Author Rebuttal for An Image is Worth 32 Tokens for Reconstruction and Generation

Figure 1: Uncurated 256×256 TiTok-L-32 samples. Class labels (class ids) from left to right and top to down are: "macaw" (88), "volcano" (980), "coral reef" (973), "white wolf" (270).
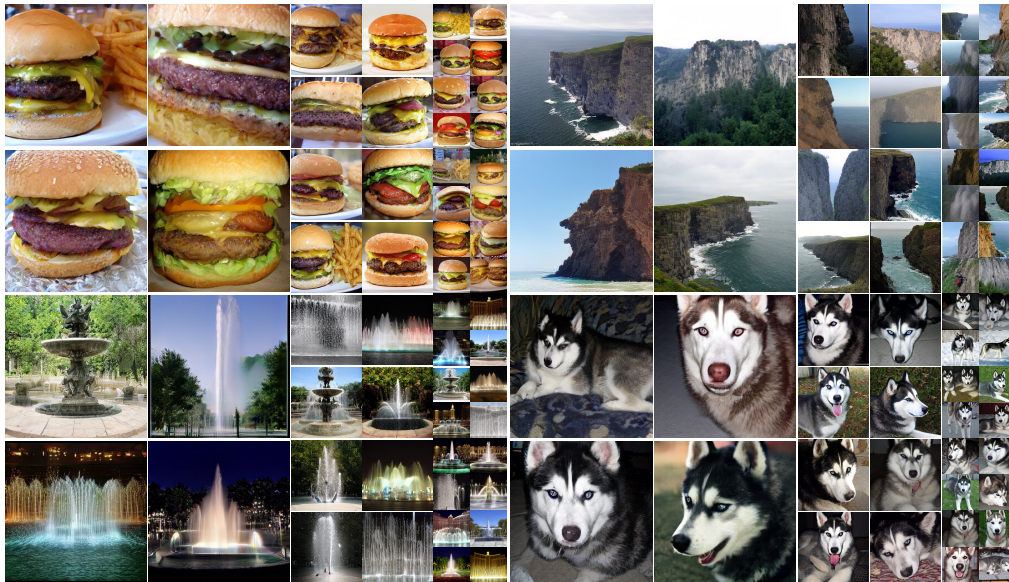

Figure 2: Uncurated 256×256 TiTok-L-32 samples. Class labels (class ids) from left to right and top to down are: "cheeseburger" (933), "cliff" (972), "fountain" (562), "Siberian husky" (250).