Response to Reviewers' Comments

We would like to thank the reviewer for all the constructive comments, which have improved the manuscript significantly. A detailed response to all comments can be found below, where the comments are in regular font and our point-to-point responses are in bold font. Line numbers correspond to the revised manuscript.

Comments:

- Review of ESSD-2020-122 This manuscript presents a study where a deep convolutional neural network was designed and built to purposefully map marine agriculture at 16 meter resolution. The manuscript is well written and presents clear objective. My comments are:
  Response: Thanks for the positive comments and all the suggestions, which significantly helps we improve our manuscript.
- First, the authors should sharpen the subject of this study. Currently the focus of this study is on comparison of the new algorithm and other methods, while the feature or content of GF-1 data, as a new data source, is better to provide more info.
  Response: Revised as suggested. We added more introductions of the GF-1 to the data section (bellow and in the revised paper).
  Line 104-110:
  'The GF-1 satellite, which is the first satellite of the China high-resolution earth observation satellite program, was launched by the China Aerospace Science and Technology Corporation in April 2013. This satellite carries four integrated WFV sensors, providing multi-spectrum data with a two-day revisit cycle and a swath width of 800 km when the four sensors are combined. Each WFV sensor has four multi-spectral bands at 16 m spatial resolution: B1 (450–520 nm, blue), B2 (520–590 nm, green), B3 (630–690 nm, red), and B4 (770–890 nm, near infrared).'
- Another question is about the fine-tuning of various methods in the comparison. To make a fair comparison, are all the methods fine-tuned to their optimal status for the classification.Currently, this is not very clear.
  Response: Revised as suggested. All the compared methods are trained from scratch. We made this point clear in the comparison section.
  Line 199-201:
  'In the training phase, all of the above models, including the proposed HCHNet, were trained from scratch using the same patches and experimental settings as in the HCHNet method. '
- Also, the advantage of the proposed algorithm over the U-Net and HCNet is not very clear (also see minor comment # 5). This point deserves more clarification and discussion.

Response: Revised as suggested.We firstly revised the Fig.7 to show the advantages of the proposed algorithm. As in the last column of the Fig.7 , UNet or HCNet tend to recognize the MAC as others, leading to lower recall values.

