

## Response to the Reviewers Comments

Author reply to RC2 egusphere-2025-77 (NHESS)

Paper title: Identification of nighttime urban flood inundation extent using deep learning

Authors: Jiaquan Wan et al.

The paper presents an interesting contribution on the use of deep learning to identify flood extents in urban areas from night-time images. As such it is relevant to NHESS, but it has a strong focus on the deep learning methods. Many readers of the journal, like myself, will not be familiar with the detailed language and concepts used in deep learning and the paper needs to be re-written with this audience in mind. Sections 2.1 and 2.2 need particular attention.

**Response:** We sincerely appreciate the reviewer's recognition of our study and the constructive feedback provided. In the revised manuscript, we will make comprehensive adjustments, particularly in Sections 2.1 and 2.2, by simplifying the language, reducing technical jargon, and providing additional background explanations to enhance readability for readers outside the deep learning field.

Much of the language in the paper is opaque and uses terms that are not common in scientific discourse (some of these are set out below). There also needs to be greater clarification of what the authors set out to do in the research, which of the methods they developed themselves, and what the conclusions mean for those working on urban flooding. For example, on Line 82 the aims are listed, but these are in fact a description of what was done without a justification. Further Line 74 says that NWseg is "proposed" and Line 91 says that the NWseg is "contributed", but was it developed by the authors or taken from other research work?

**Response:** We sincerely appreciate the reviewer's in-depth critique and constructive suggestions regarding the language and content of this paper. In the revised manuscript, we will comprehensively refine the language, clarify the research objectives and methodology, and further elaborate on the practical significance of our conclusions.

### 1. Clarification of Technical Terminology:

To improve readability and adhere to scientific writing conventions, we will simplify complex deep learning terms such as illumination-invariant reflectance and disentanglement module. In the revised manuscript, we will provide clearer explanations in layman's terms to enhance accessibility for a broader audience.

### 2. Refinement of Research Objectives:

We will clarify our research objectives in the introduction as follows:

*This study aims to address the technical challenges of nighttime urban flood extent recognition and overcome the limitations of existing methods in low-light conditions. To achieve this, we first constructed a high-quality and representative nighttime urban flood dataset encompassing various lighting conditions and complex scenarios. Subsequently, we proposed a deep learning-based nighttime semantic segmentation model, NW-Seg, to enhance segmentation accuracy for flood-affected areas. Through comparative experiments with existing models, we systematically evaluated its effectiveness and applicability across different scenarios, providing a more reliable technological solution for nighttime urban flood monitoring.*

### 3. Clarification of Research Significance in the Conclusion:

We will refine the conclusion to better highlight our study's contributions as follows:

Existing research primarily focuses on flood extent recognition in daytime scenarios, while nighttime flooding remains understudied due to the lack of datasets, complex lighting conditions, and the limited adaptability of current models to low-light environments. This study introduces a novel technological approach for the intelligent monitoring of nighttime urban flood extent. By addressing a critical research gap, our work not only advances the field of nighttime urban flood recognition but also serves as a reference for future deep learning applications in extreme lighting conditions.

Through these revisions, we aim to enhance the clarity of our language, better define our research objectives, and highlight our contributions more explicitly, while ensuring a clearer impact on urban flood studies. Once again, we sincerely appreciate the reviewer's valuable feedback, which has significantly improved the quality and relevance of our manuscript.

Significantly more than half of the references that I tried to read online returned an error message or a message in Chinese characters. Further, many of these are to non-peer reviewed sources. Whilst the latter is acceptable in a few cases, the former is not at all acceptable in an international journal.

**Response:** We sincerely appreciate the reviewer's meticulous examination and constructive criticism regarding the references. Upon thoroughly reviewing the original reference list, we found that some links were invalid or directed to Chinese-language pages. This may be due to our citation of sources available only in Chinese academic databases or links that have become inactive over time. In the revised manuscript, we will remove all references that are inaccessible online or available only in Chinese and replace them with internationally recognized, peer-reviewed journal or conference papers accessible in English. Additionally, we have significantly reduced citations from non-peer-reviewed sources to ensure the accessibility and academic rigor of the references, aligning with the standards of international journals. We sincerely thank the reviewer for this valuable suggestion, which has significantly improved the academic quality of our paper.

