Reviewer 1 Cover latter

First of all, thank you for reviewing the paper. We will answer your questions.

Q.1 It is said that the low-level feature information is lost in the FCOS pyramid of the base network. In section. 3, it is then asserted that the HIS block can recover this information. However, it seems that the base network already combines low-level features with up-sampled deeper features. Could you please be more specific about what is lost by FCOS ? The HIS blocks are more evolved than blocks used in the original FPN, and it is not obvious to me that the performance gain is due to the reason asserted by the authors.

A.1 As mentioned earlier, the information lost in FCOS is the characteristic loss that occurs because lateral connection is not used when down-sampling is used unlike the existing FPN. At this time, in the case of HIS block, it is a part that reduces the loss that can occur by composing spatial information and channel-emphasized information in a parallel structure. This part has been corrected in the paper.

Q.2 The motivation for the lightweight detection head is related to computing performance. However, the introduction of the HIS blocks looks more demanding in computing facilities (Table 2 reports 32.6M vs 32.1M weights). In section. 4, only the weight of the detection head is shown whereas the weight of the whole network is more meaningful to use HISFCOS on compute-limited devices. It should be stated that HISFCOS is heavier than the base network even with the new detection heads.

A.2 As you said, HISFCOS is about 0.5M heavier in parameters compared to the existing method (FCOS). However, in terms of GFLOPs (Computation cost) listed in the table, it is required to be lower, so it is marked as such because it has a lower computational amount compared to the previous (slightly higher) parameters.

Q.3 The ablation study in section. 4 mix the modifications HIS blocks + head together, it would have been better to see the performance of the HIS blocks alone first, and then with the proposed detection heads.

A.3 The table of the temperance study has been revised by referring to what you said.

Q.4 Typos

A.4 The typo was corrected after checking.

- Line 56: Minimizes → reduces

  I corrected the typo.

- Line 59-61: "that reduces the amount of computation while maintaining the computational cost" → unclear and misleading

  that reduces the amount of computation while maintaining the computational cost →

  We propose a lightweight detection head that reduces the amount of computation by improving the structure of the conventional detection head.

- Line 67 affected → affects

  I corrected the typo.

- Line 114 minimizes → reduces

  I corrected the typo.

- Line 142 ~~each symbol~~ ... represent~~ing~~

  I corrected the typo.

- Line 147 feature loss was ~~improved~~ *prevented*

  I corrected the typo.

- Line 166 Figure 6 leaks on the left margin

  I corrected the margin in Figure 6.

- Line 185 Entorpy → Entropy

  I corrected the typo.

- Line 187 g → y + missing conditions in the case for p_t definition

  I've added a description for symbols.

- Line 194 [[25]] (extra brackets)

  I removed the extra brackets.

- Line 202 what are Y and Y' ?

  I've added a description for symbols.

- Line 261-263 to detect [WHAT?] because

  Figure \ref{fig7} shows the detection results for FCOS and HISFCOS in the PASCAL VOC dataset. Figure \ref{fig7}(a) shows how FCOS makes it difficult to detect overlapping objects, such as images. During feature extraction, it is difficult to detect because of the loss of feature information. Additionally, in the case of the bottle class at the bottom, it is difficult to grasp the characteristics of an object, and it is difficult to detect when feature information is lost. However, the proposed method minimizes spatial and channel feature losses using the HIS block. Additionally, it was confirmed that the detection performance of the overlapping object, which was difficult to detect, was improved by improving insufficient low-level information by reconstructing the feature pyramid structure based on the results.
  →
  Figure 7(a) shows that FCOS is difficult to detect overlapping objects due to spatial and low-level feature information loss during feature extraction. In addition, the bottle class at the bottom of Figure 7 is a class that is difficult to detect when there is a feature loss. The proposed method uses HIS blocks to reduce spatial and channel loss. In addition, by reconstructing the feature pyramid structure, low-level information was improved to enhance the detection performance of difficult-to-detect overlapping objects. In Figure 7(b), compared to the conventional method, the proposed method shows improved detection results for objects that are difficult to detect such as overlapping objects.

- Line 280 Table 3 too large

  I think the size of the table is reasonable compared to the contents.

- Line 304 Table 4 too large

  I think the size of the table is reasonable compared to the contents..

- Line 305 Capitalize (or lowercase) title

  Lightweight detection head Analysis → Lightweight detection head analysis

- Line 318 Table 5 too large

  Table 5 fixed small size