1. One of the potential limitations of the proposed method is its reliance on an independent depth completion network to generate ground truth visual depth. As acknowledged by the authors, this network may introduce noise into the training data, which could affect the model's overall performance. More research focusing on denoising techniques or more accurate depth estimation could further improve the robustness of this approach.

没改,去噪的工作是深度预测的,这个如果加了都可以出一篇单独的论文了

2. While the method performs well on the KITTI and Waymo datasets, it would be beneficial to see further evaluations across a broader range of datasets and environments. This would help understand the approach's generalizability, especially in more complex and varied driving conditions.

没改,两个数据集够了.顶会都是两个,在文章里说明了 nuscenes 有多个摄像头,我做的数据集都是单摄像头的

3. The paper does not provide a detailed discussion on the computational efficiency of the proposed methods, particularly in real-time applications, which are critical for intelligent vehicle systems. More research could explore optimizations that make the approach much more feasible for real-world deployment in autonomous systems.

在 KITTI 测试集部分补充了 latency

4. The title is not well selected. Moreover, it does not clearly indicate the technical contributions of the manuscript.
现在的标题是更改过得

5. - Abstract and conclusion should be written in the same way, they are totally different in concept.
- An acceptable Abstract should only contain 1/3 background and problem statement, and the remaining 2/3 should be composed of objectives, contributions, and significance.
让 GPT 修了一下,我感觉写的挺好的,不知道啥问题

6. There are several very long paragraphs! In general, one paragraph should contain only one idea.
将一些长段落拆分了一下

7. If the whole sentence refers to the reference, the references should be placed at the end of the sentence.
这个不知道咋改,我感觉我这么写也可以吧,在文章中间

8. paper structures should be added to the end of the Introductions section
我看顶会都是写在方法段落的开头

9. writing is poor in case of equations
不知道哪里有问题

10. Equations are a part of the text, so every equation must end with either . or , according to the next sentence.
这个改了,它的意思是在每个公式后面要加逗号或者句号

11. When referring to an equation, no need to call it “equation (x)” or “eq. (x), just (x) is sufficient.
没改,感觉都可以

12. Old references imply that the authors didn’t study state-of-the-art. Try to study the most recent published related papers and cite them. e.g., 2022~
现在 2022 以后的差不多 30 篇

13. Provide a detailed theoretical or mathematical justification for why the Multi-Scale Grid Attention (MSGA) mechanism improves detection performance, particularly in addressing feature disparity across different RoI regions.
深度学习给不了理论的解释, 做了消融实验

14. Include visualizations of the attention maps generated by the MSGA module, comparing them to non-attention-based methods, to clearly demonstrate how it discriminates between foreground and background pixels.
Fig.5. 里面有

15.Provide a more detailed explanation or derivation of the probabilistic post-processing (PPP) method, possibly including pseudo-code or a step-by-step breakdown, to clarify its role in improving depth estimation.
我觉得这不是个算法, 没列算法. 文章第三 C 部分专门说这个,你看看说不说的清楚

16. Explore alternative methods or strategies to reduce dependency on the depth completion network, and propose ways to mitigate the noise or errors introduced by this network during training.
还是之前的问题,这个是后续工作地,不应该展开说

17. Expand the ablation study to include a comparison between MSGA and other attention mechanisms (e.g., vanilla attention, self-attention), and explore different scales for the MSGA module.
加了对比其他的 attention 的实验

18. Consider benchmarking the proposed method on additional datasets like nuScenes or ApolloScape to demonstrate the generalizability of the approach beyond KITTI and Waymo datasets.
在实验部分说了不用 nuScenes 的原因,我觉得两个数据集够了

19. Provide an analysis of the failure cases, particularly where the method struggles with depth estimation or misclassifies objects due to occlusion or long-distance issues.
没加,我看现在很多文章都不写这个部分了, 少数写的是为了凑篇幅

20.Address potential overfitting in the model, given the discrepancy between the performance on the KITTI test set and the validation set. Applying regularization techniques or cross-validation may help improve the generalizability of the method.
感觉并不是 overfitting. 是数据分布的问题,因为在验证集和测试集都超过了别人,而不是只有验证集超过

21. Compare the proposed method to the latest transformer-based monocular 3D object detection methods (e.g., MonoDETR or DETR3D) in terms of performance and efficiency.
MonoDETR 之前就比了, DETR3D 用的是 multi-view 的方法,和我这个不是一类的

22. Justify the choice of using the Laplace distribution for modeling depth in the PPP module and provide evidence or reasoning that demonstrates its superiority over alternatives like Gaussian or uniform distributions.
这个很简单,就是和 DID-M3D 比较.他就是用的这个分布

23. Conduct a sensitivity analysis of the δ parameter used in the PPP module, and present the impact of varying δ values on the model's performance.
这个之前文章那个里就做了消融实验

24. Expand the related work section to include recent advances in monocular 3D object detection, especially methods from 2023, and clarify how the proposed approach differs or extends those works.
加了四五篇

25. Include a detailed comparison of the computational cost and latency of the method, with metrics such as inference time, model size, and FLOPs, and compare these to other approaches.
加了 latency,其他的两个并不能直接反应时延,现在大部分论文都是直接说 latency

26. Include a flow diagram that visually explains the process of generating and refining RoI features through the MSGA and PPP modules, showing the intermediate steps and how they interact.
Fig.4 的总览图里面有交互的

27. Compare the proposed method to point-cloud-based 3D object detection methods in terms of accuracy and efficiency, and elaborate on the trade-offs between using monocular methods versus LiDAR-based methods for 3D object detection.
两个都不是一个赛道,从来没看过文章这么比的