

111 學年度大學部專題競賽



國立清華大學資訊工程學系

Department of Computer Science, National Tsing Hua University

影片人臉風格轉換研究

范姜晴、黃竑翔、郭珈源

簡介

目前2D的人臉風格轉換技術已經非常成熟，而我們想深入了解影片上的風格轉換，於是研究並比較目前發表的影像生成、風格轉換技術，研究其中原理、架構並設計實驗。

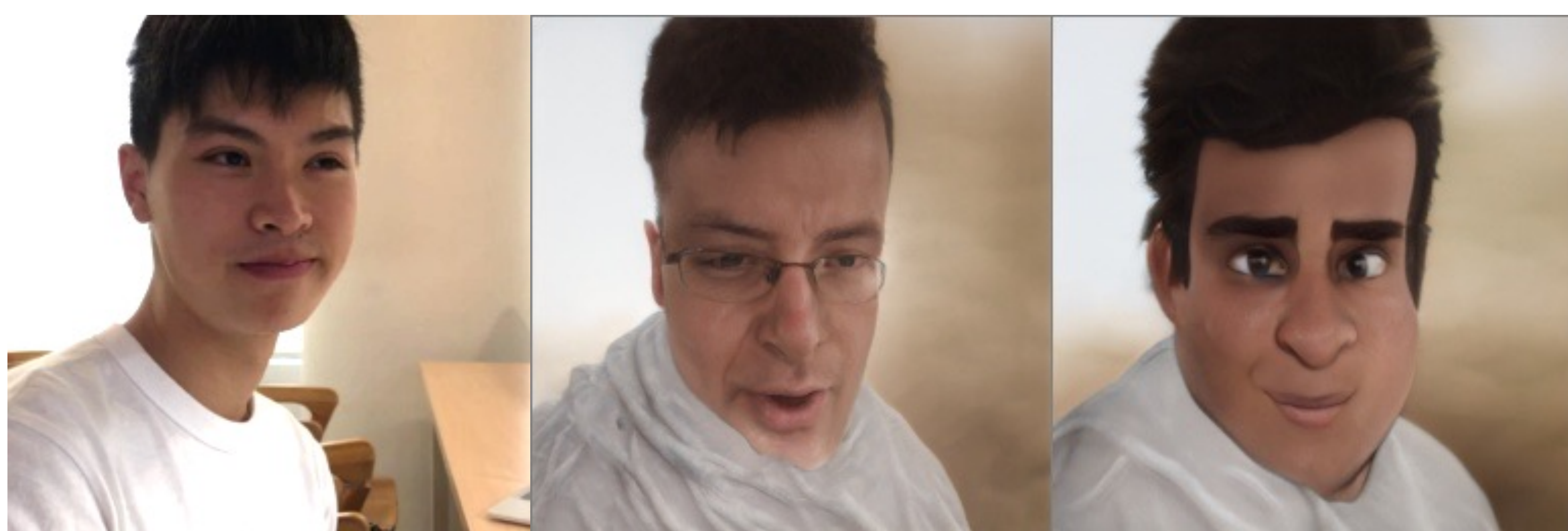
我們使用能產生高解析度及風格控制彈性的DualStyleGAN作為第一個研究的架構。

為了找到更適合影片風格轉換的技術，我們研究了第二個架構VToonify。

DualStyleGAN實作內容

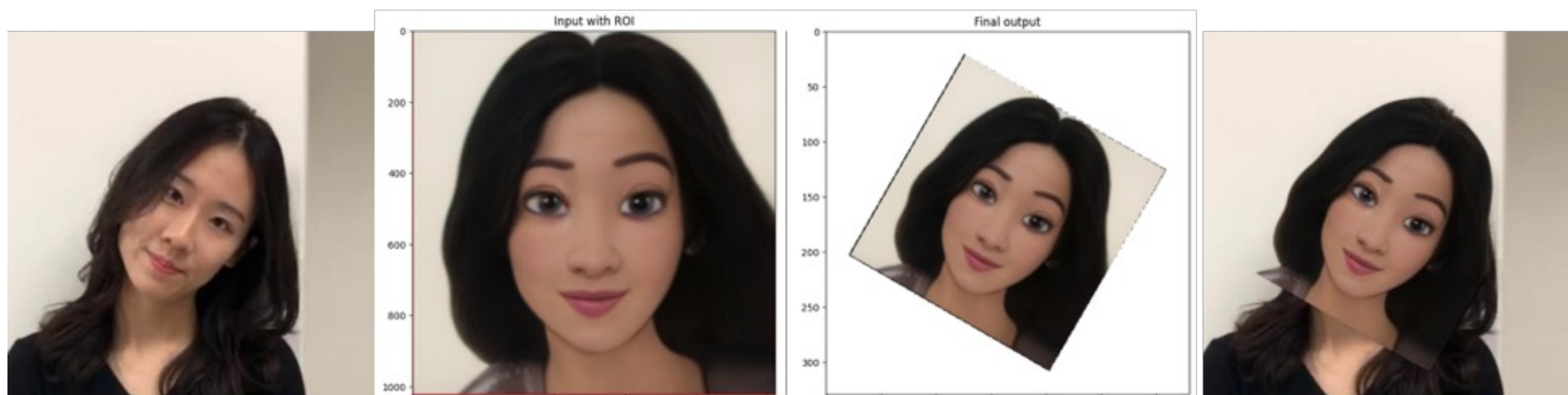
DualStyleGAN受限於 StyleGAN-based framework 的共同問題 fixed input size 以及 face alignment，轉換後的後的frame跟frame之間畫面不連續，並不適合用在動態影片上做人臉風格化。我們設計了一些實驗設法解決畫面閃爍的問題。

