

Input Videos

Control Branch

Channel-Concat

Denoising DIT

Output Video

Counterfactual
Video

Counterfactual
Tracks

Target
Tracks