#### **Specific points are:**

Line 50: a more scientific term that "remarkable" would be more appropriate.

**Response:** We will adjust the presentation of line 50 to ensure that the wording is more scientific and consistent with academic standards. In addition, we will check similar wording in the full text to ensure scientific and consistent language. We thank the reviewers for their valuable comments, and this suggestion significantly improves the academic rigor of the paper.

Line 54: it is stated that such surveillance is "ubiquitous", but whilst this may be true in the authors' experience is not true in all countries. This should be acknowledged and it limits the usefulness of these methods.

**Response:** We sincerely appreciate the reviewer's careful attention to the wording in Line 54 and the constructive suggestion. In response, we have revised the corresponding expression in the manuscript and incorporated a discussion on this limitation.

In the original text at Line 54, we assumed the widespread presence of surveillance systems based on the research background, which primarily focuses on China's high-density urban monitoring networks. However, we overlooked the fact that in certain countries or regions, such as developing countries or low-density urban areas, surveillance coverage may be limited. To address the reviewer's concern, we have revised Line 54 to: 'Simultaneously, with the increasing level of urbanization, surveillance systems have been extensively deployed across most urban areas.'

Additionally, we will expand the discussion section to analyze the current limitations of the model: The NWseg model relies on surveillance systems or mobile image data, which may limit its applicability in regions with weak monitoring infrastructure. Future research could explore the integration of drone imagery or other portable data sources to enhance its adaptability.

Line 55: of these three references two (Cheng, Yang) have links that do not work and one is to a publication that is not peer-reviewed. The remaining one does not mention whether this methodology has been tested in more than one country. Please clarify.

**Response:** We sincerely appreciate the reviewer's meticulous examination of the references. We acknowledge that some cited sources have accessibility issues, and some are not peer-reviewed. To address this, we will replace non-compliant references, prioritizing peer-reviewed journal articles and conference papers to enhance the reliability and academic rigor of the citations.

Line 89: the term "ablation" is common in machine learning, but is an example of a term that needs explaining to a different audience.

**Response:** We sincerely appreciate the reviewer's valuable feedback. We recognize that while 'ablation' is a common term in the field of machine learning, it may not be familiar to all readers of this journal. To enhance readability, we will provide the following explanation when the term is first introduced in the manuscript: Ablation refers to the systematic removal or modification of specific components of a model to evaluate their impact on overall performance.

Line 99: there is no subject in this sentence so we cannot see who proposed the model.

**Response:** We appreciate the reviewer's correction. We will revise the sentence to explicitly state that the model was proposed in this study, ensuring clarity and accurately attributing the research contribution.

Line 102: many terms in this paragraph need clarifying for a non-expert in machine learning, some examples are: SOD, illumination-independent reflectance, semantically supervising the training of the de-entanglement module, Retinex, Illumination-Aware Parser (IAParser), pyramid pooling module and a convolutional layer to construct an attention mask.

**Response:** We sincerely appreciate the reviewer's valuable feedback. We recognize that this paragraph contains multiple technical terms from the field of machine learning, which may pose challenges for readers unfamiliar with this domain. To address this, we will provide explanations for key terms in the revised manuscript, ensuring that the technical descriptions are simplified without compromising scientific rigor, thereby enhancing clarity and overall readability.

Line 115: a diagram showing how all these methods fit together would help readers understand what you are doing.

**Response:** We sincerely appreciate the reviewer's valuable suggestion. Line 115 of the manuscript primarily introduces other mainstream segmentation models to provide background for the subsequent performance comparison with our proposed NWseg model. The reviewer suggested using a diagram to illustrate the relationships among these methods for better comprehension.

We have carefully considered this suggestion; however, since this study focuses on the performance comparison between NWseg and other models rather than an integrated framework of different approaches, we believe that the textual description is already sufficiently clear and that an additional diagram is unnecessary. Furthermore, we will revise the abstract and introduction to explicitly clarify the comparison logic in the revised manuscript. We appreciate this suggestion, as refining these sections has significantly improved the logical clarity of the paper.

Line 189: what is a "Labelme tool"?

**Response:** We appreciate the reviewer's question. Labelme is an open-source image annotation tool widely used in image segmentation and object detection tasks. It enables users to manually annotate target regions in images by drawing polygons, thereby generating datasets for deep learning applications. In this study, Labelme was used to annotate the contours of flood-affected areas to construct a high-quality training dataset, enhancing the accuracy of the segmentation model. In the revised manuscript, we will further elaborate on this tool and its role to ensure better understanding for readers unfamiliar with it.

