

This paper proposes the novel task of video generation conditioned on a SINGLE semantic label map, which provides a good balance between flexibility and qu