trainval and Test, in which there were 11,540 named images and 27,450 data from named objects.

(2) MS coconut

MS Coco (Common Objects in Context) is a common information collection provided by Microsoft to the recognition of protest in images. It is considered one of the information collections of discovery of more definitive items in the field of PC vision. Coco Dataset is a huge discovery of articles, division of objects and subtitle information. Since the ImageNet challenge stopped, Coco Challenge has become the most legitimate challenge in the field of article location and the recognition of articles. It is also the main global challenge in this field that unites Google, Microsoft, Facebook and many of the main colleges and inventive efforts in the country and abroad. Cocos contains 91 items, 328,000 images and 2,500,000 brands. The largest informative collection of semantic division has been delivered to this point, with 80 classifications, in excess of 330,000 images, of which 200,000 are discussed, and the absolute number of people in the information collection exceeds 1.5 million. There are more photos in each class, which helps with a greater capacity to find a specific scene in each classification. Contrasted with the data set of Pascal VOC, it has more classes and images, and a greater number and extension of little articles.

(2) VEDAI

VEDAI (Detection of vehicles in aerial images) is an informative collection of aerial images, which is an informative index to prove little calculation of detection of articles in the climate without restrictions. Despite its small size, the articles in this set of data also have several perspectives as different headers, light changes, specular reflection and impediments. In addition, each image is furnished with an assortment of ghostly groups and goals, as well as a precise judgment conspire. Vedai has great relocation and speculation capabilities and can change comment records according to the requirements of the task. Convert the square addresses marked in the upper left organization of the rectangular box and the width and length of the rectangular box. It is used for the control and the flat examination of the small calculations of recognition of objectives of the Earth vehicles of Earth Earth and has achieved great results.

d. DOTA

DOTA (a huge set of scope data for the detection of objects in the aerial images) is a large set of data for the detection of objects in the aerial images and can be used to evaluate the small models of detection of objects in images. The data set provides 2,806 high images of several sensors and stages, in size from 800 * 800 to 4000 * 4000 pixels. Contains objects of different scales, headers and shapes, and contains

names of 15 normal items, including the name of 188,282 s.

III. SMALL OBJECT DETECTION

Multi-scale Discovery of objects Particularly small location of the item is a difficult task. From the impulse research, it can be isolated very well in the accompanying strategies to administer the issue of small recognition of articles. One is the exemplary technique is to combine the data of the map component of the element extracted by the Element Extraction Organization. One is to use the deconvolution to extend the depth of including the map; Another is to change the informative index to identify small articles, more precisely. Something, these calculations address the issue of the scale change of articles and further develop the accuracy of location of small articles, but there is still a ton of opportunity to improve, which can not meet the needs of the exact identification of small children articles. For the location of small articles, many explorers have been completed at home and abroad from recent times, and some successful algorithms have been proposed.

In 2016, [13] proposed a model of calculating small bird articles that depend on a profound convolutional network, which is to address the issue of the location of small articles in the huge area of foundation. This technique consolidates the calculation of discovery of articles with the semantic division, prepares a deep convolutionary neural organization (FCN) and FCN variations, and incorporates the results through the Help Machine (SVM), with the purpose of Perform the execution of the identification of the small increase in the article. [14] He proposed a network of depth agency of the square region at the pixel level to understand the grouping of a single pixel in ether images and configure the land cover guide to understand the identification of small elements. [15] proposed a technique to eliminate the provincial pixel elements by pixel, and use the convoluted neural organization to characterize and recognize the most outstanding aspects. I know that, as can, the productivity of this strategy is excessively low, and the multiple size of the article is not versatile.

In 2017, [16] proposed a characteristic pyramid organization (FPN) with a level association, using multi-scale feature data to achieve the location of the small location of the article. The calculations of recognition of previous objects only use the elements of the upper layer of the organization for the discovery, although it contains rich semantic data, however, after the endless supply of the grouping activity will lose innumerable data of elements Defined. Through the combination of highlight, the FPN algorithm effectively combines semantic data in rich in the deep layer of the organization with data from rich components in the organization's surface layer to work on the impact of the recognition of small items. [17] proposed a more developed SSD calculation DSSD, which used RESNET101 as the skeletal organization, and used the possibility that FPN identifies the protests through the semantic data interface of several layers and multi-layered brand data melting . [18] Use PGGAN to further develop the rate of identification of small articles. The impact of discovery of small articles is deficient, while the enormous items are acceptable. Primary thinking is to use the PGGAN organization to allow The generator to familiarize himself with a planning, plan the elements of the small article to the elements of the huge article, and then further develop the impact of identification of the small articles.

