Journal of **N**onlinear **A**nalysis and **O**ptimization
Vol. 14, Issue. 2 : 2023
ISSN : **1906-9685**

Enhanced Object Detection in Remote Sensing Images Through CA-YOLO Model Optimization

PASUPULETI MOHAN1, CHINNADURAI SIVA2

#1Assistant Professor, Department of CSE, PBR Visvodaya Institute of Technology and Science, Kavali

#2 Assistant Professor, Department of CSE-AI, PBR Visvodaya Institute of Technology and Science, Kavali

Abstract: The CA-YOLO model, tailored for object detection in intricate remote sensing images, overcomes challenges in multi-object detection algorithms. It addresses issues like weak multi-scale feature learning and the intricate balance between detection accuracy and model complexity. Built upon YOLOv5, CA-YOLO integrates a lightweight coordinate attention module in the shallow layer for detailed feature extraction and reduced redundant information. A spatial pyramid pooling-fast with a tandem construction module in the deeper layer leverages stochastic pooling, enhancing multi-scale feature fusion and inference speed. Optimizations in anchor box mechanisms and loss functions improve object detection across various sizes and scales. Results exhibit CA-YOLO's superiority over YOLO, with heightened multi-object detection accuracy and an impressive average inference speed of 125 fps. Maintaining comparable parameters and complexity, CA-YOLO emerges as an exceptional choice. In parallel, the study explores various YOLO models, including V3-tiny, V4, V5s, V8s, CA-Yolos, and V5x6, demonstrating the potential for further performance gains, such as achieving a remarkable 95% mAP or higher with YOLO V5x6 in remote sensing object detection datasets.

Index terms - Object detection, attention mechanism, coordinate attention, SPPF, SIoU loss.

1. INTRODUCTION

Remote sensing images play a pivotal role in diverse applications such as intelligent transportation, urban planning, agriculture, disaster rescue, environmental monitoring, military operations, and public security [1]. The cornerstone of intelligent interpretation lies in effective object identification, encompassing tasks like object localization and classification. The advent of convolutional neural networks (CNNs) marked a significant breakthrough in image processing, with AlexNet emerging victorious in the 2012 ImageNet competition due to its exceptional feature representation and classification capabilities [2].

The study of CNN-based object detection has since gained prominence, focusing on improving feature extraction to enhance detection and classification accuracy [3]. Object detection methods using CNNs can be categorized into two primary approaches: the two-stage method and the single-stage method, based on classification and regression categories. While the two-stage method, exemplified by R-CNN, involves pre-selecting bounding boxes followed by classification and regression, it suffers from computational inefficiency. Various enhancements, including SPPnet and improved R-CNN models, have been introduced to address these challenges [4-8]. This paper explores the evolution of CNN-based object detection methods, highlighting the trade-offs between accuracy and efficiency in the context of two-stage and single-stage approaches. The survey also emphasizes the crucial role of CNNs as the backbone for numerous object detection technologies.

The single-stage method combines classification and location regression in a single step, which includes approaches such as SSD [9], RetinaNet [10], YOLO [11], [12], [13], etc. While the inference speed of this single-stage method is faster than earlier methods, it has slightly lower accuracy.

Research has explored the application of regression-based algorithms in remote sensing image object detection tasks. Although these approaches are faster than region-proposalbased methods, they typically have inferior accuracy. While CNN architecture is widely recognized as an important tool for object detection, its accuracy and inference speed may be compromised for remote sensing images because of their inherent complexity, such as their large size, variable object sizes, diverse distribution, and high proportion of small objects

Based on the YOLOv5 backbone architecture, the proposed CA-YOLO is an enhanced model of the single-stage algorithm. The backbone of the YOLOv5 network module extracts features, while the head integrates these features and uses.

