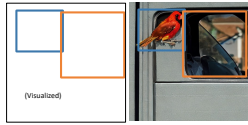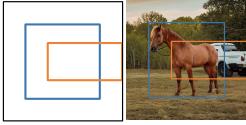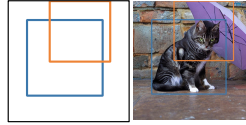# Attribute Binding

Input (to LayoutGPT): A red bird fluttering next to a car window.
Output (of LayoutGPT):
a red bird with wings fluttering
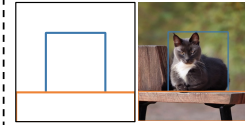a car window with the window open

(Visualized)

Input: A brown horse standing next to a white truck.
Output:
a brown horse standing still
a white truck with four black wheels

Input: A gray and brown cat sitting under a purple umbrella.
Output:
a gray and brown cat in a sitting position
a purple umbrella

Text-based Inpainting

Input: Grey and black cat sitting under a bench.
Output:
a grey and black cat with its tail tucked in and its head down, looking at the viewer
a wooden bench with two legs

Input: A cat is sitting on a basket under a bench.
Output:
a black and white cat with its tail curled up in the basket
a light brown basket with a handle on the side
a wooden bench with a curved backrest
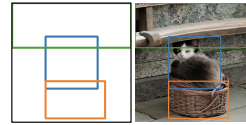
Input: A cat is lying on top of a suitcase.
Output:
a black and white cat lying on its side with its eyes closed
a purple suitcase with a design of a blue, yellow, and white flower
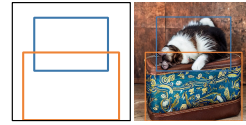
Figure 1: (Left) LayoutGPT correctly binds attributes to each **bounding boxes**. (Right) LayoutGPT can creatively expand the local descriptions beyond global prompts. Bounding boxes with colors correspond to sentences with the same color for each example. For each image pair, the left is LayoutGPT while the right one is LayoutGPT+ReCo.

"A bedroom with a double bed, two wardrobes and a pendant lamp."
Room Size: max length 262px, max width 207px

"A bedroom with a double bed, a wardrobe, two tables, a cabinet and a pendant lamp."
Room Size: max length 214px, max width 306px

"A living room with three coffee tables, a tv stand, a multi seat sofa, a dining table, four dining chairs, a console table and two pendant lamps."
Room Size: max length 256px, max width 490px

"A living room with a multi seat sofa, a coffee table, a tv stand, a dining table, four dining chairs, a ceiling lamp and a pendant lamp."
Room Size: max length 413px, max width 256px

Figure 2: When conditioned on captions that enumerate the furniture, LayoutGPT faithfully generate all objects' layout without missing or hallucination.
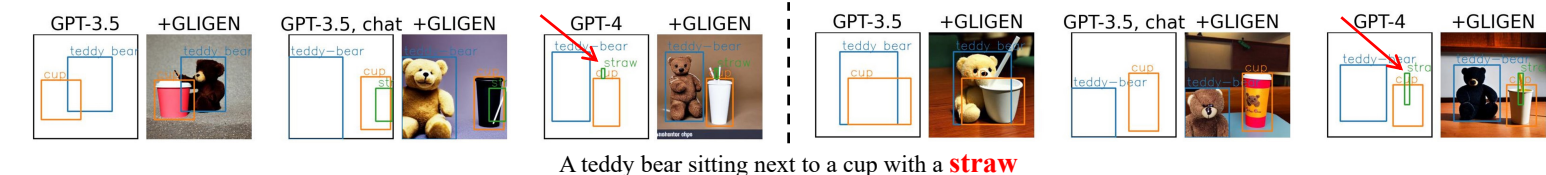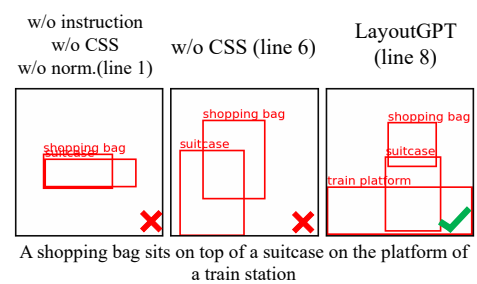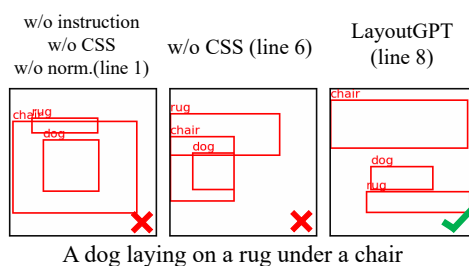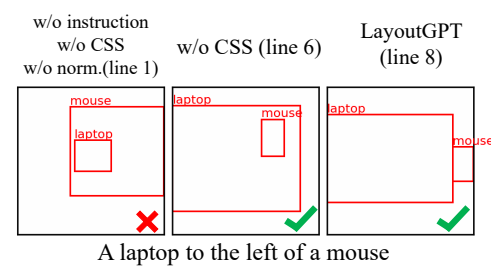

