We firstly performed a query on Scopus, which is the largest database of peer-reviewed literature. The strategies/protocols for searching the papers are as follows:

1. The query is based on the predefined keywords of “building detection” or “building extraction” in the article title, abstract, and keywords under the category of “remote sensing” and using a time range from January 1, 2000, to December 31, 2021.
2. The articles whose main text was written in non-English language are removed.
3. The books, reviews, and the informal publications including conference and conference review papers are removed. In other words, only journal articles and are reviewed.
4. We then read each one of the remaining works, manually removed those articles meet following criterions：

4.1) the studies do not use VHR optical images. Here VHR optical image refers to remote sensing data from visible to infrared spectrum, with spatial resolution ranging from 0.05m to 4m.

4.2) the application studies do not focus on the building detection technique.

4.3) these studies only focus on the interpretation of damaged buildings or building facades. Damaged buildings and building facades have different characteristics from ordinary buildings. The main data sources of damaged building detection / change detection and building facade are active remote sensing data, that is, SAR is used for the former and point cloud is used for the latter, rather than optical images.

4.4) although the studies address land cover classification, it does not highlight the characteristic of building in the text.

4.5) the studies address built-up related studies. Although built-up is dominated by buildings, the spectral, contextual characteristic are different from buildings.