實驗1：將原圖不經過 face alignment 直接丟進 encoder，但由於encoder會根據圖片的相對位置取得人臉特徵，而未經align的face會誤導encoder，取到錯誤的68 facial landmarks，導致轉換結果異常。



(a) 原圖不經過alignment (b) 經過pSp encoder (c) DualStyleGAN output

實驗2：將轉換後的圖片旋轉平移、去背再貼到原圖上，但會有銜接不自然的問題。

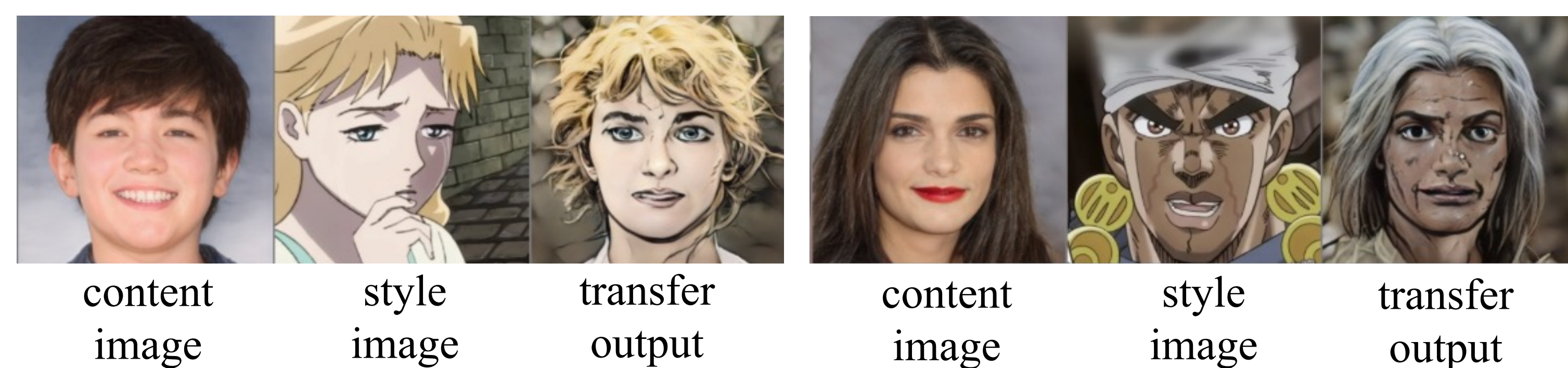


(a) 原圖 (b) DualStyleGAN output (c) 旋轉平移回原本的位置 (d) 去背後貼回原圖上

實驗3：希望利用取得原始影片frame跟frame之間的 optical flow，並將first frame進行style transfer，再將optical flow套在轉換好的frame上，依序得到後面的平滑transfer frame。但嘗試了幾個模型發現品質並不穩定，再者並無適合之model提供image與video進行相對關係之比對。