Figure 3: In-depth analysis on the 2D reasoning ability of LayoutGPT. (Top) All components (instruction, CSS, normalization) matter for layout generation. (Bottom) GPT-4 is potentially much stronger in spatial understanding and reasoning beyond quantitative superiority on NRS-1K.
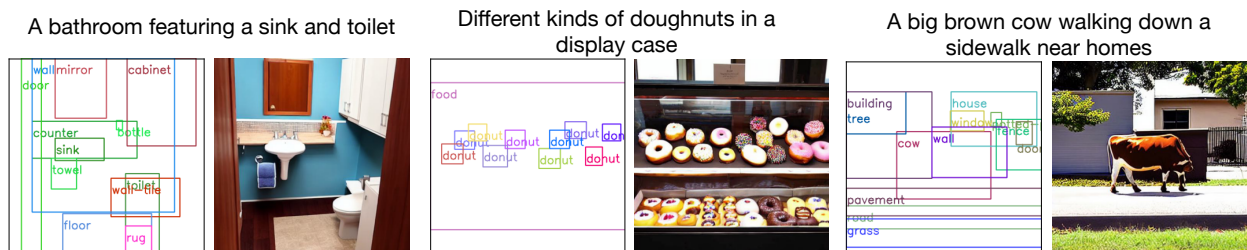

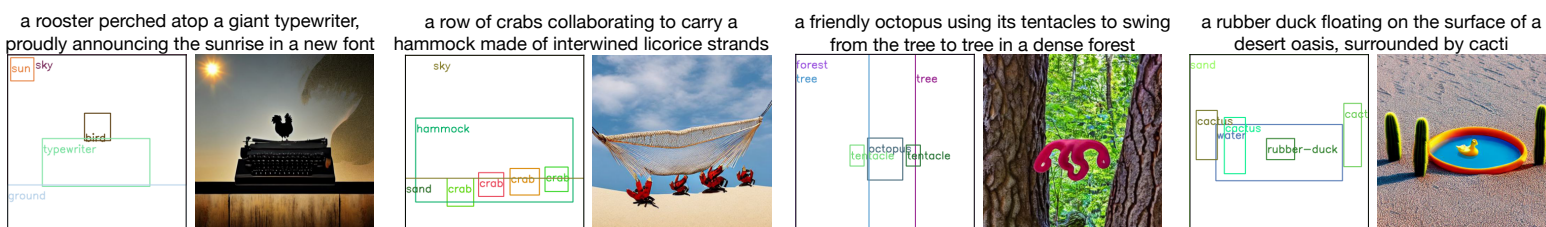Figure 4: LayoutGPT performing dense layout planning on MSCOCO2017 Panoptic task.

Figure 5: LayoutGPT's performance on counterfactual prompts provided by ChatGPT.