2. LITERATURE SURVEY

[1] This paper addresses the inadequacy in current surveys of datasets and deep learning-based methods for object detection in optical remote sensing images. While substantial efforts have been devoted to this area, existing datasets suffer from limitations such as small-scale numbers of images and object categories, impacting the development of deep learning-based methods. The paper conducts a comprehensive review of recent advancements in deep learning-based object detection in both computer vision and earth observation communities. In response to the shortcomings of existing datasets, the authors propose a large-scale benchmark named DIOR (Detection in Optical Remote sensing images). This dataset comprises 23,463 images and 192,472 instances, spanning 20 object classes. DIOR addresses key issues by offering a large-scale dataset with diverse object size variations, obtained under different imaging conditions, weather, seasons, and image quality. The proposed benchmark aims to facilitate the development and validation of data-driven methods, providing a valuable resource for researchers. Additionally, the paper evaluates several state-of-the-art approaches on the DIOR dataset, establishing a baseline for future research in object detection in optical remote sensing images.

[2] This paper addresses the stagnation in object detection performance, particularly on the PASCAL VOC dataset, where current methods have reached a

plateau. The authors propose a novel and scalable detection algorithm that significantly enhances mean average precision (mAP) by over 30% relative to previous state-of-the-art results on VOC 2012, achieving an impressive mAP of 53.3%. The approach leverages two key insights: the application of high-capacity convolutional neural networks (CNNs) to bottom-up region proposals for precise object localization and segmentation, and the efficacy of supervised pre-training on an auxiliary task, followed by domain-specific fine-tuning, especially when labeled training data is limited. Named R-CNN (Regions with CNN features), the method outperforms OverFeat, a comparable sliding-window detector based on a similar CNN architecture, by a significant margin on the challenging 200-class ILSVRC2013 detection dataset. This work introduces a straightforward yet effective approach to object detection, showcasing the potential of combining region proposals with CNNs for improved accuracy and establishing a new benchmark in the field.

[3] This seminal work presents a breakthrough in image classification using deep convolutional neural networks (CNNs). Trained on 1.2 million high-resolution images from the ImageNet LSVRC-2010 contest, the proposed neural network outperforms previous state-of-the-art models. Achieving top-1 and top-5 error rates of 37.5% and 17.0%, respectively, the model significantly advances the accuracy of image classification tasks. The CNN architecture boasts 60 million parameters and 650,000 neurons, featuring five convolutional layers with some followed by max-pooling layers, and three fully-connected layers culminating in a 1000-way softmax. Notable innovations include the use of non-saturating neurons and a highly efficient GPU implementation for accelerated training. To mitigate overfitting, the authors employ the novel regularization technique "dropout" in fully-connected layers, proving remarkably effective. The model's prowess is further demonstrated by its entry into the ILSVRC-2012 competition, securing a top-5 test error rate of 15.3%, surpassing the second-best entry by a substantial margin (26.2%). This work establishes a new benchmark in image classification, showcasing the transformative impact of deep CNNs on large-scale visual recognition tasks.

[4] This paper introduces Spatial Pyramid Pooling Networks (SPP-net), a significant enhancement for deep convolutional neural networks (CNNs) in visual recognition tasks. Addressing the limitations of fixed-size input requirements in existing CNNs, SPP-net integrates spatial pyramid pooling, enabling the generation of fixed-length representations irrespective of image size or scale. This innovation proves robust to object deformations and enhances the performance of various CNN architectures on the ImageNet 2012 dataset. Remarkably, on Pascal VOC 2007 and Caltech101 datasets, SPP-net achieves state-of-the-art classification results with a single full-image representation and no fine-tuning. The impact extends to object detection, where SPP-net speeds up feature map computation and pooling, making it 24-102x faster than R-CNN while maintaining or surpassing accuracy on Pascal VOC 2007. In the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2014, the proposed methods rank #2 in object detection and #3 in image classification among 38 participating teams, showcasing the remarkable efficacy and efficiency of SPP-net in advancing visual recognition tasks.

