

透過 AnimateDiff 與 ControlNet 實現可控的文字轉影片生成

Controllable Text-to-Video Generation via AnimateDiff and ControlNet

指導教授：蕭宏章

專題成員：張弈鉸

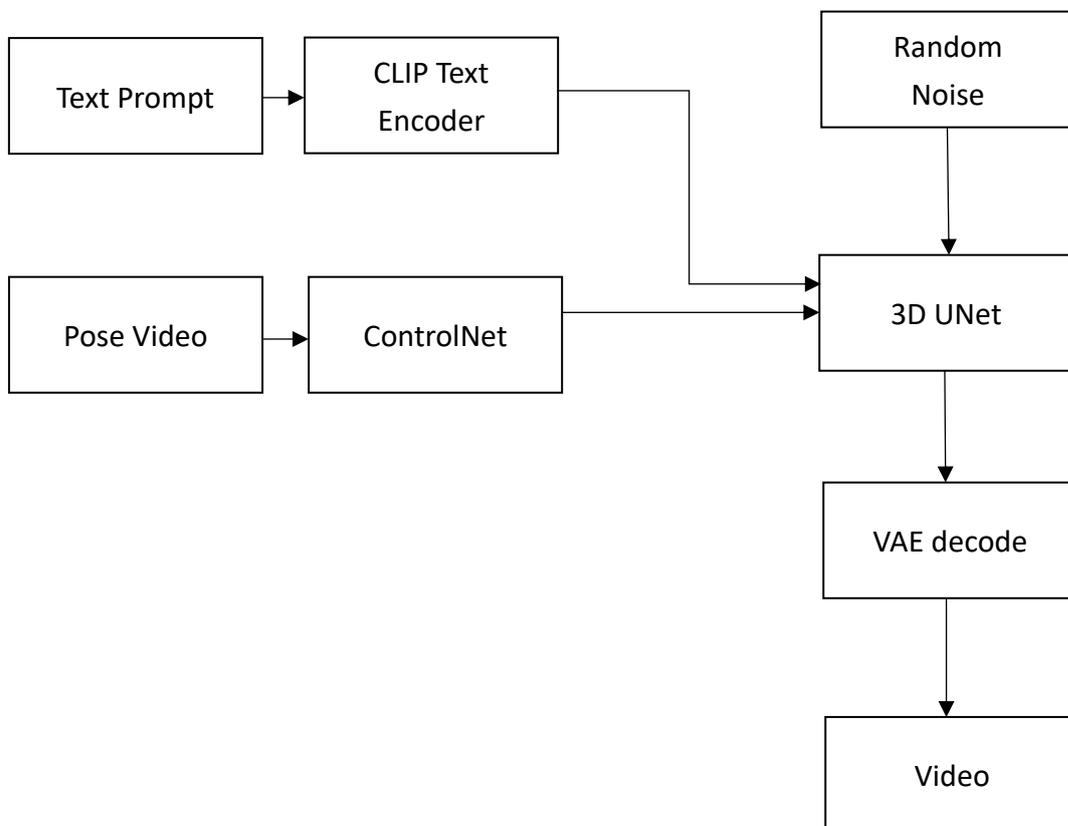
開發工具：Python、PyTorch

測試環境：Linux

一、簡介：

此專案目的在於解決生成動畫時難以精準控制生成內容的問題，透過融合 ControlNet 提供的控制條件 (softedge)，使用者可以對 AnimateDiff 生成的每一幀畫面做明確的控制，進而達到預期的動畫。

以下為系統架構圖：



二、測試結果：

文字提示：1girl, looking at viewer, upper body, contemporary, dress



說明：上方四張照片為我自行錄製的 pose video 畫面，下方為使用該影片作為 ControlNet 輸入條件後，搭配文字提示所生成的動畫畫面。透過 ControlNet 的姿勢引導，可以明確控制角色的動作，達到動畫動作與原始影片對齊的效果。