We would like to thank the reviewers for their time in thoroughly reviewing the work and providing helpful comments. We have addressed the comments below:

Comments from the Editors and Reviewers:  
  
***Reviewer 1****: I liked the paper a lot. It was very easy to read, systematic, and comprehensive in its description of the research variables. I have no critical comments against the paper in its current form as I believe it is of very high quality and a clear contribution to the literature.*

**Reviewer 1**: *I only have two non-compulsory comments for the authors to consider:  
- In Section 4.3 the following is mentioned, "We also note that applying a trigger to the input image of the computer vision model is not a very practical attack in the application of Bauza USV." Please elaborate on why the attack is not practical. This could drive the discussion regarding the limitations of the attacks to define the types of systems at risk.*

**Response:** We agreeproviding some more context to this statement will improve the work. We have first clarified that we are discussing a digital patch rather than a physical patch “*We also note that applying a digital trigger to the input image of the computer vision model*”, we have also expanded on the paragraph to add “as making a digital manipulation would require access to the model (which would be aboard the vessel) at the time of inference ”. p.g. 14

**Reviewer 2:** *The shown attacks are highly fitted to the use case at hand, although it is intended to demonstrate the framework, a discussion regarding this limitation is expected. Another aspect, at a more general level, is related to the types of investigated attacks, poisoning, backdoor, evasion, etc. The nature of these attacks is highly reliant on the AI systems and datasets they are developed with. Making the generalizability of such attacks very limited, unless for instance they are evaluated for a heterogeneous fleet relying on the same technology. From a Red team framework perspective, developing test cases for these attacks for each AI system might be resource and time-intensive. I believe the reader could really benefit from such a discussion.*

**Response:** We agree this is a valuable discussion to add. We have now added to the discussion section “Although the attacks are fitted to the use case to demonstrate the framework, they would likely need to be adapted for other technologies. More complex attacks may be required as the technology becomes more sophisticated (e.g., more robust AI could make the development of attacks more challenging). The use of multimodal input AI (e.g., utilising Radar, AIS, FLIR and optics data to make predictions) may also require more sophisticated attacks against different modalities. The development of these attacks could also be time and resource-intensive, one way to mitigate this limitation is for the community to responsibly develop AI attacking tools which can be applied to as many systems as possible with minimal adaptation. ” p.g., 21  
  
***Reviewer 3****: 1. Literature review together with analysis on maritime autonomous systems should be thoroughly presented in the background and related works.*

**Response:** SecureAI maritime autonomy encompasses a large number of research domains. However, this paper focuses on the most relevant work and provides a literature review for the security of AI, existing secure by design AI frameworks, red-team AI frameworks (which focus on LLM security). Marine AI is used for a large number of purposes for MAS so we have provided some surveys that focus on common areas which we can refer interested readers to instead of an exhaustive list which would be beyond the scope of the paper (to focus on the security of AI). We have added:

“The current landscape considering marine AI encompasses many different technologies, these include marine AI for MAS to support greater situational awareness \cite{ thombre2020sensors}, object detection and anticollision \cite{shao2022multi, molina2021autonomous, zhang2021collision} and marine AI to support the navigation of MAS \cite{ brandsaeter2018towards, wang2023novel}. The works of \cite{rodseth2020taxonomy , azimi2020survey, jovanovic2024review} provide comprehensive reviews for further details into the multidomain of MAS marine AI”.

We also provide a discussion on real-existing MAS in Section 3.2. - p.g. 3

***Reviewer 3****: Especially for marine vessel object detection, there exist many typical works including 10.1016/j.conengprac.2020.104458, 10.1016/j.apor.2023.103835. So, what is the advantage of the proposed method?*

**Response:**

The key novel contribution of our work is the development of a framework for testing marine AI against adversarial AI. The works you list attempt to benchmark marine vessel detection but do not consider the security implications of attacks on AI systems from adversarial AI. Further, highlighting the purpose of this paper that it is important to raise awareness of the security risks of these technologies and provide a framework that can be used to assess the risk to marine AI technologies against adversarial AI. This is not a paper about the development of marine AI, but about considering the security of marine AI from adversarial AI risk.

***Reviewer 3****: A holistic paradiagm sketch and/or pseudo-code should be provided for complete understanding for this work.*

**Response:** We have provided both a holistic paradigm sketch Figure 1 and also provided a checklist in Appendix A which is more direct and simplified without the additional test case detail surrounding it.

***Reviewer 3****: Simulation or experimental results and even comparisons should be provided to validate the ideas.*

**Response:** We provide a real experimental test case in Figures 5,6,7 to highlight how these attacks could work against MAS, the effects of the attacks and the frameworks. After validating our ideas with an experiment and taking photos of the effects we have chosen to include them in the paper as Figures 5,6,7. No numerical data exists. We are not developing any new AI technologies which need benchmarking, we have developed a methodology and have shown the effects of the methodology/validated all ideas/findings with a real-world case study.