Response to Reviewer 1 Comments

We would like to thank the reviewers for their valuable time and thoughtful evaluation spent on this manuscript. The paper has been carefully revised according to the reviewers’ comments, and our response to each question is provided below, individually.

This study aims at detecting building corner in aerial images via adopting fully convolutional networks. This paper firstly uses FCN to give an initial segmentation of building, and then applies Fully-connected CRF and morphological operation to refine the segmentation results. Secondly some traditional methods are used for contour extraction. Finally, the simplified evolved method in scale space are used for corner detection. Experimental result validated the proposed method outperform the existed corner detection method.

**Point 1:** This study used the complicated FCN for building segmentation. However the final purpose of this study is corner detection, which is implemented by the traditional methods. I think that the simple way is to apply CNN for direct building curve extraction or corner detection.

**Response 1:** We agree with the reviewer that it would be simple if CNN could be directly applied for direct building curve extraction or corner detection. However, according to our practice and experience, such approaches are not able to deliver superior performance in building corner detection. In the revised paper, a subsection (Section 4.3) is added to discuss this issue.

For verification, a reference approach is developed in Section 4.3, which directly extracts building contours by using a CNN model trained with the same dataset as our proposed approach. Performance of the reference approach is demonstrated in Figure 8 (Page 10). It can be seen that the extracted building contours could be blurred, fragmented or miss-detected (see the last column of Figure 8). By comparing with the results of our proposed approach (the last column of Figure 5), it can be concluded that the corner detection scheme proposed and used in our work represents a more effective way for detecting building corners from aerial images. Note that the corner detection results produced by using the reference approach have also been presented in Figure 9 and Figure 10, where it is further verified that our proposed approach is more suitable for building corner detection.

**Point 2:** As showed in the segmentation results, if only evaluating the tripap with small width, the performance of the segmentation results will be decreased, which manifest the better performance of the segmentation not always generate better result for the final corner detection.

**Response 2:** In the revised paper, Section 4.2 is added to verify the fact that better performance of the segmentation can generate better result for the final corner detection. To
be specific, in this added subsection, corner detection is conducted on the segmentation produced with the existing FCN and FCN-CRF models, and our proposed FCN-CRF-MOP model, individually. The results are presented in Figure 7 (Page 9), where it can be seen that our proposed model delivers better results over the existing models in corner detection.

The comment that “if only evaluating the tripap with small width, the performance of the segmentation results will be decreased” is a misunderstanding of what we have expressed, i.e., “the measured accuracy of each method increases as the trimap width increases”. Note that this phenomenon we have expressed is only related to the change of evaluation conditions. In other words, a model could have a low measured accuracy when a stricter evaluation condition is imposed. On the other hand, the model always has “fixed” performance for a given image. To avoid confusion, in the revised paper a description sentence is corrected (Lines 140-141, Page 6): “It is seen that the measured accuracy of each compared method increases as the trimap width increases (that is, when a less-strict evaluation condition is imposed).”

Point 3: Although the author provided that the post-processing: fully-connected CRF and MOP can improve the performance of trimap evaluation, it seems that no intuitive effort is explored for improving the segmentation results of the building boundary.

Response 3: In the paper, Figure 3 demonstrates the progressive improvement achieved by using the CRF and MOP. Since the improvement effect might not be intuitive enough for the readers, we have rewritten the corresponding paragraph for a more specific description (Lines 119-123, Page 5): “however, the segmentation maps are somewhat blurred and inaccurate. By using the CRF as a post-processing, the boundary edges of buildings become sharper, as shown in the fourth column. Unfortunately, some segmentation flaws, such as spur-like artifacts, are still observed, which could incur many false positives in the subsequent corner detection process.”

In fact, the progressive improvement effect has also be clearly verified with an objective evaluation. Table 1 shows that the FCN has a mean IoU of 92.83%, the FCN-CRF has a mean IoU of 93.16%, and the FCN-CRF-MOP has a mean IoU of 93.21%, respectively.

Point 4: The experimental results are compared with the traditional corner detection method, which are not based CNN models.

Response 4: To our best knowledge, at present there are no CNN models that have been proposed and used for corner detection, either for a generic corner detection task or for a specific corner detection task. In fact, according to our experience, it could be a very hard problem to design and train such a CNN model. Therefore, there are no CNN-based corner detectors that can be compared with our detector.
**Point 5:** How about the compared results of the corner detection using the building segmentation with FCN, FCN-CRF, and FCN-CRF-MOP.

**Response 5:** In the revised paper, experiments are conducted to evaluate the corner detection performance of these three models. The results are presented in Figure 7, where it is shown that the F-measure values of corner detection achieved by using the FCN, the FCN-CRF and the FCN-CRF-MOP are 0.74, 0.81 and 0.83, respectively.