Comment on essd-2022-19
Anonymous Referee #2


The paper presented a nationwide dataset of roadside noise barriers (RNB) for China, which was mapped from street view photos using machine learning algorithms. According to the manuscript, the dataset was with high spatial accuracy and has potentials for urban studies. While the methodology and algorithm assessments seem reasonable to me, the paper is more like a technical report instead of an introduction of a useful dataset. My specific comments are:

1) Why the dataset is important? Are there any specific reasons why you created the dataset? Instead of some vague statements like “useful to a variety of urban studies”, it would be more convincing to list some specific applications the dataset would have in China.

2) As mentioned, the current version of this manuscript is more like a technical paper, without comprehensive assessment of the dataset itself. Users would like to know more details about reliability and limitations of the dataset, such as spatial variations of mapping accuracy at the city scale, limitations in cities where limited street view photos are available, and the timing of detected RNB across cities. Given that the input street view photos were collected from 2014 to 2020 with uneven spatial distribution, knowing whether mapping accuracy was impacted by data completeness is particularly important. Without such information, it is hard to know whether the total 2227 km of RNB is reliable or not, as well as for each province.

3) Technically, readers and users would like to know what types of RNB were mapped in the dataset and how the authors visually interpreted training, validation, and testing photos. It would be clearer to have some examples of RNB and detail logics used to manually label samples.
4) It seems the deep learning architecture used in this study was adopted from others. I
do not criticize the algorithm (although I do not think it’s novel), but more specific design
related to RNB detection should be considered and evaluated. Table 3 and 4 do not
support that incorporating context information contributed to higher mapping accuracy.