

Summary of Research

This paper demonstrates the novel application of deep convolutional neural networks to determine the presence of flooding in CCTV footage. In addition, the extent of the flooding can be roughly determined using a new SOFI index. The paper's key findings include:

- The successful segmentation of CCTV images to identify areas covered by water.
- The development of the SOFI index to quantify flooding extent and analyse water level fluctuations.
- SOFI loosely correlates with water depth, which may prove useful to the future calibration of flood models.

The paper demonstrates its results over six CCTV sequences, taken from a variety of locations and test sites. Two examples are also accompanied by water level recordings, providing an objective comparison for the technology.

Context within current research

The presented work appears to be new and novel, contributing to a small but growing pool of work on the subject. Other work in the field has concentrated on detecting flood depth from CCTV images or application to still images (particularly from social media). Although the application of deep learning (convolutional neural networks) is widespread across computer vision problems, this is a novel application, and the technology's implementation has been thoroughly explained in this paper.

Strengths and weaknesses

Overall, I feel the research presented in this paper is of a high quality, and provides a plethora of insights into the technology and its potential applications. The results are presented in a clear manner, which should be accessible to all readers. However, I was a little surprised by the choice of journal for publication. The paper doesn't feel like it fits perfectly with the journal's target audience, even though the work presented within is of extremely high quality. I tend to agree with the previous reviews, that the technical elements of the methodology may be hard to follow for readers not familiar with deep learning. Even so, the technical content is well referenced, enabling a reader to further explore and understand the more complex elements.

Other reviews have questioned the usefulness of the SOFI descriptor and the conversion to water depths. I tend to agree with the author that SOFI works well as a distinct tool. This does limit its usefulness for existing flood modelling, but it can still be used for contextual validation, even if that is only in a binary manner (flooding present or not). The translation water depths from CCTV footage is a very different and extremely complex problem, given the tremendous volume of noise in 'wild' CCTV footage.

Comments

Key points

Page 7 Line 21: You describe the 'Fine-Tuning' process, however you don't comment on the viability of this for mass implementation. This could be particularly problematic as you have moved from a single holistic DCNN to many (independent) machines. Furthermore, this does imply that you have footage containing a flood for that camera feed, which is extremely unlikely, especially if someone planned to roll this out to tens of thousands of CCTV cameras. Generally speaking, I wouldn't rely on this fine-tuning process and believe it to be extremely situational in its usefulness.

Page 8 Line 27: Following on from the scalability issues with 'Fine-Tuning', the manual definition of ROIs would not be viable for mass implementation. Even though you found the use of ROIs to be unnecessary, it may be worth the investigation of automatic ROI generation (in future work). Not only would this improve the scalability of the technique, but could improve the calculation of a SOFI index in video containing water multiple water sources.

Other notes

From the case studies provided, the technology appears to have been demonstrated on largely still/slow moving water. It would be good if you could comment on the techniques application to moving water (particularly relevant in flash flooding) as still/white water are visually quite distinct.

Another potential application for this technology may be in key infrastructure/assets (i.e generators/pumping stations/power stations) that are particularly at risk of damage to flooding. Quite often these assets will have CCTV cameras installed for security. This technology could act as an additional alarm/early warning system for these at-risk assets and asset failures.

Page 8 Line 19: You discuss issues arriving with your SOFI descriptor if the scene changes suddenly (something is moved by flood water or a vehicle parks in the scene). However, this should be visible in the SOFI curve? Assuming you could work with dynamic thresholds/filtering, this issue could be could be tackled in future work.

Page 8 Line 30: Sentence starts 'provides characteristics of these videos' I assume a word is missing? Otherwise I would advise re-wording.