DualStyleGAN 新風格訓練

使用作者提供的face detection model及 training model將動漫影片中角色們的臉擷取下來並訓練成一套新的風格。

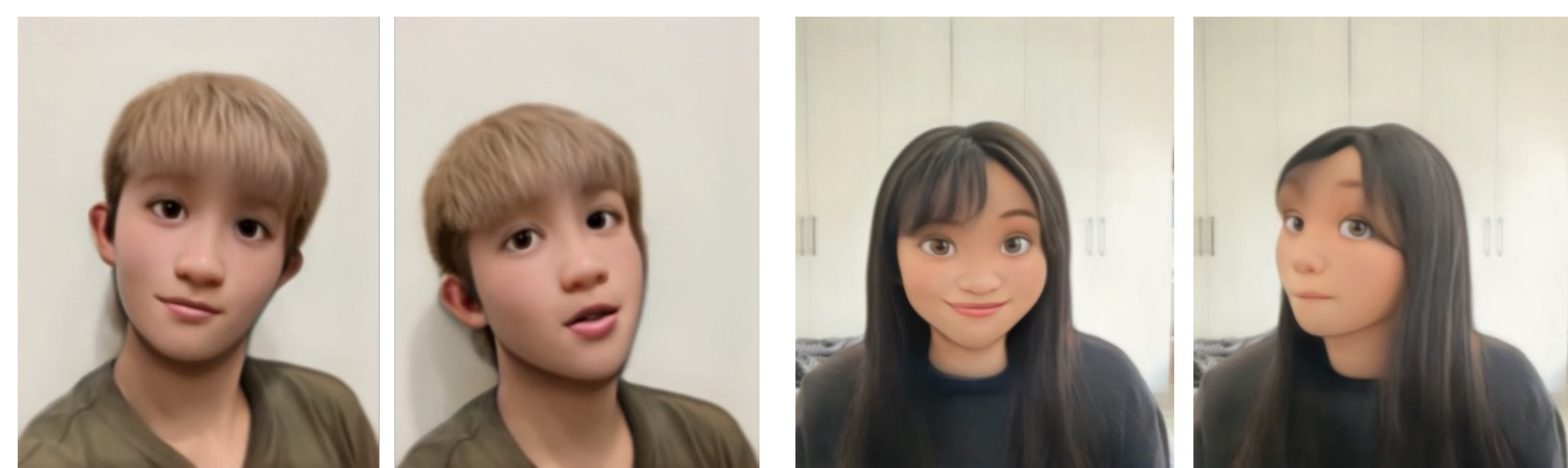


VToonify架構上的改進

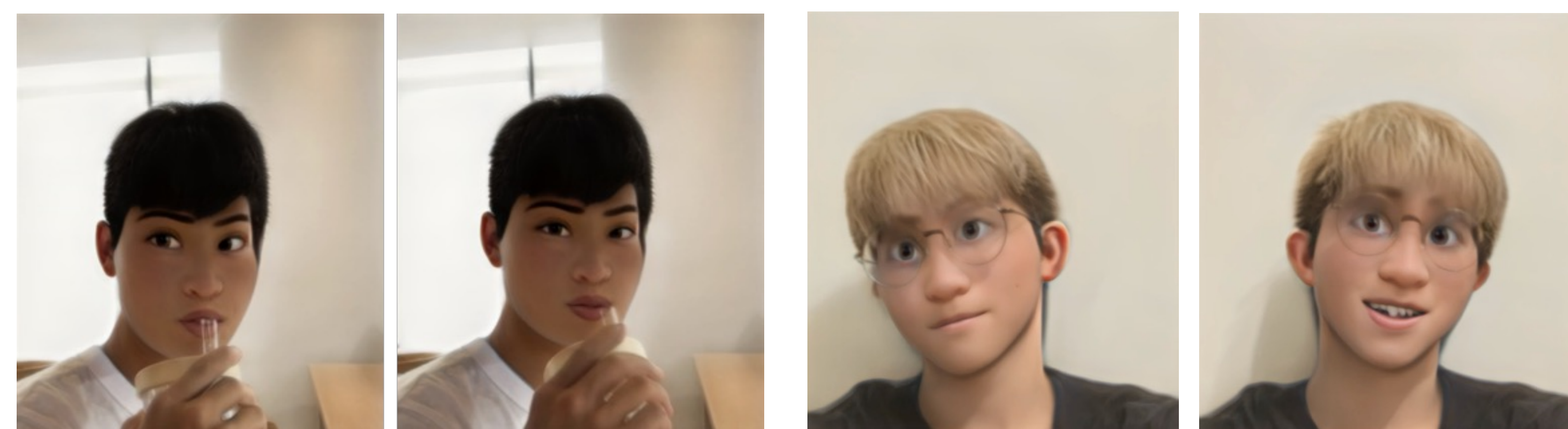
1. 支援不同的影像輸入尺寸
2. 改良StyleGAN fixed-crop limitation，支援 unaligned face
3. 利用 multi-scale content condition 及 flicker suppression loss 改善閃爍問題，維持畫面連續性
4. 繼承DualStyleGAN的優點，彈性控制結構、顏色上的轉換程度

基於以上，VToonify比先前的架構都更適用於影片上的肖像風格轉換。

VToonify成果展示



大部分的臉部角度轉換順暢，但極端角度時會出現不自然現象



在臉部前的手指變的透明，及眼鏡周圍有點模糊

結論

我們透過研究各個影像生成、風格轉換之技術，認識Machine Learning不同類型的學習模式，學習運用許多Python library、OpenCV的影像處理功能，了解整個影像風格轉換流程、架構，在現有模型中找出問題並設計實驗來驗證、改善問題。