



基於雙向生成對抗網路的臉部年齡轉換程式

Facial Age Transformation Based on Modified BiGAN

組別：5-5

指導教授：連震杰 教授

專題成員：趙珈葦、白耀升

Outline

1. 專題動機與目標

1.1 動機與目標

1.2 專題貢獻

2. 模型介紹

2.1 BiGAN 架構 & 訓練流程

2.2 Modified BiGAN

3. 實驗結果

3.1 訓練資料集

3.2 實驗結果

3.3 實驗優化



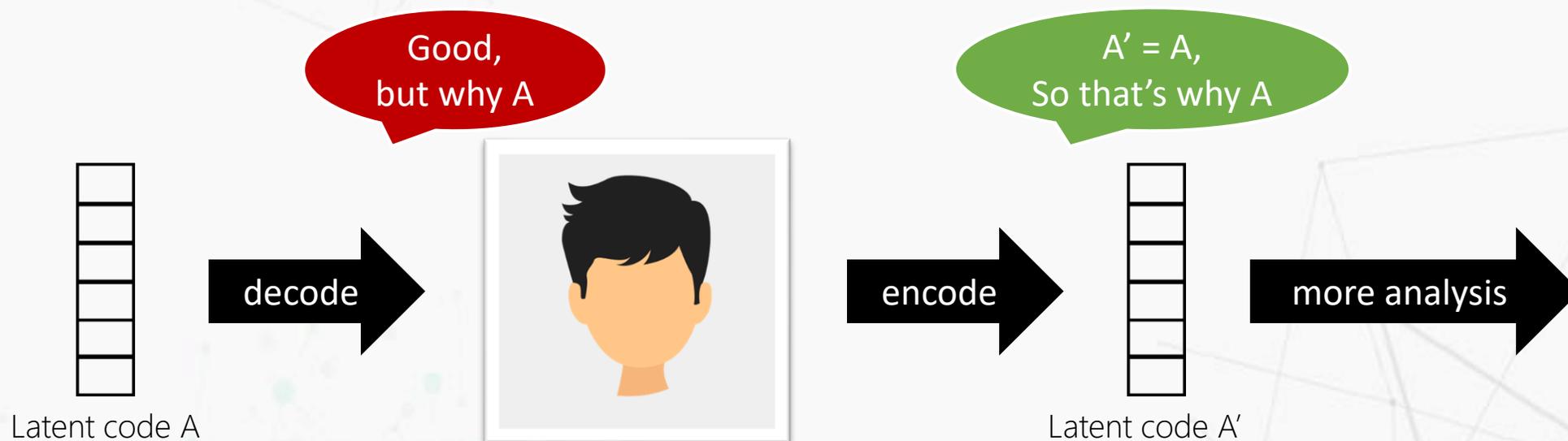
1. 專題動機與目標

Motivation & Goal

1.1 專題動機與目標

- 專題動機

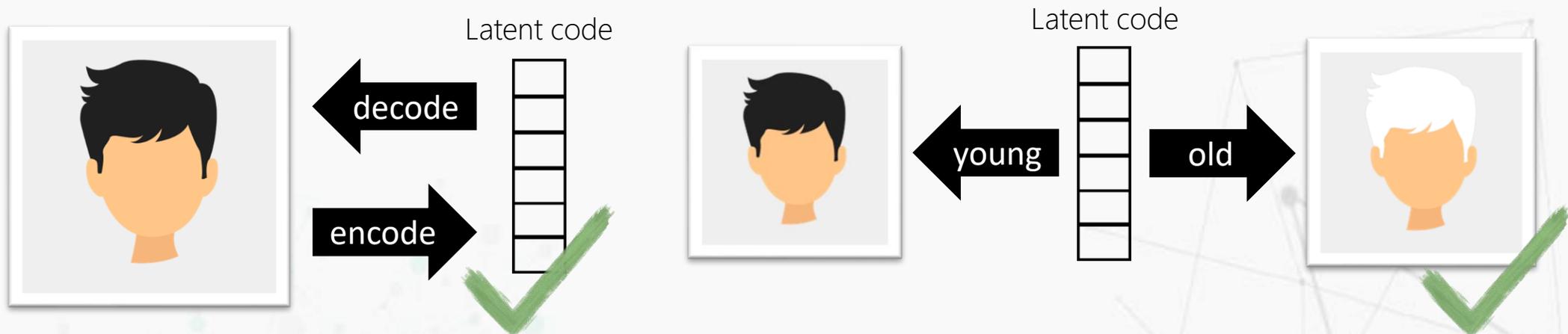
目前多數生成對抗網路 (**Generative Adversarial Networks**, 簡稱 **GAN**) 皆由隱性分類碼 (latent code) 單向映射生成擬真圖像並成效優異，然而若能再次回推編碼，則或許能藉該編碼進行數值分布、群落分析，以對圖像得到更多可用資訊。



