General comments:

This study presents the performance and limitations of computer vision methods in detecting natural (vegetation) and urban debris around hydraulic infrastructure (bridges, culverts).

Given (a) the increasing intensity of rainfall events that lead to extreme floods, (b) the importance of bridges and culvers for transportation services and emergency response efforts, and (c) the potentially detrimental effects of debris on infrastructure safety and the level of flooding in nearby areas, the topic of the manuscript is significant and timely. Moreover, the findings of the paper could potentially prove useful to public agencies responsible for infrastructure maintenance, especially in flood-prone areas, making their efforts more efficient and cost-effective.

Generally the article is well written and technically sound. Although the “water” part, which is the main focus of the particular journal is limited in this article, its topic fits perfectly the section “New Sensors, New Technologies and Machine Learning in Water Sciences”, to which it has been submitted.

Therefore, the reviewer suggests the manuscript for publication with Minor revisions. A few comments are shown below, with the hope that they will help the authors improve the manuscript further.

Comments/revisions:

1. The authors mention in the Abstract and in Section 5 that they used real images of debris blocking the hydraulic structures extracted from Wollongong City Council (WCC) records AND simulated images of scaled debris blocking the culverts collected from hydraulics laboratory experiments. Please clarify or discuss in the manuscript:
   (a) Why did you have to use both types of images?
   (b) Wouldn't the accuracy of the computer visions method (CVM) be affected by the simplification of the background in the flume experiments? Moreover, since most of hydraulic experiments are usually scaled (what scale exactly did you use), wouldn’t the differences in the size of the debris in flume experiments and real life affect the calibration and performance of the CVM? For example, if someone was to calibrate the CVM only based on laboratory results...would they be able to detect accurately debris in real life (which will have different background and size)?
   (c) Which images did you use for training/testing and what criterion did you use in order to select the ones for training and the ones for testing? For example, did you use 50% of the lab images and 50% of the real images for training (with random selection), and the remaining images for testing?

2. **Line 26-39:** The authors mention in this paragraph the catastrophic effects of floods on properties and agriculture systems and at the end say that in three recent floods (in Australia, UK) hydraulic structures failed due to debris effects. Please note that:

   - Critical hydraulic infrastructure (e.g. bridges) can fail during extreme floods, even if debris is <u>not</u> present. In fact, extreme flows can cause <u>scour</u> of the bridge piers (there are plenty of studies in the last 20 years, e.g. Chiew et al., 1987) and very <u>large hydrodynamic loads</u>

especially during transient flows (this topic has attracted attention more recently, e.g. Istrati et al, 2022). So, generally speaking all parameters or a combination of them (scour, hydrodynamic loads, debris) might be responsible for the sustained damage. Therefore, the authors are advised to consider including 1-2 new sentences, which will mention the aforementioned reasons (scour, hydrodynamic loads, debris) with appropriate references, in order to provide an objective overview of the most common reasons related to floods that can cause damage to hydraulic structures.

- It would be valuable for the reader to explain in one or two sentences, why the debris can be so catastrophic for hydraulic structures (e.g. piers, bridges etc). For example, past work showed that during extreme flows/floods water-borne debris can cause (i) large impulsive loads on simplified structures (Haehnel et al, 2004) and bridges (Hasanpour et al, 2022), and (ii) damming in front of the bridge that affects the flow properties (e.g. flow velocities) around the structure and the applied moments, such as, pitching, yaw roll, increasing the demand in individual components and consequently the likelihood of failure (Istrati et al, 2020).

Chiew et al (1987). Local scour around bridge piers. Journal of hydraulic research, 25(1), 15-26. https://doi.org/10.1080/00221688709499285
Istrati et al (2022): "Numerical Investigation of Dam Break-Induced Extreme Flooding of Bridge Superstructures", 3rd Intl Conf on Natural Hazards & Infrastructure, 2021-2022, Athens, Greece, https://www.researchgate.net/publication/362108279_Numerical_Investigation_of_Dam_Break-Induced_Extreme_Flooding_of_Bridge_Superstructures
Haehnel et al (2004) Maximum impact force of woody debris on floodplain structures. J. Hydraul. Eng. 2004, 130, 112–120. https://doi.org/10.1061/(ASCE)0733-9429(2004)130:2(112).
Hasanpour et al (2022): "Multi-Physics Modeling of Tsunami Debris Impact on Bridge Decks", 3rd International Conf on Natural Hazards & Infrastructure, 2021-2022, Athens, Greece, https://www.researchgate.net/publication/361943381_Multi-Physics_Modeling_of_Tsunami_Debris_Impact_on_Bridge_Decks
Istrati et al (2020). "Numerical investigation of tsunami-borne debris damming loads on a coastal bridge." 17th World Conf Earthq Engineering, Sendai, Japan.

3. **Line 42-43:** The authors mention here the inability of conventional flood models to account for non-linear factors such as debris. The reviewer agrees totally with this statement. However, please note that recently multi-physics numerical approaches that combine particle-based methods and mesh-based methods, such as the coupled SPH-FEM, have been used successfully for simulating water-borne floating debris effects during tsunami-induced floods, including both the debris movement and the loads on the structure (Hasanpour et al, 2021). Therefore, this modeling approach would be easily applicable to similar phenomena during riverine floods. Moreover, other multi-physics methods, such as the coupled SPH-DEM, have been successfully used for simulating debris in soil flows (Trujillo-Vela et al, 2020), which is a different application than the one investigated in the current manuscript, but could potentially be adjusted in the future for investigating water-borne debris flows. Therefore, the authors are advised to either rephrase this sentence or add a new sentence, in which they will clarify/mention that although conventional floods models cannot capture the debris effects, multi-physics methods have shown a promising performance/accuracy in the aforementioned studies (and perhaps in additional

studies that the authors can find on the web), in order to provide an up-to-date overview of the state-of-the-art in the numerical modeling of water-borne debris.

Hasanpour et al (2021). Coupled SPH–FEM Modeling of Tsunami-Borne Large Debris Flow and Impact on Coastal Structures. Journal of Marine Science and Engineering, 9(10), 1068.
Trujillo-Vela et al (2020). Smooth particle hydrodynamics and discrete element method coupling scheme for the simulation of debris flows. Computers and Geotechnics, 125, 103669.

4. **L246-250:** The authors explain here the stages that involved the training and testing of the computer vision methods. Please clarify which images (real or lab) were used for training and which for testing. (this comment is also related to my first comment)