Response to Reviewer 2 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.

The paper proposed a method for building corner extraction in high-resolution satellite/aerial images by combinations of building segmentation and contour curve analysis. Although the individual approaches are not very new, the combination is innovative for the task of building corner extraction with the robustness supported by the established methods. I would recommend this for publication in Sensor though I would like to request authors for a revision with some more information and discussions as below.

**Point 1:** Please highlight some false cases to discuss limitation and possible countermeasures for future works.

**Response 1:** In the revised paper, Section 4.7 has been added to discuss the above-mentioned issues (Pages 11-12, Lines 275-284). In brief, it is highlighted that for buildings with complex structure, the extracted contours might be blurred and some false negatives could be yielded in corner detection. A possible way to cope with this problem is to add the curvature values of each labeled building contour to the FCNs so that an edge-preserving network model can be trained for building contour extraction. This topic would be one of our future works.

**Point 2:** Abstract and conclusion - Please refer to the accuracy indicators from the results.

**Response 2:** Corrections are made as suggested. In the revised paper, it has been clarified both in the abstract and conclusion that the proposed corner detector can achieve an F-measure of 0.83 on the test image set. (Page 1, Line 13; Page 12, Lines 292-293)

**Point 3:** L85 - Please indicate the spatial resolution of the input data. It is crucial for the parameters in the algorithms that I would request descriptions as the comment below.

**Response 3:** Corrections are made as suggested. In the revised paper, the spatial resolution of the input data has been specified: “The dataset of aerial images used in the following for training (11,700 images of 321×321 pixels) and testing (450 images of 321 × 321 pixels) is produced from the Vaihingen.” (Page 4, Lines 88-89)

**Point 4:** Section 3.1.2 - Please indicate parameters in the algorithms. As long as I understand correctly, it should be with patch/window size and batch size.
Response 4: Corrections are made as suggested. In the revised paper, the parameters used in the algorithm have been specified: “The ResNet-101 model contains a $7 \times 7$ convolutional layer and 5 residual blocks. Each residual block includes several $3 \times 3$ or $1 \times 1$ convolutional layer. The latter two blocks are re-purposed by atrous convolution … where the batch size is taken to be 1.” (Page 4, Lines 110-112 and Line 114)

Point 5: L183 - Please indicate parameters in the compared algorithms if applicable.

Response 5: Corrections are made as suggested. In the revised paper, Table 3 is added to summarize the parameter settings of the compared algorithms, and necessary description is also added (Page 10, Lines 244-248).