

11-877 Advanced Topics in Multimodal Machine Learning

Week 9: Brain and Multimodal Perception

Due date: 11PM EST, Wednesday, March 16 2022

Submission: <https://forms.gle/7cHDXywYMdhHBPZD8>

We designed the reading assignments to help you prepare for the live discussions. Discussion probes were drafted related to this week's topic. These were written to help conceptualize the problem and guide your thought process. Take the time to read them first. The goal is not to answer each of these questions and probes individually, but they are meant to be taken as a whole. We also selected research papers relevant to this topic. Required papers should be read completely. Suggested papers should at least be skimmed. The purpose of the reading assignment is to start your critical thinking process, so your responses should demonstrate constructive thoughts, with a good understanding of the current research in this area, and express your own insights.

Your response to this reading assignment should be submitted in the online Google Form (see link above). Your response should consist of four main components:

- (1) **Scouting:** As you start thinking about the discussion probes, it is always good to also scout papers, blog posts, and other resources related to the topic. We ask that you search for related resources and share with us 2 extra links to these new resources. For each extra link, include 1-2 sentences explaining the value and relevance of this extra resource.
- (2) **Reading notes:** As you read the required papers, suggested papers, and the extra resources you scouted, please write down at least 4-6 notes related to the discussion probes. Each note should be 1-3 sentences long. These can be empirical results you observed, ideas or theories expressed by other researchers, or any interesting fact that is worth noting when summarizing your reading.
- (3) **Your thoughts:** Separate from your reading notes, we ask that you reflect more holistically about the discussion probes. Please write 3 discussion points you would like to share on this topic. Each discussion point should be one paragraph (3-5 sentences). These discussion points should go beyond the reading papers, and try to address as many aspects of the discussion probes as you can. We do not expect that you answer all discussion probes. For example, it would be ok to focus on only 1 or 2 probes if these bring the most ideas and thoughts for you.
- (4) **Clarification requests [OPTIONAL]:** Please take a moment to suggest parts of the papers where clarifications would be useful. Try to be as specific as possible in your clarification requests. These requests will be shared with the Reading Leads in charge of creating a short presentation for the beginning of Friday's course and answering other requests directly on Piazza.

Week 9 discussion probes:

- What are the main takeaways from neuroscience regarding unimodal and multimodal processing, integration, alignment, translation, and co-learning?
- How can these insights inform our design of multimodal models, following the topics we covered previously (cross-modal interactions, co-learning, pre-training, reasoning)?
- To what extent should we design AI models with the explicit goal to mirror human perception and reasoning, versus relying on large-scale pre-training methods and general neural network models?
- What different paradigms for multimodal perception and learning could be better aligned with how the brain processes multiple heterogeneous modalities?
- How does the human brain represent different modalities (visual, acoustic)? Are these different modalities represented in very heterogeneous ways? How is information linked between modalities?
- What are several challenges and opportunities in multimodal learning from high-resolution signals such as fMRI and MEG/EEG?
- What are some ways in which multimodal learning can help in the future analysis of data collected in neuroscience?

Required papers (you should read these papers in detail)

- [The Neurophysiological Foundations of Mental and Motor Imagery. chapter 1: Multimodal Images in the Brain](#) (pages 3-16)
- <https://www.sciencedirect.com/science/article/pii/S0010945217302277>

Suggested papers (you should skim through these papers, at the minimum)

- <https://academic.oup.com/cercor/article/11/12/1110/492310>
- <https://arxiv.org/abs/1711.07998>
- <https://arxiv.org/abs/2011.09850>
- <https://arxiv.org/abs/1911.03268>

Other relevant papers:

- <https://aclanthology.org/2020.lrec-1.85.pdf>
- <https://nobaproject.com/modules/multi-modal-perception>
- <https://arxiv.org/abs/1810.10974>
- http://userpage.fu-berlin.de/rmcichy/publication_pdfs/Cichy_et_al_CC_2016.pdf
- <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.269.2537&rep=rep1&type=pdf>, chapter 5
- https://www.cs.cmu.edu/~hyunahs/papers/SDM2017_1.pdf
- <https://arxiv.org/abs/1905.11833>