1.1 專題動機與目標

- 專題目標

- 訓練 BCGAN 模型，以生成與重建擬真人臉及手寫數字作為驗證模型能力的指標
- 修改模型架構，允許使用者操作
 - 1) 生成人臉的年齡、性別
 - 2) 同樣筆跡下的不同數字
- 驗證 latent code 對於生成圖像的影響力，以及是否符合被賦予的實質意義



1.2 專題貢獻

- 貢獻

- 1) 開發 **Modified BCGAN** ([Bidirectional Conditional GAN](#)) 架構，使其可基於特定條件學習圖像生成
- 2) 在其他領域 (如生物領域) 中，存在將圖片轉換為編碼並衍伸其他分析研究工作的需求，本實驗應可作為其他研究的基石
- 3) 本實驗使用 MNIST 手寫數字集及 Flickr-Faces-HQ 人臉資料集進行模型效能測試，驗證可成功達成**圖像重建以及產生條件生成差異**的功能

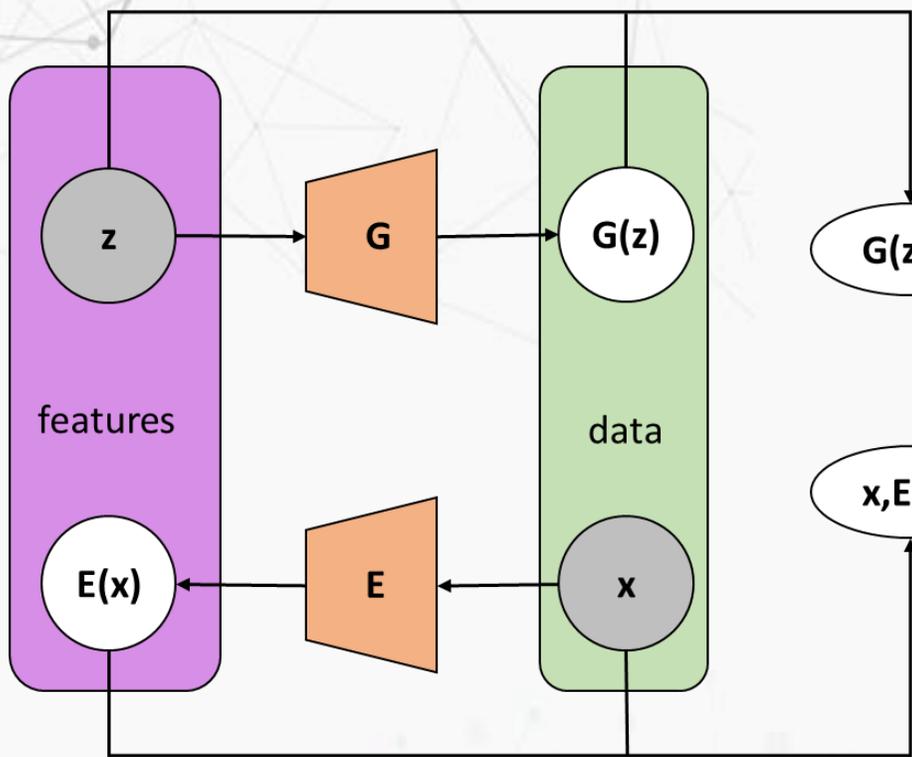


模型介紹

Introduction of BiGAN

2.1 BiGAN 架構 & 訓練流程

- 模型架構



Symbols :