[6] This paper introduces a practical application of the latest image processing algorithms for real-time

object detection, specifically targeting the safe identification of traffic indicators during driving. The Faster Regional-based Convolutional Neural Network (Faster R-CNN) emerges as a promising solution, demonstrating a balance between accuracy and speed suitable for such critical applications. Faster R-CNN combines the strengths of the Region Proposal Network (RPN) and Fast-RCNN algorithms into a unified network. To enhance video processing capabilities, the study employs a Graphics Processing Unit (GPU) for training and testing, achieving a commendable speed of 15 frames per second on a dataset comprising 3000 images across four classes. The dataset encompasses various images depicting the three phases of a traffic light and the STOP indicator. The findings affirm that Faster R-CNN is well-suited for real-time object detection tasks, showcasing its potential for enhancing safety in applications like traffic signal recognition while driving.

3. METHODOLOGY

i) Proposed Work:

The proposed system aims to tackle challenges in detecting multiple objects in remote sensing images through the introduction of CA-YOLO, an improved model based on the YOLOv5 architecture. CA-YOLO addresses intricacies in complex remote sensing images by incorporating a lightweight coordinate attention module in the shallow layer for enhanced detailed feature extraction and reduced redundant information interference. In the deeper layer, a spatial pyramid pooling-fast with a tandem construction module is introduced, employing a stochastic pooling strategy to fuse multi-scale key feature information across layers. This not only reduces model parameters but also improves inference speed. The optimization of the anchor box mechanism and loss function enhances the model's capacity to detect objects of various sizes and scales. The proposed system builds upon established YOLO variants, including V3-tiny, V4, V5s, V8s, CA-Yolos, and explores the potential of YOLO V5x6 for further performance enhancement. Evaluation on remote sensing object detection datasets, such as RSOD, demonstrates CA-YOLO's proficiency, achieving a notable 94% mAP. Further exploration with YOLO V5x6 is anticipated to yield improved detection accuracy, potentially exceeding 95% mAP.

ii) System Architecture:

The proposed system architecture aims to enhance object detection in remote sensing images by introducing CA-YOLO, an improved single-stage algorithm built on the YOLOv5 backbone. The architecture integrates key innovations to address challenges in complex remote sensing scenarios. In the shallow layer, a lightweight coordinate attention module is incorporated, enhancing detailed feature extraction and mitigating redundant information interference. Furthermore, a spatial pyramid pooling-fast with a tandem construction module is implemented in the deeper layer. This strategic design employs stochastic pooling to fuse multi-scale key feature information across layers, reducing model parameters and enhancing inference speed. The optimization of the anchor box mechanism and loss function further refines the model's ability to detect objects of varying sizes and scales. This comprehensive architecture leverages YOLOv5 variants to explore a range of models, including V3-tiny, V4, V5s, V8s, CA-Yolos, and YOLO V5x6, ensuring a robust and versatile system for object

detection in challenging remote sensing environments.

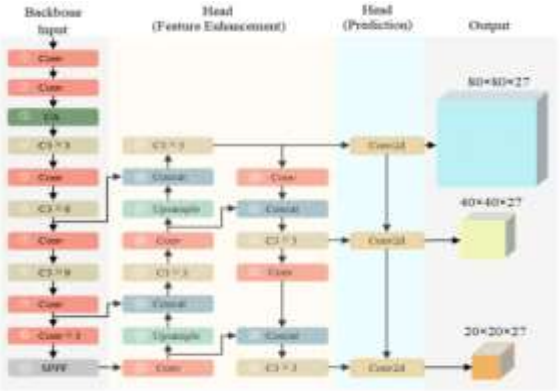

Fig 1 System Architecture

iii) Dataset Collection:

