Kaiwen Duan, Song Bai et al., [1], CenterNet, a keypoint-based object detection framework that addresses the problem of incorrect object bounding boxes, is introduced in this paper. CenterNet improves precision and recall by detecting objects as triplets of keypoints and employing custom pooling modules. CenterNet achieves an AP of 47.0% on the MS-COCO dataset, surpassing existing one-stage detectors by at least 4.9 percent. It also performs similarly to the best two-stage detectors while maintaining a faster inference rate. CenterNet's inventive methodology and promising outcomes make it a significant commitment to the field of article identification.

Zuo-Xin Li, Fu-Qiang Zhou., [2], The paper introduces the enhanced SSD object detection algorithm known as FSSD (Feature Fusion Single Shot Multibox Detector). FSSD proposes a lightweight feature fusion module that significantly improves performance without sacrificing speed to address SSD's feature pyramid detection limitation. FSSD produces a novel feature pyramid by joining features from various layers with varying scales together. On the Pascal VOC 2007 test, FSSD accomplishes 82.7 Guide at 65.8 FPS utilizing a solitary Nvidia 1080Ti GPU. It beats customary SSD and many cutting edge object identification calculations regarding both exactness and speed.

Kye-Hyeon Kim et al., [3], The paper proposes a novel strategy for detecting multi-category objects with the highest possible accuracy at the lowest possible computational cost. The network achieves impressive results on well-known benchmarks for object detection by redesigning the feature extraction portion of the pipeline and employing methods like concatenated ReLU, Inception, and HyperNet. The organization's profound and slender engineering, joined with clump standardization, lingering associations, and learning rate planning, adds to its unrivaled exhibition. Surprisingly, the proposed network accomplishes equivalent outcomes to ResNet-101 with just 12.3% of the computational expense, making it exceptionally productive and successful for object discovery assignments.

Songtao Liu and Yunhong Wang [4], The paper tends to the compromise among precision and computational expense in object location models. To improve feature discriminability and robustness, it proposes a novel RF Block (RFB) module based on the structure of RFs in human visual systems. The proposed RFB Net detector maintains real-time processing speed while maintaining performance comparable to that of advanced deep detectors by incorporating RFB modules into the SSD framework. The RFB Net is a promising approach to achieving a balance between accuracy and efficiency in object detection tasks, as demonstrated by the experiments carried out on major benchmarks.

Pieze Sun, Rufeng Zhang et al., [5], An entirely sparse method for image object detection is introduced by Sparse R-CNN. Sparse R-CNN uses a fixed set of learned object proposals, which eliminates the need for manually designed candidates and many-to-one label assignment, in contrast to other methods that use dense object candidates. On the challenging COCO dataset, the proposed method achieves comparable accuracy, runtime, and training convergence performance to established detectors. Sparse R-CNN challenges the use of dense priors in object detection and makes a valuable contribution to the field at a significantly lower computational cost.

Mingxing Tan et al., [6], A weighted bi-directional feature pyramid network (BiFPN) and compound scaling are among the optimizations suggested in this paper to address the requirement for effective object detection models. The resulting EfficientDet models, which are based on EfficientNet backbones, outperform previous detectors in terms of performance while being significantly smaller and requiring fewer computational resources. With amazing outcomes on COCO test-dev, EfficientDet shows the viability of the proposed enhancements for further developing effectiveness in object location undertakings.

Zhi Tian et al., [7], The paper presents FCOS, a completely convolutional one-stage object locator that works in a for each pixel expectation design. Not at all like different indicators depending on pre-characterized anchor boxes, FCOS is anchor without box and proposition free, working on the identification interaction. FCOS improves detection accuracy by eliminating the need for anchor box computations and hyper-parameter tuning. FCOS surpasses previous one-stage detectors with an impressive AP of 44.7% in single-model and single-scale testing with ResNeXt-64x4d-101. A significant advancement in object detection techniques, FCOS is a powerful alternative for a variety of instance-level tasks because it provides a simpler and more adaptable detection framework.

Jospeh Redmon and Ali Farhadi, [8], The paper introduces modifications to YOLOv3, resulting in enhancements to the design and training of a more precise network. The updated YOLOv3 continues to be fast, running at 22 milliseconds at 320 x 320 resolution despite its slight increase in size. It accomplishes tantamount exactness to SSD yet is multiple times quicker.

YOLOv3 outperforms RetinaNet in terms of the.5 IOU mAP detection inference times, making it a promising option for object detection tasks.

Alexey Bochkovskiy et al., [9], The paper calls for practical testing and theoretical justification of the effectiveness of features that improve the accuracy of Convolutional Neural Networks (CNNs). Weighted residual connections (WRC), Cross stage partial connections (CSP), Cross mini-Batch Normalization (CmBN), Self-adversarial training (SAT), and Mish activation are examples of universal features that can be applied to a wide range of models, tasks, and datasets. On Tesla V100, the proposed method achieves a real-time speed of 65 frames per second in real time and an impressive AP of 43.5% (65.7% AP50). This demonstrates that these features are useful for enhancing object detection performance.

Xingkui Zhu et al., [10], TPH-YOLOv5 introduces a new prediction head for different-scale objects, employs Transformer Prediction Heads (TPH) with self-attention, and integrates convolutional block attention model (CBAM) for dense object scenarios to address the difficulties of object detection in drone-captured scenarios. On the DET-test-challenge dataset, it outperforms previous cutting-edge methods by 1.81 percent and achieves comparable results on the VisDrone Challenge 2021. TPH-YOLOv5's improvements over the baseline model make it a promising option for drone-based object detection.

Reference:

[1] "CenterNet: Keypoint Triplets for Object Detection" by Xingyi Zhou, et al. (2019)

[2] "FSSD: Feature Fusion Single Shot Multibox Detector" by Zhiqiang Shen, et al. (2017)

[3] "PVANET: Deep but Lightweight Neural Networks for Real-Time Object Detection" by Kaiming He, et al. (2016)

[4] "RFBNet: Receptive Field Block Network for Accurate and Fast Object Detection" by Songtao Liu, et al. (2018)

[5] "Sparse R-CNN: End-to-End Object Detection with Learnable Proposals" by Peize Sun, et al. (2020)

[6] "EfficientDet: Scalable and Efficient Object Detection" by Mingxing Tan, et al. (2020)

[7] "FCOS: Fully Convolutional One-Stage Object Detection" by Zhi Tian, et al. (2019)

[8] "YOLOv3: An Incremental Improvement" by Joseph Redmon and Ali Farhadi (2018)

[9] "YOLOv4: Optimal Speed and Accuracy of Object Detection" by Alexey Bochkovskiy, et al. (2020)

[10] "YOLOv5: Improved Real-Time Object Detection" by Alexey Bochkovskiy, et al. (2020)