(a) Pre-trained on high-resolution Stable Diffusion 2.1 significantly enhances the legibility of small text.
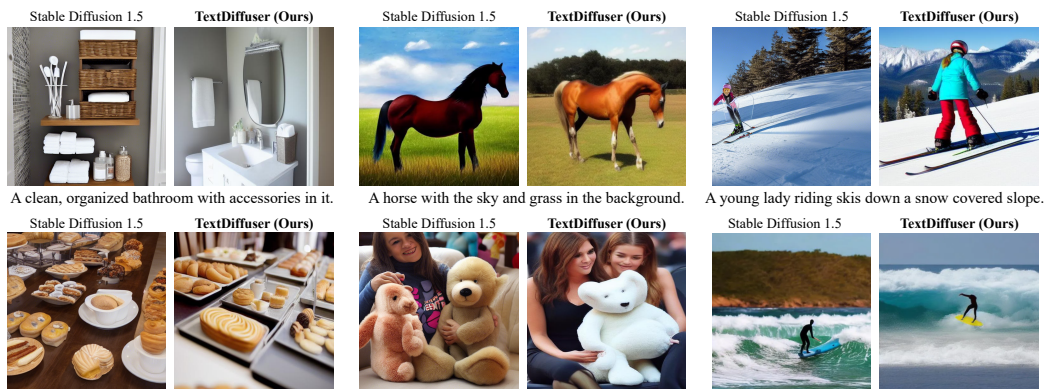
a bear holding a board with a red 'hello world'   a bear holding a board with a yellow 'hello world'   a bear holding a board with a purple 'hello world'

(b) Demonstration of using language descriptions to control the style of text.

A clean, organized bathroom with accessories in it.   A horse with the sky and grass in the background.   A young lady riding skis down a snow covered slope.

A table with servings of pastries on trays.   A Daily Mirror featuring stories on celebrity stuffed animals.   A man riding on top of a wave on a surfboard.

(c) Visualizations of general images generated by Stable Diffusion 1.5 and the proposed TextDiffuser.

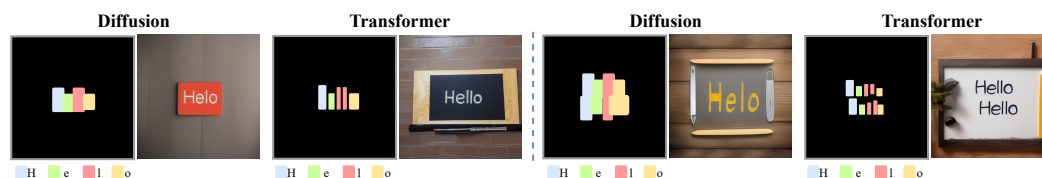A vintage postage stamp with the message: california all dreams welcome   A sign that says "Please refrain from arguing with the chimpanzees"   The cover for the album "Elusive Interludes" by the band The Melting Snowmen.

Photo of "Gas Station"   a 3d model of a 1980s-style computer with the text "my old habit" on the screen   A t-shirt with the message "There is no planet" written on it

(d) Comparison with Character-Aware Model [42] and the *concurrent* GlyphDraw [45]. The code is not available for [42], the checkpoints and datasets are not available for [45]. So we compare the samples in their papers.

Text 'Hello' on a board   Text 'Hello Hello' on a board

(e) Layout generation using diffusion model and Transformer.