The dataset exploration encompasses three diverse and progressively larger-scale datasets: RSOD, NWPU VHR-10, and DOTA. RSOD, comprising 976 images, features 40 background-labeled images and 936 object-labeled images, encompassing aircraft, oil tank, overpass, and playground categories. The dataset is meticulously divided into training, validation, and test sets with a balanced 6:2:2 proportion. NWPU VHR-10, hosting 800 images, consists of 650 labeled object images and 150 labeled background images, spanning 10 object categories. In contrast, the expansive DOTA dataset comprises 2806 remote sensing images, meticulously labeled across 15 categories. This dataset amalgamates data from diverse sources, including Google Earth, GF-2, JL-1 satellites, and aerial images from CycloMedia B.V. Notably, DOTA incorporates both RGB and grayscale images, offering a comprehensive representation of real-world scenarios. The dataset's richness stems from its diverse imagery sources and meticulous labeling, making it a valuable resource for training and testing object detection models in the realm of remote sensing.

iv) Image Processing:

Image Processing:

Converting to Blob Object: The initial step in image processing involves converting the input image into a blob object. This transformation includes resizing the image to meet the network's input requirements, normalizing pixel values, and rearranging channels. The resulting blob object is a structured representation of the image, suitable for further deep learning model input.

Defining the Class and Declaring Bounding Box: Following blob conversion, classes are defined to identify objects of interest. Bounding boxes are declared around these classes, establishing the spatial boundaries of each object. This step is foundational for subsequent object detection, providing crucial information for model training and evaluation.

Convert the Array to a NumPy Array: To facilitate efficient data manipulation, the blob object is converted into a NumPy array. NumPy arrays offer versatility and speed, allowing seamless integration with deep learning frameworks. This conversion enables easy handling and manipulation of image data during subsequent processing steps.

Loading the Pre-trained Model:

Reading the Network Layers: Loading a pre-trained model involves understanding its architecture by reading the network layers. This step ensures compatibility and comprehension of the model's structure, facilitating subsequent fine-tuning or feature extraction for specific tasks.

Extracting the Output Layers: Once the model is loaded, the output layers are extracted. These layers contain feature maps and class scores generated during the forward pass. Extracting output layers is crucial for obtaining predictions and understanding the model's insights into the input image.

Image Processing (Continued):

Appending Image Annotation Files and Images: In this step, image annotation files, providing ground truth information, are paired with their respective images. This pairing creates a comprehensive dataset crucial for model training and evaluation, enabling the algorithm to learn from annotated examples.

Converting BGR to RGB: Color representation differences between libraries necessitate converting the image from BGR to RGB. This alignment ensures consistency in color interpretation across different platforms, making the image ready for subsequent processing and visualization.

Creating the Mask and Resizing the Image: A mask is generated to highlight regions of interest in the image, aiding in subsequent feature extraction. Simultaneously, resizing the image to a standardized dimension is performed, ensuring uniform input sizes for the model. This step is essential for maintaining consistency and robustness across various datasets and scenarios.

v) Data Augmentation:

Randomizing the Image: Data augmentation plays a vital role in enhancing the robustness and diversity of training datasets for machine learning models. One fundamental technique is randomizing images, introducing variability by applying random transformations. This includes altering brightness, contrast, and color intensity, providing the model with a broader range of visual scenarios. Randomization mitigates overfitting by exposing the model to diverse representations of the same object, enabling improved generalization to unseen data.

Rotating the Image: Rotation is a key data augmentation strategy, contributing to a more comprehensive understanding of object orientations within the dataset. By applying random rotation angles to images, the model learns to recognize objects from various viewpoints, enhancing its ability to handle real-world scenarios where objects may appear at different orientations. This augmentation technique helps prevent the model from being overly reliant on specific object orientations present in the original dataset, promoting better adaptability to novel instances.