In 2018, [19] proposed SNIP, focal thinking is to do as close as conceivable for the information broadcasting model and preparation of dispersion. First, SNIP reticed on ImageNet, and then modified the image acquired from the first image with the calibrated model to foresee the image after the exam. The impact of the update of this technique is very amazing. [20] It was proposed that the CNN R-CNN multiple object location system proposed to take care of the problem that the impact of identification was weakened with the increase of the IOU edge. R-CNN course consists of a progression of several IOU EDGE search engines, with more options for the detection box to approach false positives. The same waterfall applies to reasoning, assuming a closest coincidence between the dough of the detector at each stage. Redmon [9] proposed the Yolov3 improved model of Yolo9000. Yolov3 uses the DarkNet53 network as the characteristic extraction network, and the DARKNET-53 network borrows the idea of Renet and adds residual module. Effectively resolve the problem of gradient explosion and disappearance of gradients in a

deep network.

In 2019, [21] he proposed a model of calculation of small bird articles that depend on a deep convolutional network, which is to address the issue of the location of small articles in the huge area of foundation. This technique consolidates the calculation of discovery of articles with the semantic division, prepares a deep convolutionary neural organization (FCN) and FCN variations, and incorporates the results through the Help Machine (SVM), with the purpose of Perform the execution of the identification of the small increase in the article. [14] He proposed a network of depth agency of the square region at the pixel level to understand the grouping of a single pixel in ether images and configure the land cover guide to understand the identification of small elements. [15] proposed a technique to eliminate the provincial pixel elements by pixel, and use the convoluted neural organization to characterize and recognize the most outstanding aspects. I know that, as can, the productivity of this strategy is excessively low, and the multiple size of the article is not versatile.

In 2017, [16] proposed a characteristic pyramid organization (FPN) with a level association, using multi-scale feature data to achieve the location of the small location of the article. The calculations of recognition of previous objects only use the elements of the upper layer of the organization for the discovery, although it contains rich semantic data, however, after the endless supply of the grouping activity will lose innumerable data of elements Defined. Through the combination of highlight, the FPN algorithm effectively combines semantic data in rich in the deep layer of the organization with data from rich components in the organization's surface layer to work on the impact of the recognition of small items. [17] proposed an additional developed Calculation of SSD DSSD, which used RESNET101 as the skeletal organization, and used the possibility that FPN identifies the protests through the multi-layer semantic data interface and the commercial brand data of several layers. [18] Use PGGAN to further develop the rate of identification of small articles. The impact of discovery of small articles is deficient, while the enormous items are acceptable. Primary thinking is to use the PGGAN organization to allow the generator to become familiar with planning, plan the elements of the small article to the elements of the huge article, and then further develop the impact of identification of small articles.

In 2018, [19] proposed SNIP, focal thinking is to do as close as conceivable for the information broadcasting model and preparation of dispersion. First, SNIP restricted on ImageNet, and then modified the image acquired from the first image with the calibrated model to foresee the image after the exam. The impact of the update of this technique is very amazing. [20] It was proposed that the CNN R-CNN multiple object location system proposed to take care of the problem that the impact of identification was weakened with the increase of the IOU edge. Course R-CNN consists of a progression of various search engines of ISOU.

IV. SUMMARY AND LEAFLET

Since Alex Net resulted in 2012, the remarkable calculation achievements of the convolutional neuronal organization in the field of PC vision have established the increase in deep learning research. The impact of the location of the calculation of the identification of the item that depends on deep learning is a long way that beyond the customary calculus, which replaces the conventional calculation and grows rapidly. After a great performance of advance and improvement, its organizational structure has been continuously improved, the accuracy and speed of the location have been further improved, and the common, military and different angles have been effectively applied.

The calculation of the article that depends on a deep learning has been improved persistently and its organization structure has been improved. Fundamentally calculation groups identification of a stage and calculation of two-stage discovery, among which the calculation of a stage enjoys more benefits in the recognition speed and calculation of two stages enjoys more benefits in the accuracy of the location . With the extensive knowledge of knowledge, situations of application of the identification of the article extend further. As a significant article in situations of complex applications, small objectives have been constantly a problem in the field of the discovery of articles due to its limited scale, under target, absence of data from components and different

variables. From now on, most of the small calculations of identification of articles are optimized and improved the discovery structure of traditional professional learning objects. To further develop discovery accuracy, the location speed and balance have become the focal point of the small research identification research. Likewise, to further develop the impact of identification of small objectives, numerous organizations have given public information collections. to little objectives.

