Summary:
This paper proposes a multi-modal LLM, called any-to-any MM-LLM, to extend the multi-modality of LLM to a state where there is no limitation on the input and output modality combinations. To achieve this goal, the authors (1) propose a lightweight alignment learning technique to achieve en effective semantic alignment across different modalities with limited trainable parameters and (2) annotate a modality-switching instruction tuning dataset. The displayed results and visualizations suggest the promising performance of the tuned any-to-any MM-LLM.

Soundness: 3 good
Presentation: 4 excellent
Contribution: 3 good
Strengths:

- Extending the multi-modal LLMs free of limitation on the input/output modalities is an important research question that can facilitate a wider range of applications.
- The introduced dataset, if made publically available, would be a good contribution to the community.
- Various evaluation benchmarks are used to benchmark the proposed model with existing solutions.
- The writing is clean and easy to follow

Weaknesses:

- The proposed alignment learning technique is a bit naive and does not consider much about the challenge introduced by the any-to-any modality, such as how to balance the performance across different modalities.
- Although introducing contents from different modalities during tuning is considered to improve the overall performance of the model, in the experiment section, it seems introducing these additional modalities actually leads to worse performance on benchmarking datasets. Does this indicate the alignment technique is not effective enough as expected?

Questions:
Will the pretrained model and dataset be released to the public?

Flag For Ethics Review: Yes, Discrimination / bias / fairness concerns, Yes, Privacy, security and safety, Yes, Responsible research practice (e.g., human subjects, data release)
Details Of Ethics Concerns:
The proposed datasets's content may need a deeper look from experts to check its content. And the content generated by the model may need further checking to make sure there are no harmful contents generated.

Rating: 6: marginally above the acceptance threshold
Confidence: 4: You are confident in your assessment, but not absolutely certain. It is unlikely, but not impossible, that you did not understand some parts of the submission or that you are unfamiliar with some pieces of related work.
Code Of Conduct: Yes