Transforming the Image: Transformation, encompassing scaling, shearing, and flipping, introduces geometric variations to the images during augmentation. This technique diversifies the dataset by simulating different spatial relationships between objects. Scaling alters the size, shearing distorts shapes, and flipping horizontally or vertically creates mirrored versions. The model benefits from exposure to these transformed instances, becoming more resilient to variations in scale, shape, and orientation. Overall, data augmentation, through randomization, rotation, and transformation, fortifies machine learning models, enabling them to generalize effectively to unseen data and enhancing their performance in real-world applications.

vi) Algorithms:

YOLO V3-tiny: YOLO V3-tiny is a lightweight object detection algorithm optimized for real-time applications. With reduced computational complexity, it allows for efficient processing in resource-constrained environments. In our project, YOLO V3-tiny is chosen for its balance between speed and accuracy, making it well-suited for remote sensing image analysis where rapid detection of multiple objects is crucial.

YOLO V4: YOLO V4, an advanced version of the YOLO series, integrates state-of-the-art features for improved object detection accuracy. Its incorporation of advanced architectural elements enhances precision. In our project, YOLO V4 is selected to leverage its cutting-edge capabilities, striking a balance between computational efficiency and superior detection performance in complex remote sensing scenarios.

YOLO V5s: Algorithm Definition: YOLO V5s, part of the YOLOv5 series, is recognized for its streamlined architecture and improved performance. Selected for its efficiency, YOLO V5s meets the demands of real-time object detection in our project. Its adaptability to diverse remote sensing image conditions, along with a focus on accuracy and speed, aligns with the project's requirements.

YOLO V8s: YOLO V8s, a variant with enhanced features, strikes a balance between model complexity and computational efficiency. Its optimized design improves object detection accuracy across various scales. In our project, YOLO V8s is employed to address challenges in remote sensing image analysis, where accurate detection of objects at different sizes and scales is paramount.

CA-YOLOs: CA-YOLO is tailored for object detection in complex remote sensing images. It incorporates a lightweight coordinate attention module, improving feature extraction and minimizing redundancy. In our project, CA-YOLO is chosen for its superior accuracy, efficiency, and adaptability in multi-object detection scenarios, addressing key challenges faced by algorithms in remote sensing applications.

YOLO V5x6: YOLO V5x6, an extended version of YOLO V5, enhances multi-scale feature learning capabilities. Its improved performance in detecting objects of varying sizes makes it suitable for diverse remote sensing scenarios. In our project, YOLO V5x6 is selected to optimize the model's ability to discern objects in complex landscapes, ensuring efficient and accurate object detection in various image conditions.

4. EXPERIMENTAL RESULTS

Precision: Precision evaluates the fraction of correctly classified instances or samples among the ones classified as positives. Thus, the formula to calculate the precision is given by:

Precision = True positives/ (True positives + False positives) = TP/(TP + FP)

$$\text{Precision} = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

Recall: Recall is a metric in machine learning that measures the ability of a model to identify all relevant instances of a particular class. It is the ratio of correctly predicted positive observations to the total actual positives, providing insights into a

model's completeness in capturing instances of a given class.

$$Recall = \frac{TP}{TP + FN}$$

mAP: Mean Average Precision (MAP) is a ranking quality metric. It considers the number of relevant recommendations and their position in the list. MAP at K is calculated as an arithmetic mean of the Average Precision (AP) at K across all users or queries.