Line 191 says that the work was done by three graduate students. Rather than describing who did the work, it is necessary to explain how they did it and how the quality of the analysis was checked and ensured.

**Response:** We sincerely appreciate the reviewer's detailed suggestions regarding Line 191. In the revised manuscript, we have removed references to who conducted the work and instead provided a detailed description of the workflow and quality control measures to enhance the scientific rigor and transparency of the study. In Section 3.1 Construction of Dataset, we have added the following clarification:

*We assigned three researchers with expertise in urban flood studies to independently annotate 4,000 nighttime flood images using the professional annotation tool Labelme, delineating the inundated areas. To further enhance annotation accuracy, all annotated images were cross-checked, and discrepancies were reviewed and adjusted to ensure*

consistency. Finally, the annotated dataset was validated by experts in flood research to guarantee data quality.

Line 193: "waterlogged" refers to soil saturation. I think "inundated" would be a better word.

**Response:** We sincerely appreciate the reviewer's detailed suggestions regarding the wording in Line 193. In the revised manuscript, we will correct any inaccuracies to ensure that the terminology aligns with the context of urban flood inundation.

Line 196: this and other figures captions need more details.

**Response:** We sincerely appreciate the reviewer's detailed suggestions regarding figure titles. In the revised manuscript, we will carefully refine the titles of all figures, incorporating additional details to enhance clarity and scientific precision.

Section 4.1: I think there should be less discussion of the three inferior methods as the differences between them are minor compared to their differences to NWseg.

**Response:** We sincerely appreciate the reviewer's suggestion regarding Section 4.1. In the revised manuscript, we will streamline this section by reducing the detailed comparison and analysis of suboptimal methods, instead emphasizing the overall superiority of NWseg and further highlighting its advantages.

Line 250: I don't see any experimental results in the text i.e. data that was collected through physical measurements on site.

**Response:** We sincerely appreciate the reviewer's thorough examination and valuable suggestions regarding Line 250. The original text states, 'Training results of NWseg, ResNet50-FCN, LRASPP, and U-Net models on the training set,' which describes the performance of these models on a night-time flood image dataset rather than data obtained through physical field measurements (e.g., water depth or flow rate).

Additionally, the experimental data in this study are derived from an annotated surveillance image dataset rather than on-site physical measurements, aiming to assess the models' capability in identifying flood-affected areas. This aligns with the design objective of the NWseg model, which is to perform semantic segmentation based on surveillance imagery. In Section 4.3, Real-world scenes prediction comparison, we evaluated the models' performance in various complex scenarios. To further enhance the reliability of our assessments, we will manually annotate the flood extent in each scene and compare these annotations with the model predictions, providing a more comprehensive evaluation of model performance.

Line 252: what is "social inundation"?

**Response:** We sincerely appreciate the reviewer's careful examination and inquiry regarding the wording in Line 252. In our study, the intended term should be nighttime urban flooding dataset rather than 'social inundation.' In the revised manuscript, we will correct this terminology to ensure accuracy and scientific rigor while providing clearer context for the dataset.

Line 254: "exceptional" is too strong a word here.

**Response:** We sincerely appreciate the reviewer's detailed suggestions. In the revised manuscript, we will carefully review similar wording throughout the text to ensure scientific accuracy and consistency. We sincerely appreciate the reviewer's valuable suggestion, which has significantly enhanced the academic rigor and credibility of our paper.

Line 285: these images, and later one, are too small.

**Response:** We sincerely appreciate the reviewer's detailed suggestions regarding Line 285 and the related images. In the revised manuscript, we will adjust these images to enhance their clarity and readability, ensuring an effective presentation of the experimental results.

Line 328: if the conclusions state that there is a high computational demand, this should be investigated and reported on in the results section. How much greater is it? How long did it take? What sort of computer was used? Does this allow for practical use of NWseg?

**Response:** We sincerely appreciate the reviewer's thorough examination and valuable suggestions regarding Line 328. In response, we have added an analysis of NWseg's computational requirements in the results section and further clarified its practical implications in the conclusion to directly address the reviewer's concerns.