$x \rightarrow$ 真實資料 ; $z \rightarrow$ random noise

$G \rightarrow$ generator ; $E \rightarrow$ encoder ; $D \rightarrow$ discriminator

$E(x) \rightarrow$ 經過 Encoder 編碼的 latent code

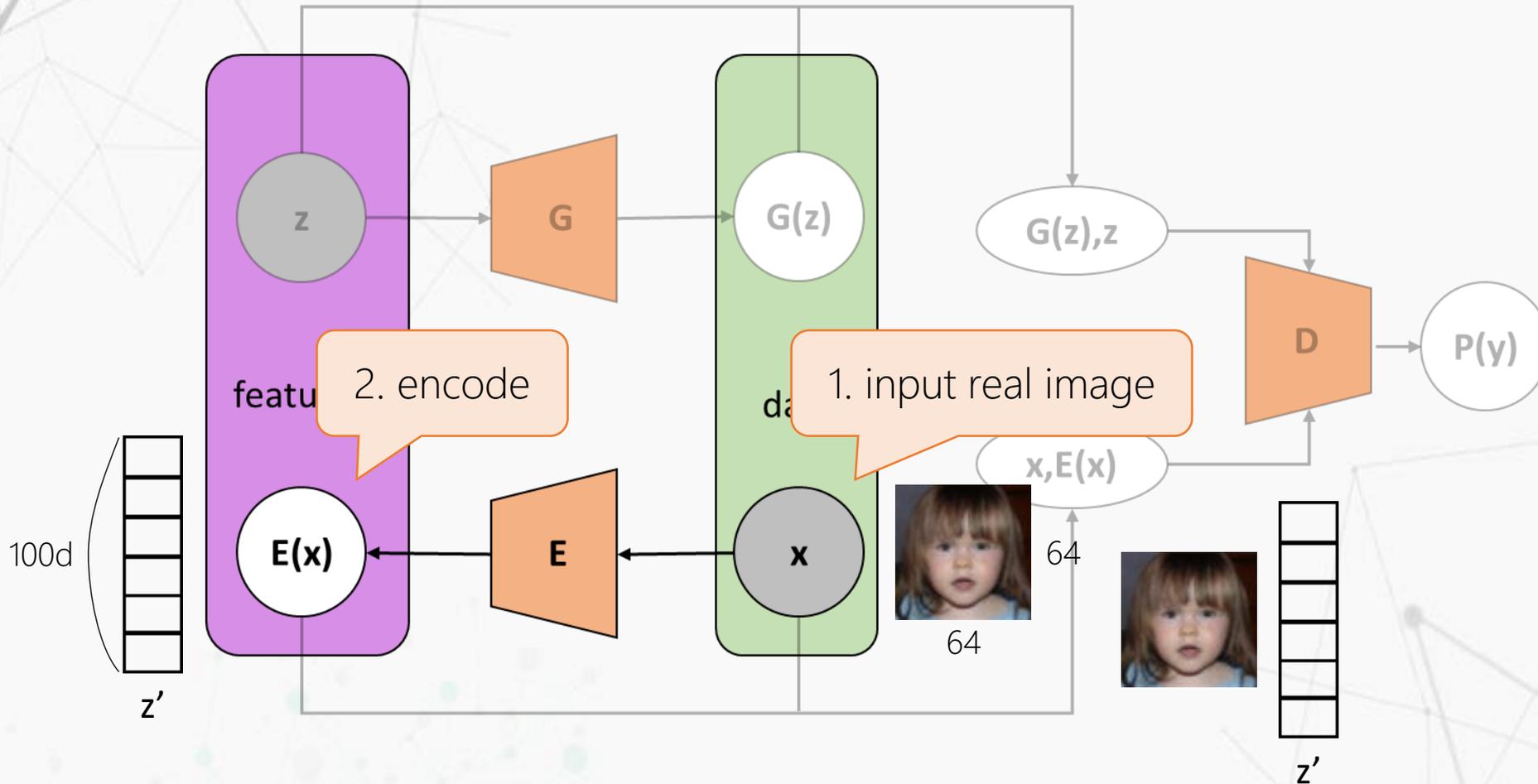
$G(x) \rightarrow$ 經過 Generator 產生的圖像

Objective :

$$\begin{aligned} & \min_{G,E} \max_D V(D, E, G) \\ & = E_{x \sim P_x} [E_{z \sim P_E(\cdot|x)} \text{Maximize } \log D(x, z)] + E_{z \sim P_z} [E_{x \sim P_G(\cdot|z)} \log(1 - \text{Minimize } D(x, z))] \\ & = E_{x \sim P_x} [\text{Minimize } \log D(x, E(x))] + E_{z \sim P_z} [\log(1 - \text{Maximize } D(G(z), z))] \end{aligned}$$

