



**Ahmedabad  
University**

Report 3

CSE541 Computer Vision Section-1

Group members:

Kushalkumar Suthar (AU2140122)

Dhruvin Prajapati (AU2140064)

Rohit Rathi (AU2140023)

Nish Parikh (AU2140039)

## Analysis of Model: QueryDet

QueryDet is a novel query mechanism designed to accelerate the inference speed of feature-pyramid-based object detectors. It first predicts the coarse locations of small objects on low-resolution features and then computes the accurate detection results using high-resolution features sparsely guided by those coarse positions. This method not only utilizes the benefits of high-resolution feature maps but also avoids unnecessary computation for the background area. On the VisDrone dataset, which contains more small objects, QueryDet sets a new state-of-the-art while achieving a  $2.3\times$  high-resolution acceleration on average.

One notable observation addressed by QueryDet is the uneven distribution of computational costs across different layers of the FPN, particularly impacting the detection of small objects. To mitigate this issue, QueryDet introduces a sparse query approach. This involves predicting rough object locations on coarse feature maps and then focusing computational efforts on refining these predictions on finer feature maps. A query head is integrated into the model to predict these coarse object locations alongside the classification and regression heads.

During inference, locations with predicted scores above a set threshold are designated as queries. These queries are then mapped to key positions on the preceding layer's feature map. Sparse convolution techniques are employed to process only these selected positions, significantly reducing computational costs. Additionally, a cascade sparse query (CSQ) strategy is implemented, wherein queries for each layer are derived from key positions on the previous layer, thus avoiding exponential increases in computational complexity.

The query head is trained using FocalLoss as well, with binary target maps generated based on distances between feature positions and small object centers. To ensure balanced learning across all layers, the loss of each layer is appropriately weighted, particularly crucial with the addition of higher-resolution features.

In comparison to related approaches such as RPN-based detectors and PointRend, QueryDet distinguishes itself through its focus on classification in coarse prediction, sparse and selective computation, and distinct methods for query generation. Overall, QueryDet represents a significant advancement in improving the efficiency and accuracy of small object detection.

Model	AP[%]	AP50[%]	AP75[%]
Query Det	28.32	48.14	28.75

Table 1

Table 1 shows the performance values in terms of AP of QueryDet on the Visdrone-2019 dataset. Here AP50 shows the average precision value when the IoU threshold is 0.5 and similarly AP75 shows the average precision value when the IoU threshold is 0.75.

## References:

Yang, C., Huang, Z., & Wang, N. (2022). Querydet: Cascaded sparse query for accelerating high-resolution small object detection. Retrieved from <https://arxiv.org/abs/2103.09136>