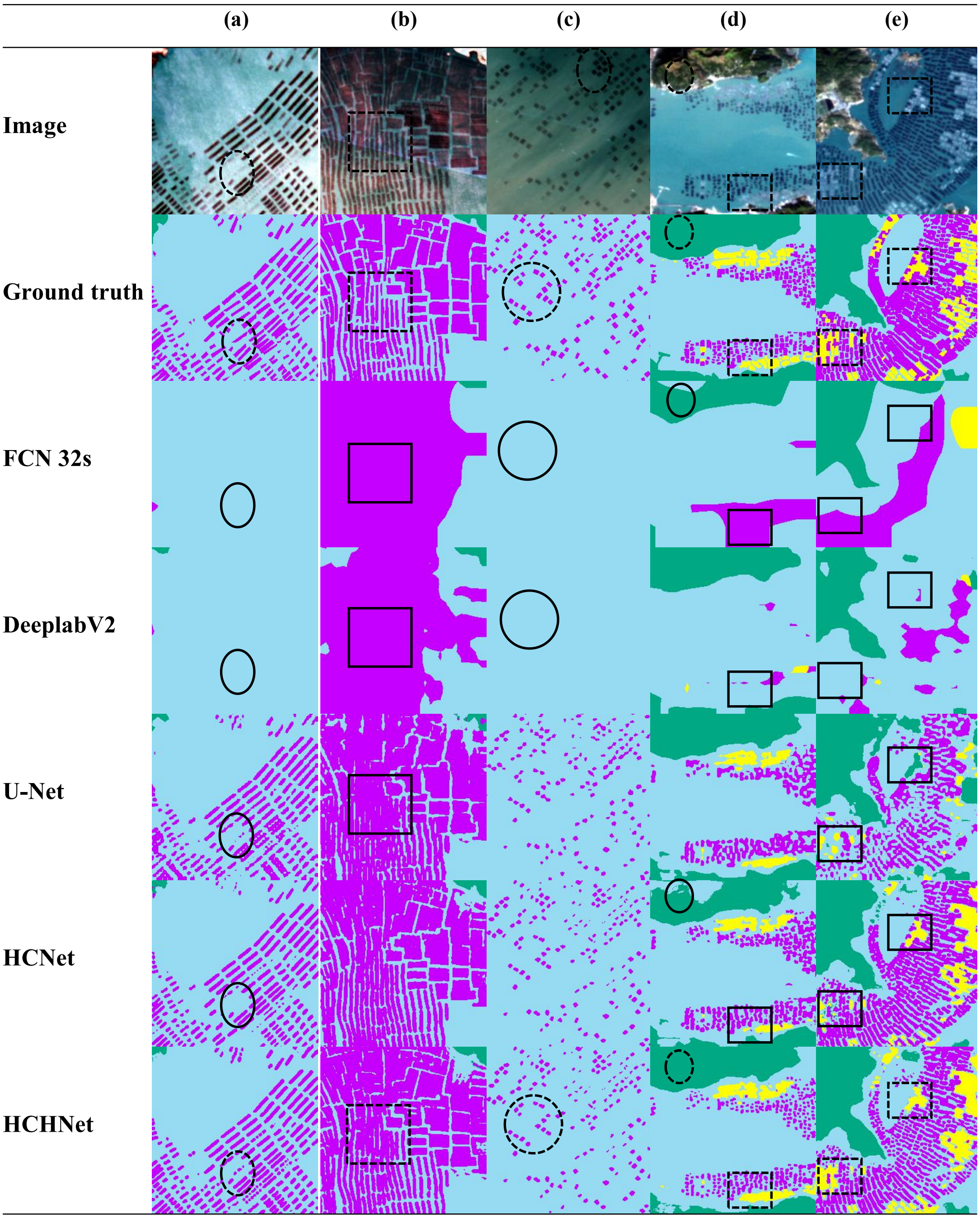

Figure 7: The classification results of MPC and MAC areas comparing the proposed HCHNet method with other approaches. The black solid outlined areas indicate where HCHNet obtains better results. The dotted line shows same locations in other images.The purple, yellow, blue, and green areas in the classification maps represent the MPC, MAC, sea, and land areas, respectively.

In addition, we also make this point clear in the results analysis part.

Line 266-268:

'Besides, the HCHNet also achieved a good balance between precision and recall values of MAC, identifying more accurate and existing MAC areas. The difference between them is less than 4% for the HCHNet, while the difference values of other methods are more than 28%.'

And then, we discussed more to analysis the reasons.

Line 289-294:

'The HCHNet achieved the best classification performance for three reasons: (1) all of the pooling operations were removed to avoid the shrinking of features, which helps improve the identification of smaller foreground objects; (2) the hierarchical structure was used to enlarge the receptive field to capture more contextual information, which is helpful for reducing the influence of local variance; (3) the weighted loss function was employed to solve the classes imbalance problem.'

Minor comments:

- 1. In the caption of Figure 7, the MPC area should be described as purple, instead of red.

  Response: Revised as suggested.

  'The purple, yellow, blue, and green areas in the classification maps represent the MPC, MAC, sea, and land areas, respectively.'
- 2. Line 94: 'environment' should be plural, i.e. 'environments'.

  Response: Revised as suggested.

  'the features of MPC in remotely sensed images are usually influenced by different environments (Fig. 2b, e, g, h , j), making it difficult for classification.'
- 3. Line 194: 'To perform the accuracy assessment' is a very general purpose. If not mistaken, the stratified sampling is done to ensure representativeness of each class in the whole sample population. I recommend state it more specifically.

  Response: Revised as suggested.

  'To ensure representativeness of each class in the whole sample population for accuracy assessment, we followed the widely used random stratified sampling method (Padilla et al., 2014; Ramezan et al., 2019) to generated 4000 randomly selected points in the coastal zone.'
- 4. Line 235: should be 'performed for 1000 iterations'.

  Response: Revised as suggested.

  'The bootstrapping was performed for 1000 iterations, and the mean of the distribution used for the evaluation and the confidence intervals was set as 95% quantile.'
- 5. Line 248: 'HCHNet identified more MPC and MAC areas than U-Net and HCNet,(Fig. 7a,d,e).' This statement is not visually obvious, especially in the d and e cases between UNet and HCNet. I suggest the authors rephrase or be more

specific when making this comparison.

Response: Revised as suggested.We revised the Fig.7 to show the advantages of the proposed algorithm. The revised Fig.7 and statements can be found in previous comments.

- 6. Line 276: should be '. . . more difficult to be implemented

  Response: Revised as suggested.

  > **'making such approaches more difficult to be implemented operationally for national-scale studies.'**