V. CONCLUSION

Object identification is considered as principal step in arrangement of self-driving vehicles and advanced mechanics. In this paper, we demystified the job of profound learning procedures in light of CNN for object identification. Profound learning systems and administrations accessible for object location are additionally talked about in the paper. Benchmarked datasets for object limitation what's more recognition delivered in overall rivalries are additionally covered. The pointers to the areas wherein object discovery is appropriate has been talked about. Best in class profound learning-based item recognition procedures have been evaluated and analyze

REFERENCES

[1] Y. Lecun, L. Bottou, Y. Bengio, P. Haffner, "Gradient-based learning applied to document recognition," Proceedings of the IEEE, vol.86, no.11, pp.2278-2324, 1998.

[2] A. Krizhevsky, I. Sutskever, G. Hinton, "ImageNet classification withdeep convolutional neural networks" International Conference on Neural Information Processing Systems, pp.1097-1105, August 2012.

[3] R. Girshick, J. Donahue, T. Darrell, J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," Conference on Computer Vision and Pattern Recognition, pp.580-587,May 2014.

[4] K. He, X. Zhang, S. Ren, J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," IEEE Transactions on Pattern Analysis & Machine Intelligence, vol.37, no.9, pp.1904-1916,2015.

[5] R. Girshick, "Fast R-CNN," Computer Science, vol.18, no.3, pp.5-7, 2015.

[6] S. Ren, K. He, R. Girshick, J. Sun, "Faster R-CNN: towards real-time object detection with region proposal networks," IEEE Trans Pattern Anal Mach Intell, vol.39, no. 6, pp.1137-1149, 2015.

[7] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, "You only look once: unified, real-time object detection," Conference on Computer Vision and Pattern Recognition, pp.779-788, May 2016.

[8] J. Redmon, A. Farhadi. "YOLO9000: better, faster, stronger," Conference on Computer Vision and Pattern Recognition, pp.6517- 6525, January 2016.

[9] J. Redmon, A. Farhadi. "YOLOv3: an incremental improvement," Computer Vision and Pattern Recognition, pp.1-6, January 2018.

[10] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C. Fu, "SSD: single shot multibox detector. European Conference on Computer Vision," pp.21-37, January 2016.

[11] S. Bell, C. Zitnick, K. Bala, R. Girshick, "Inside-outside net: detecting objects in context with skip pooling and recurrent neural networks," IEEE Conference on Computer Vision and Pattern Recognition, pp.2874-2883, June 2015.

[12] M. Topi, P. Matti, "Texture analysis with local binary patterns," Handbook of Pattern Recognition & Computer Vision, vol.3540, pp.115-118, 2005.

[13] Takeki A, T.Trinh, R. Yoshihashi, R. Kawakami, M. Iida, T. Naemura, "Combining deep features for object detection at various scales: finding small birds in landscape images," IPSJ Transactions on Computer Vision and Applications, vol.8, no.1, pp.1-5, 2016.

[14] M. Kampffmeyer, A. Salberg, R. Jenssen, "Semantic segmentation of small objects and modeling of uncertainty in urban remote sensing images using deep convolutional neural networks," 2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp.680-688, June 2016.

[15] T. Mundhenk, G. Konjevod, W. Sakla, K. Boakye, "A large contextual dataset for

classification, detection and counting of cars with deep learning" ECCV 2016, pp785-800. Sep 2016.

[16] T. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, S. Belongie, "Feature pyramid networks for object detection," Conference on Computer Vision and Pattern Recognition, pp.936-944, Aug 2017.

[17] C. Fu, W. Liu, A. Ranga, A. Tyagi, A. Berg, "DSSD: Deconvolutional single shot detector," Computer Vision and Pattern Recognition, pp.12-15, Aug 2017.

[18] J. Li, X. Liang, Y. Wei, T. Xu, J. Feng, S. Yan, "Perceptual Generative Adversarial Networks for Small Object Detection," Conference on Computer Vision and Pattern Recognition, pp.1222- 1230, Aug 2017.

[19] B. Singh, L. Davis, "An Analysis of Scale Invariance in Object Detection – SNIP," Computer Vision and Pattern Recognition, pp.3578-3587, Oct 2018.