Dear reviewers:

I sincerely thank you for your careful review and valuable comments on this paper! The issues you have pointed out are extremely constructive and crucial to improving the quality of the paper. We have carefully revised the paper one by one, and the main changes are reported as follows:

1.To talk about this topic the manuscript tries to provide a comprehensive introduction that still lacks consistency. The introduction does not account for the different types of models used or the different improvements provided by each of the mentioned models, it just mentions as a list all the models. This aspect is crucial to get the state of the art and the possible improvements led by the proposed manuscript. Therefore, a better literature review is strongly required.

105-112 are a perfect synthesis of what is required and what is currently present; however, it seems not linked well with the mentioned cited articles or with the way the cited articles have been presented.

The goal of the research could have been addressed talking about the gap and the contribution in the same point, not talking about the three gaps and then listing the three research contributions.

1)In order to make the presentation of the article clearer, we restructured the introduction to make it easier for the reader to read, and added tables to summarise the literature and improve readability.

2.Another important aspect to tackle is the connection between EV and AV. In the middle of nowhere, the authors start talking about EVs and then they disappear again. Is there any connection that authors want to emphasize?

1)We have added an elaboration of the link between electric and self-driving cars (Section 2.1, first paragraph inblue): Thanks to the deep compatibility between EV drive systems and autonomous driving architectures, they can more efficiently integrate key hardware components such as high-precision environmental perception sensors, real-time decision-making units, and by-wire actuators, providing reliable technical support for intelligent control in complex traffic scenarios.

3.The rationale behind the choice of the TTC for safety reward and for the other rewards might be better explained and detailed.

Thank you for your interest in reward function design. Time to Collision (TTC) was chosen as the core metric for safety rewards based on its combined advantages of dynamic risk quantification, multi-objective optimisation compatibility, and engineering interpretability.

TTC directly characterises the urgency of a potential collision in the time dimension by fusing the relative distance and velocity information (TTC = $\Delta d / \Delta v$), and is capable of adapting to better high-speedsafety requirements in scenarios. The advantages of TTC as a safe reward function are described in detail in literature [Zhao, R., Chen, Z., Fan, Y., Li, Y. and Gao, F., 2024. Towards robust decision-making for autonomous highway driving based on safe reinforcement learning. *Sensors*, *24*(13), p.4140.]

4.The entire chapter 3 is the core part of the manuscript. It could be more precise and clearer. Some parts are confusing and the description is not linear. Please try to include a more reader-friendly approach while describing. Not all the readers could be expert; therefore, the concepts must be designed in a friendly way.

1)We have refined the chapter titles, presented the concepts, described the algorithms in a more accessible way, and in section 3.1.1 we have detailed the inputs and outputs of the model as well as the technical routes, while we have explained the training process in detail see paragraph 4 of 3.3.1

5.Figures must be renumbered in section 4.

1)We have restructured the experimental analysis graphs and language, see (Sections 4.1, 4.2 in blue) for the results, and added new experiments with the quintic polynomial + PID + CACC from the existing method to compare with our proposed model and to analyse the data in detail.

6.The safety features introduced in the description of the model are not accounted in the results. How did you define safety benefits based on TTC?

6)The safety in the model is captured by the vehicle spacing of the fleet, whose safety benefits are integrated in the reward function of the AV, the details of which are given in the blue part of Part A in Section 2.3.

Thank you once again for your professional advice!

The above modifications have been verified through simulation experiments and updated in the manuscript, significantly enhancing the paper's completeness and practical value. If you have any other suggestions, we are willing to make further adjustments.