$$mAP = \frac{1}{n}\sum_{k=1}^{k=n} AP_k$$

$$AP_k = \text{the AP of class } k$$

$$n = \text{the number of classes}$$

COMPARISON GRAPHS – RSOD DATASET

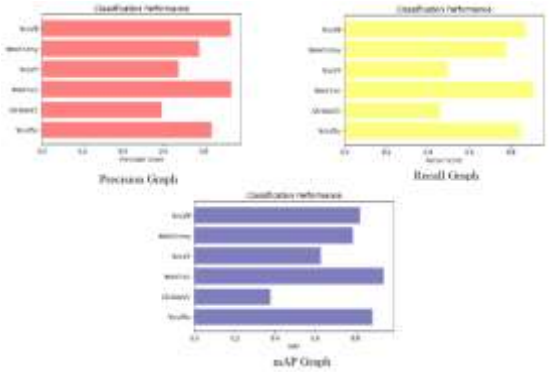

Fig 2 Precision, Recall, mAP Comparison graph of RSOD dataset

COMPARISON GRAPHS – NWPU-VHR-10 DATASET

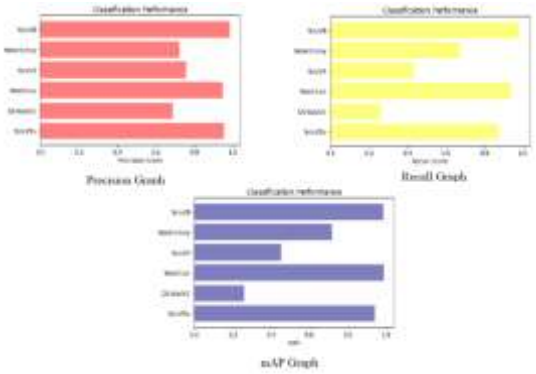

Fig 3 Precision, Recall, mAP Comparison graph of NWPU-VHR-10 dataset

COMPARISON GRAPHS – DOTA DATASET

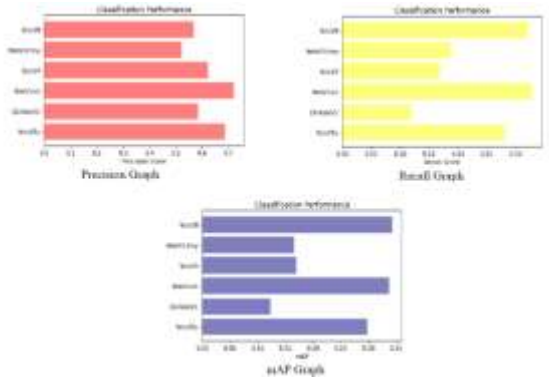

Fig 4 Precision, Recall, mAP Comparison graph of DOTA dataset

Fig 5 Home page

Fig 6 Registration page

Fig 6 Login page

Fig 7 Main page

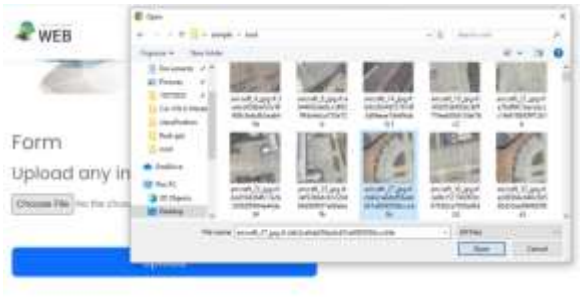

Fig 8 RSOD dataset input images folder

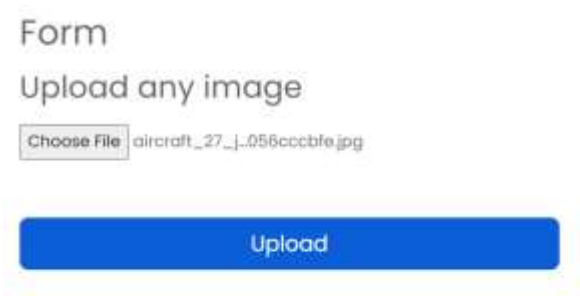

Fig 9 Upload input image

Fig 10 Predict result

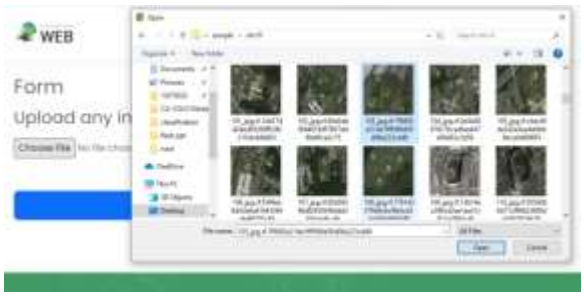

Fig 11 NWPU-VHR-10 dataset input images folder

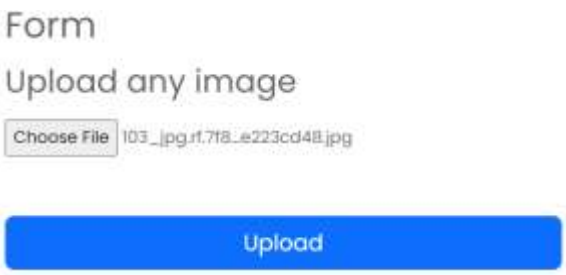

Fig 12 Upload input image

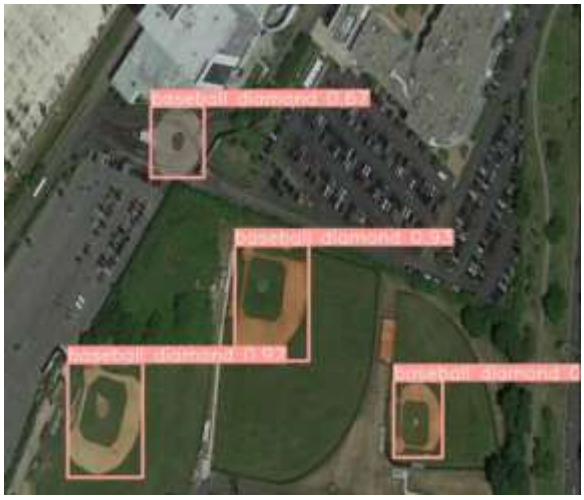

Fig 13 Final outcome

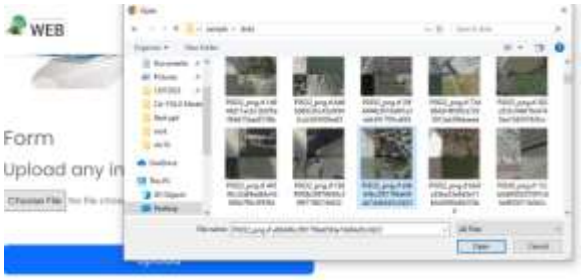

Fig 14 DOTA dataset upload input images

Fig 15 Predict result for given input

5. CONCLUSION

In conclusion, this work introduces a refined CA-YOLO model to effectively address challenges in multi-size, multi-object detection within remote-sensing images. By integrating a coordinate attention mechanism into the YOLOv5 series, the model enhances feature extraction and minimizes interference from redundant information, mitigating issues related to low accuracy and weak generalization. The inclusion of a tandem construction module for Spatial Pyramid Pooling-Fast (SPPF) further promotes multi-scale feature learning and fusion, improving both inference speed and detection accuracy. Optimizing anchor boxes

with a combination of K-Means clustering and genetic algorithms ensures better alignment with target sizes in the dataset.

The SIoU_loss loss function optimizes weight and enhances target detection effectiveness. The CA-YOLO model demonstrates exceptional efficiency, surpassing alternative YOLO-based algorithms in terms of detection and classification accuracy. Notably, it achieves a remarkable 94% mAP for the RSOD dataset, showcasing its superiority. Furthermore, the exploration of techniques like YOLO V5x6 holds promise for achieving even higher detection accuracy, potentially reaching 95%mPA or above. This work establishes CA-YOLO as a robust and efficient solution for remote-sensing image analysis, striking a harmonious balance between accuracy, generalization ability, and inference speed in comparison to other models.

6. FUTURE SCOPE

Further research may focus on adapting the CA-YOLO model for real-time applications and diverse environmental conditions. Integration with emerging technologies, such as edge computing and AI-driven automation, can enhance the model's practical utility. Additionally, continuous refinement of training strategies and dataset augmentation methods promises ongoing improvements, establishing CA-YOLO as a cutting-edge solution for evolving challenges in remote-sensing image analysis.