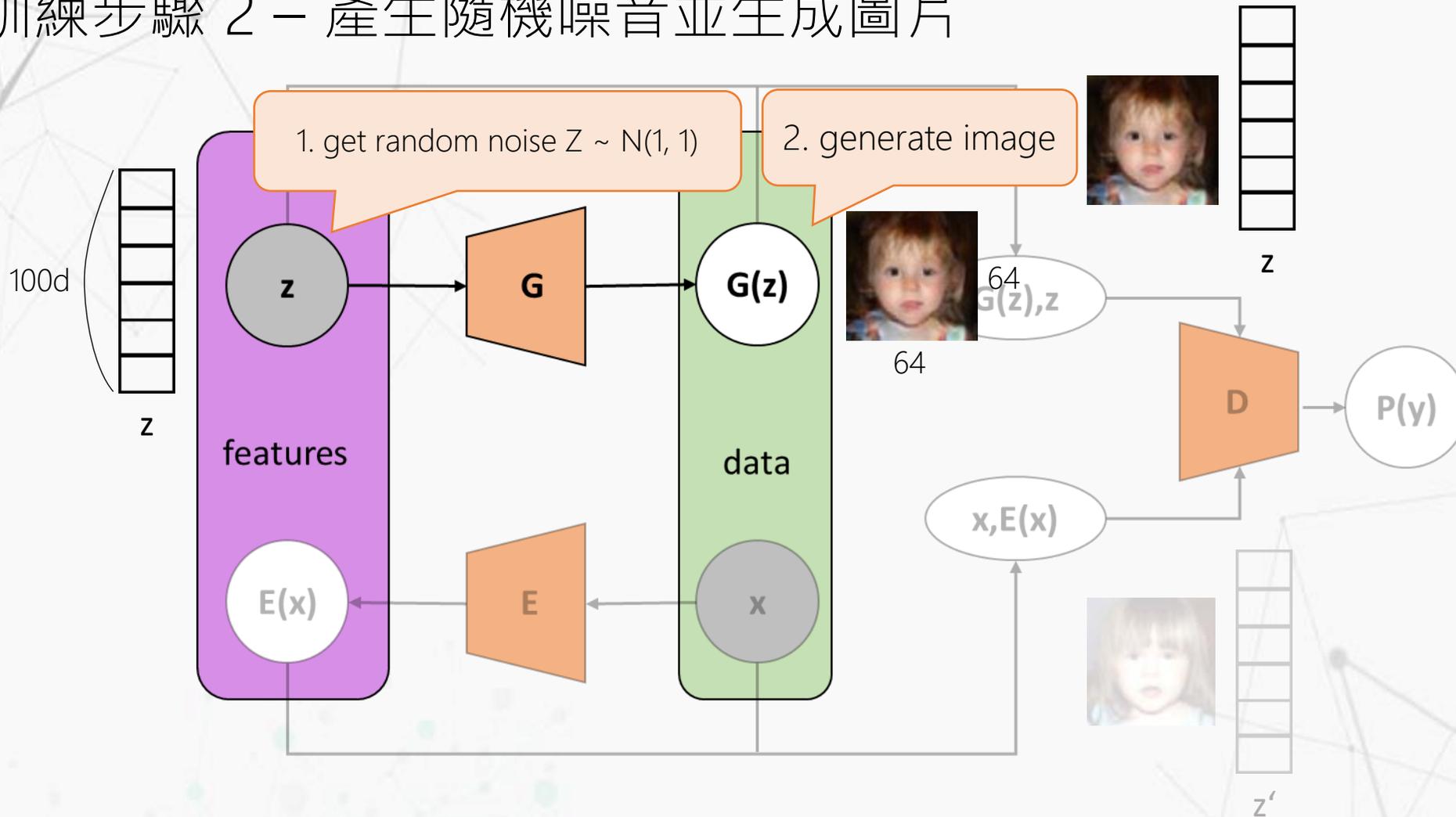
2.1 BiGAN 架構 & 訓練流程

- 訓練步驟 1 – 