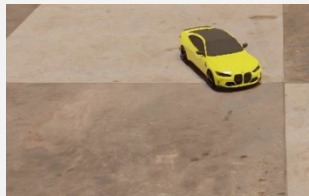


1-*th* Frame



Trajectory Mask



Traj. Guider

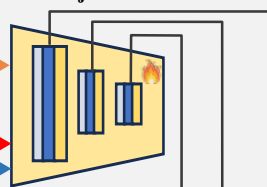
Random Mask



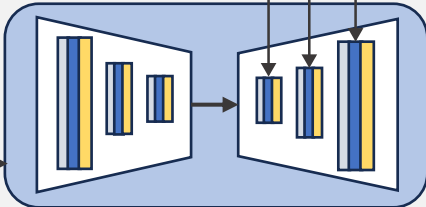
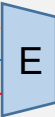
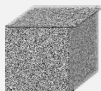
Camera Poses



Traj-ControlNet



Latent Noise



SVD



Video Sequence Frame



--- Stage One Pretraining

--- Stage Two Pretraining

--- Finetuning