This manuscript extends edge intelligence (EI) capabilities to support efficient cooperative perception (CP) deployment in heterogeneous vehicular networks. Specifically, it interprets the closed-box DNN into two explicitly interpretable and orthogonal components, namely perception coverage and feature quality, and then formulates the synchronization-efficient cooperative perception (SECP) problem. The problem is further decoupled into two subproblems, coverage-aware node selection (CAS) and heterogeneity-aware pruning with bandwidth allocation, which are solved using a greedy algorithm and alternating optimization (AO), respectively. The topic is timely and of potential interest to the community. However, I have several major concerns that should be carefully addressed.

1. The most critical weakness of this manuscript is that although multiple vehicles are considered and each vehicle is equipped with a DNN in the V2X system, the proposed algorithm only addresses CP capacity from the perspective of a single vehicle. This significantly undermines the potential of the considered V2X setting, especially given the focus on “Edge-Intelligent Cooperative Perception in Heterogeneous Vehicular Networks.” Can the proposed approach be extended to support multiple vehicles simultaneously?
2. The idea of decomposing a closed-box DNN into two interpretable and orthogonal components is interesting. However, this relies on the critical assumption that DNN weights follow a zero-mean Gaussian distribution. This assumption is central to the subsequent derivation, yet the manuscript does not provide sufficient justification or references. Please elaborate on the rationale behind this assumption or cite appropriate references.
3. In Section IV-A, it is stated that V2X communication is enabled only when the distance between two vehicles is below a certain threshold, which is indeed practical. However, this constraint is not incorporated into the optimization problem (7). Please clarify why it is omitted or explicitly include this constraint in the problem formulation.
4. Another concern regarding problem (7) is its dependence on the priori knowledge of vehicles’ locations. If the vehicles’ locations are already known, what is the necessity of the subsequent detection process? This raises a suspected circular assumption. Please revise the formulation or clarify the rationale to avoid this issue.
5. While the problem formulation is interesting, the solution approach seems rather incremental. The use of projected gradient descent is straightforward but suffers from slow convergence and suboptimal solution quality. More advanced optimization methods such as SCA or ADMM could be considered to improve solution performance.
6. In Fig. 7c, as the node selection threshold increases from zero to one, the response delay (RD) decreases. This is counterintuitive, since involving more cooperative vehicles would generally increase RD. This phenomenon should be carefully explained.
7. In Fig. 12, the benchmark FedMP achieves similar network utility to the proposed method, and the utility appears to depend heavily on the weight factor phi. This raises concerns regarding the actual advantage of the proposed method. Please include additional results showing the variation of network utility with respect to phi to clarify this point.
8. The literature [1] W. Xu, Z. Yang, D. W. K. Ng, M. Levorato, Y. C. Eldar, and M. Debbah, “Edge learning for B5G networks with distributed signal processing: Semantic communication, edge computing, and wireless sensing,” IEEE Journal of Selected Topics in Signal Processing, vol. 17, no. 1, pp. 9–39, Jan. 2023 should be cited to provide a more comprehensive overview of edge intelligence.
9. In Fig. 1, the manuscript claims that due to occlusion by a large truck (vehicle B), vehicle A and C cannot detect each other, which may lead to a collision. However, since the truck is itself a movable vehicle, collisions are more likely between A and B or C and B, rather than A and C. A static obstacle (e.g., a billboard) would be a more appropriate example.
10. The styles of Figs. 9, 10, and 13 are unclear and overly dark. The vehicles in the illustrations, particularly in Fig. 13, are barely visible. I recommend adopting a clearer style similar to Fig. 11.