輸入真實圖片並將其編碼



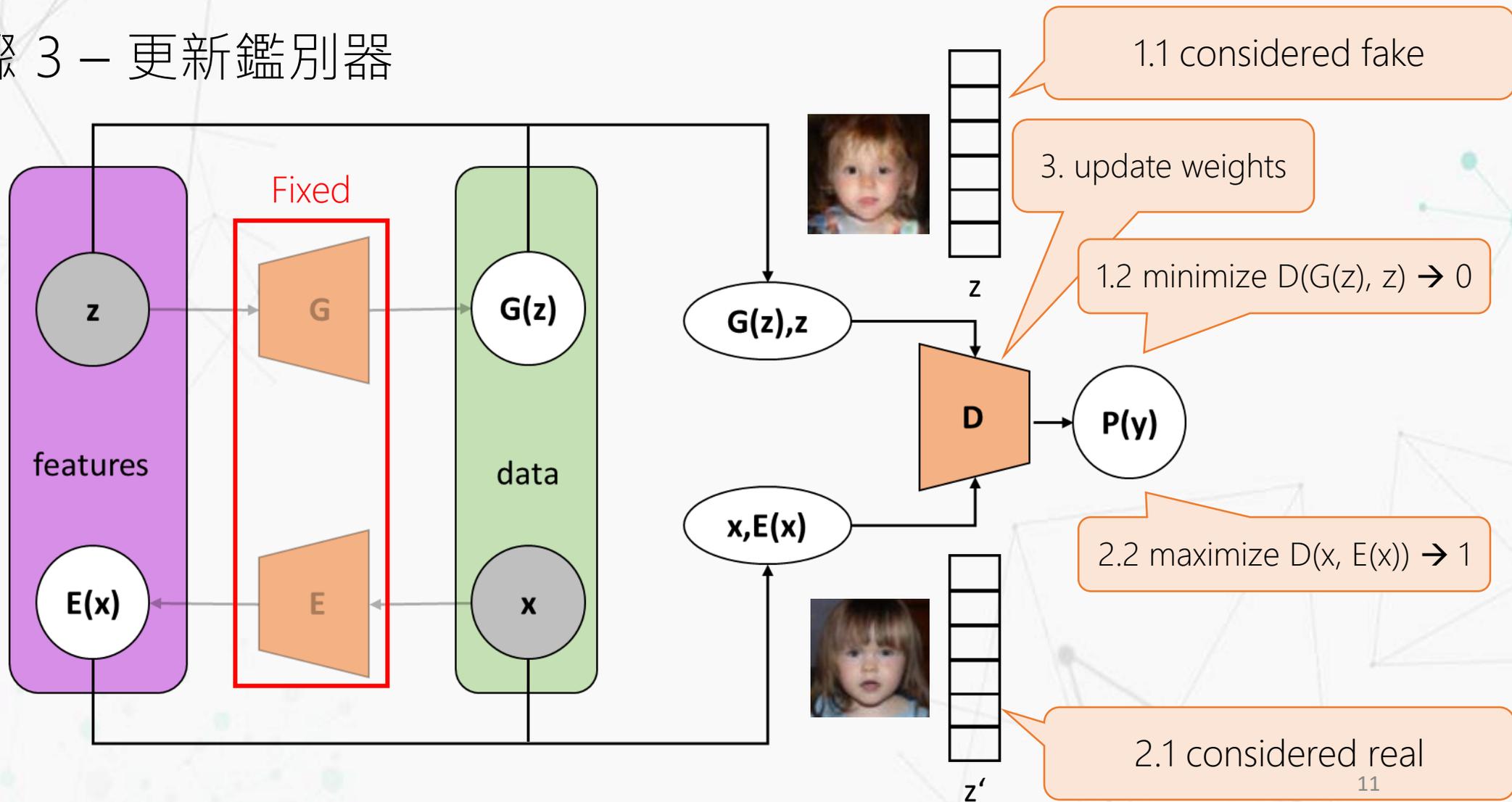
2.1 BiGAN 架構 & 訓練流程

