Summary

This paper proposed a novel method to solve point matching problem using the graph-based method. The feature of each point is generated by the position of the point and the positions of its neighbors. Then, the feature is updated by the GNN generated by the position of the point. Last, the corresponding points is got by the cross-attention score and Sinkhorn algorithm. The experiment shows that the method performs much better than current methods.

Main Review

Strengths:

1. Very good results. The experiments show that the results of the work are far better than SOTA results.
2. The method is quite simple and seems work efficiently compared with CorrNet3D.

Weakness/ Detailed Comments:

Writing:

1. I am confused about the expression of the initial feature in III.A. I think if what you wants to express is the same as [1], you can express the same as [1] to make it clear. If what your express is not the same, please make it clear and let me know the difference between them.
2. The equation (3) has a typo. I think the correct of the last of the equation should be $\frac{\sum j \neq i g_{msg}(f_{x_j}^{local})}{\sum_{j \neq i} \epsilon_{x_j}}$.
3. I think the reference [31] is not correct and I do not find any useful information. Is the paper [1] the correct reference.

Approach:

1. There exists a problem is that if the point is not shown in the second graph, the points will match with other points. I wonder if it is necessary to add a dustbin like Superglue [2] which is a method for the point matching by descriptor to prevent this situation. Could you provide why you not choose to use the dustbin.
2. Based on my reading, the graph generating method is the key that this work perform better than SOTA method for it uses the geometry information of the point. Some methods like K-NN and 3D mesh with transformer can have the same concept. Therefore, the comparison of this methods should be considered in ablation study.
3. From the experiment part, it seems that the unsupervised CorrNet3D perform better than supervised CorrNet3D. So why not consider to extend this work to the unsupervised version?
4. Could you provide more details, such as the code, on why and how this method is so efficient?

Rating: 6

References

[1] Yue Wang, Yongbin Sun, Ziwei Liu, Sanjay E Sarma, Michael M Bronstein, and Justin M Solomon. Dynamic graph cnn for learning on point clouds. *Acm Transactions On Graphics (tog)*, 38(5):1–12, 2019.

[2] Paul-Edouard Sarlin, Daniel DeTone, Tomasz Malisiewicz, and Andrew Rabinovich. Superglue: Learning feature matching with graph neural networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 4938–4947, 2020.