HIVE: Harnessing Human Feedback for Instructional Visual Editing

Shu Zhang*, Xinyi Yang*, Yihao Feng†, Can Qin†, Chia-Chih Chen†, Ning Yu†, Zeyuan Chen†, Huan Wang†, Silvio Savarese1,2, Stefano Ermon2, Caiming Xiong3, Ran Xu†

*Equal Contribution †Salesforce AI Research ‡Stanford University

Background
- Instructional image editing has emerged as one of the most promising application scenarios for content generation. We hypothesize that instructional image editing could benefit from human feedback, as their outputs may not adhere to the correct instructions and preferences of users.

Contribution
- To tackle the technical challenge of fine-tuning diffusion models using human feedback, we introduce two scalable fine-tuning approaches, which are computationally efficient and offer similar costs compared with supervised fine-tuning. Moreover, we empirically show that human feedback is an essential component to boost the performance of instructional image editing models.
- We create a new dataset for HIVE including three subdatasets: a new 1.1M training dataset, a 3.6K reward dataset for rewards learning, and a 1K evaluation dataset.
- We introduce cycle consistency augmentation based on the inversion of editing instruction. Our dataset has been enriched with one pair of data for bi-directional editing.

Proposed Methods
- **Step 1:** Instructional supervised training
  - Collect data to fine-tune GPT-3 and use fine-tuned GPT-3 to generate text edits

- **Step 2:** Collect comparison data, and train a reward model
  - Collect a reward dataset and generate sampled outputs from Step 1

- **Step 3:** Fine-tune diffusion model with learned rewards
  - Use the learned reward model to calculate reward values for each training pair

Quantitative Results
- Comparisons between InstructPix2Pix(IP2P) and HIVE. Illustration of trade-offs between consistency with the input image and with the edit. HIVE achieves higher similarity on both metrics.

Qualitative Results
- User study of comparison between InstructPix2Pix(IP2P) and HIVE. HIVE obtains 29% more votes.

More examples of HIVE with human feedback.