

TORÉ: Token Recycling in Vision Transformers for Efficient Active Visual Exploration – Supplement

November 26, 2024

We provide examples of full exploration trajectories performed by our method as illustrated in [1](#). Although image reconstruction is of poor quality, it is sufficient to guide the exploration process and results in the high accuracy of our method on the AVE task.

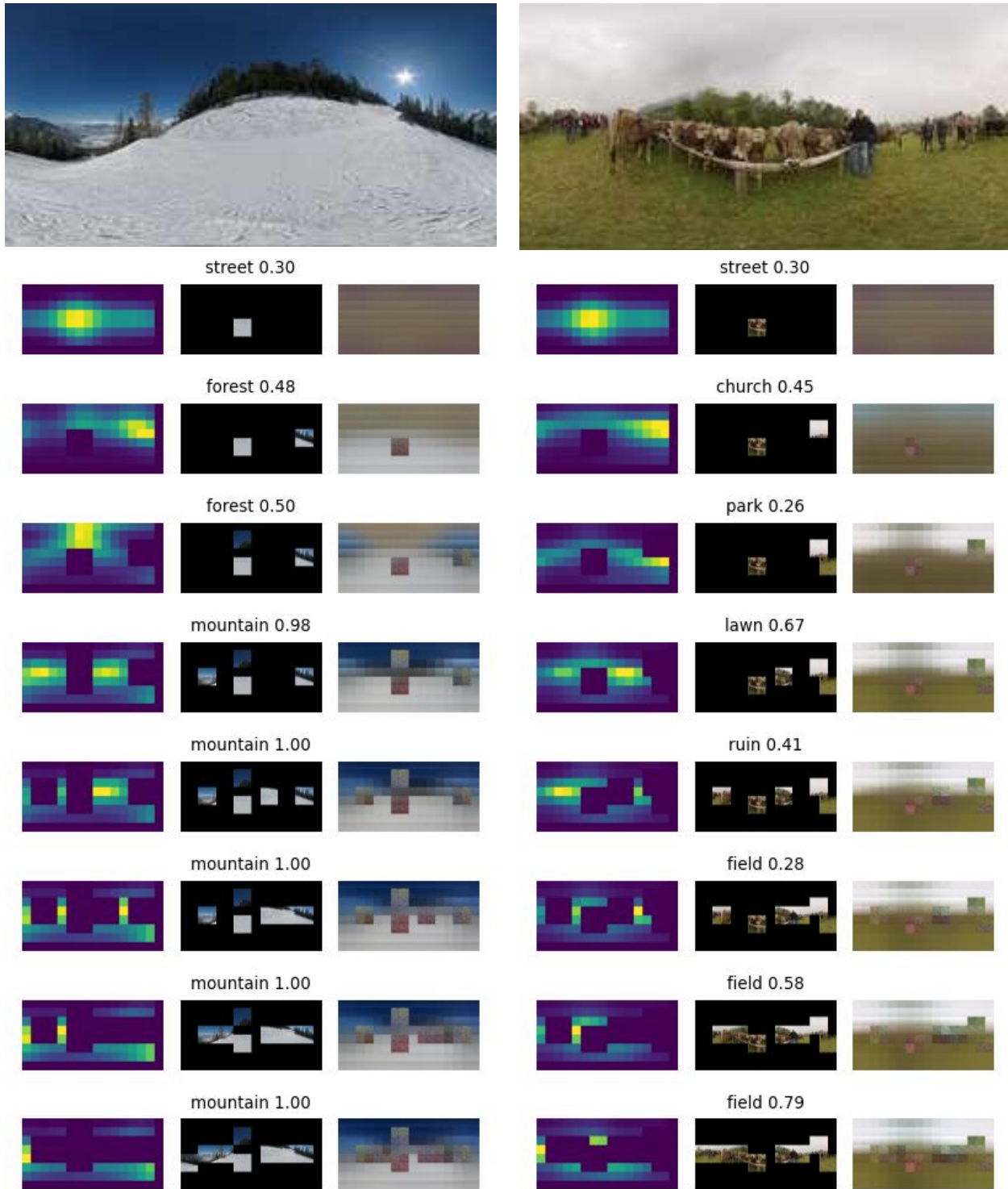


Figure 1: Visualization of active visual exploration performed by TORE. For each image, in the first column, we present entropy maps used by the selection policy, in the second column, the model successive fields of view, and finally in the last column we show the image reconstructions at each timestep. Note that we train TORE using masked reconstruction loss, therefore the model is not required to produce meaningful reconstructions for visible patches.