R-Fig1: Qualitative comparison with two recent text-driven video editing methods: TokenFlow and VidToMe. It is difficult for text-driven methods to add or delete objects. Besides, our I2V-Edit-Benchmark can be expanded by giving each video more edited first frames.