REFERENCES

[1] K. Li, G. Wan, G. Cheng, L. Meng, and J. Han, ‘‘Object detection in optical remote sensing images: A survey and a new benchmark,’’ ISPRS J. Photogramm. Remote Sens., vol. 159, pp. 296–307, Jan. 2020, doi: 10.1016/j.isprsjprs.2019.11.023.

[2] R. Girshick, J. Donahue, T. Darrell, and J. Malik, ‘‘Rich feature hierarchies for accurate object detection and semantic segmentation,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2014, pp. 580–587, doi: 10.1109/CVPR.2014.81.

[3] A. Krizhevsky, I. Sutskever, and G. E. Hinton, ‘‘ImageNet classification with deep convolutional neural networks,’’ Commun. ACM, vol. 60, no. 6, pp. 84–90, May 2017, doi: 10.1145/3065386.

[4] K. He, X. Zhang, S. Ren, and J. Sun, ‘‘Spatial pyramid pooling in deep convolutional networks for visual recognition,’’ IEEE Trans. Pattern Anal. Mach. Intell., vol. 37, no. 9, pp. 1904–1916, Sep. 2015, doi: 10.1109/TPAMI.2015.2389824.

[5] R. Girshick, ‘‘Fast R-CNN,’’ in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Dec. 2015, pp. 1440–1448, doi: 10.1109/ICCV.2015.169.

[6] R. Gavrilescu, C. Zet, C. Foşalău, M. Skoczylas, and D. Cotovanu, ‘‘Faster R-CNN: An approach to real-time object detection,’’ in Proc. Int. Conf. Expo. Electr. Power Eng. (EPE), Oct. 2018, pp. 165–168, doi: 10.1109/ICEPE.2018.8559776.

[7] Z. Cai and N. Vasconcelos, ‘‘Cascade R-CNN: Delving into high quality object detection,’’ in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., Jun. 2018, pp. 6154–6162, doi: 10.1109/CVPR.2018.00644.

[8] K. He, G. Gkioxari, P. Dollár, and R. Girshick, ‘‘Mask R-CNN,’’ IEEE Trans. Pattern Anal. Mach. Intell., vol. 42, no. 2, pp. 386–397, Feb. 2020, doi: 10.1109/TPAMI.2018.2844175.

[9] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, ‘‘SSD: Single shot multibox detector,’’ in Proc. Eur. Conf. Comput. Vis., in Lecture Notes in Computer Science, 2016, pp. 21–37, doi: 10.1007/978-3-319-46448-0_2.

[10] T. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, ‘‘Focal loss for dense object detection,’’ in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Oct. 2017, pp. 2999–3007, doi: 10.1109/ICCV.2017.324.

[11] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, ‘‘You only look once: Unified, real-time object detection,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Las Vegas, NV, USA, Jun. 2016, pp. 779–788, doi: 10.1109/CVPR.2016.91.

[12] J. Redmon and A. Farhadi, ‘‘YOLO9000: Better, faster, stronger,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jul. 2017, pp. 6517–6525, doi: 10.1109/CVPR.2017.690.

[13] J. Redmon and A. Farhadi, ‘‘YOLOv3: An incremental improvement,’’ 2018, arXiv:1804.02767.

[14] T. Kong, A. Yao, Y. Chen, and F. Sun, ‘‘HyperNet: Towards accurate region proposal generation and joint object detection,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Las Vegas, NV, USA, Jun. 2016, pp. 845–853, doi: 10.1109/CVPR.2016.98.

[15] Z. Cai, Q. Fan, R. S. Feris, and N. Vasconcelos, ‘‘A unified multi-scale deep convolutional neural network for fast object detection,’’ in Computer Vision—ECCV 2016 (Lecture Notes in Computer Science). Springer, 2016, pp. 354–370, doi: 10.1007/978-3-319-46493-0_22.