

11-877 Advanced Topics in Multimodal Machine Learning

Week 14: Multimodal Generation and Ethical Concerns

Due date: 11PM EST, Wednesday, April 20 2022

Submission: <https://forms.gle/eQpkJL7uzxsmPHaf8>

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 14 discussion probes:

- What are some challenges in multimodal generation beyond generating each modality individually? How can we synchronize generation across multiple modalities?
- What degree of multimodal modeling is required for these cross-modal generation to be possible? For example, how much do models need to learn regarding cross-modal interactions, alignment, reasoning, etc?
- What are the qualities we should consider when evaluating outputs from multimodal generation? What do you think is the best practice to evaluate these qualities? Can we efficiently evaluate these qualities, at scale?
- What are the opportunities and challenges of automatic and human evaluation? How can we combine the best of both worlds?
- What are the real-world ethical issues regarding generation? How are these risks potentially amplified or reduced when the dataset is multimodal, with heterogeneous modalities? Are there any ethical issues that are specific to multimodal generation?
- *How can we build a taxonomy of the main ethical concerns related to multimodal generation?*
- How can we update our best practices to help address these ethical concerns? Who is better placed to start this dialogue? How can we make significant changes in this direction of reducing ethical issues?

Required papers (you should read these papers in detail)

- <https://arxiv.org/abs/2102.10407>
- <https://openai.com/blog/dall-e/> and <https://openai.com/dall-e-2/>
- <https://dl.acm.org/doi/abs/10.1145/3442188.3445922>

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

- <https://www.liebertpub.com/doi/full/10.1089/cyber.2021.29208.jth>
- <https://www.theverge.com/21298762/face-depixelizer-ai-machine-learning-tool-pulse-stylegan-obama-bias>
- <https://arxiv.org/abs/2005.03201>
- <https://arxiv.org/abs/1905.12616>

Other relevant papers:

- <https://thegradients.pub/pulse-lessons/>
- <https://arxiv.org/abs/2011.03775>
- <https://aclanthology.org/D18-1084.pdf> <https://arxiv.org/abs/1810.05795>
- <https://arxiv.org/abs/1906.07901>
- <https://arxiv.org/abs/2012.07805>
- https://openaccess.thecvf.com/content_iccv_2015/papers/Suwajanakorn_What_Makes_Tom_ICCV_2015_paper.pdf
- <https://arxiv.org/abs/1710.00421>