- 訓練步驟 2 – 產生隨機噪音並生成圖片



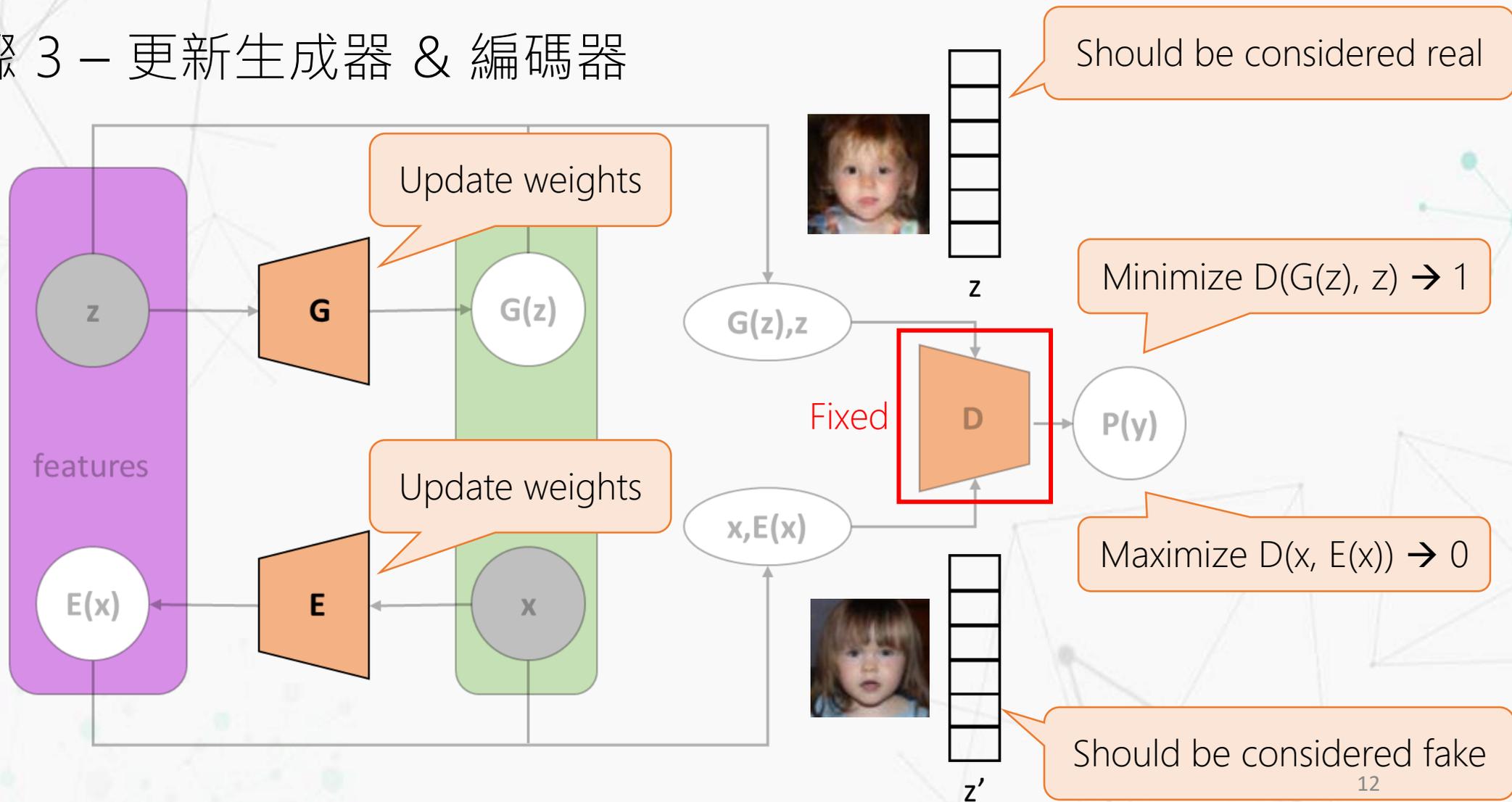
2.1 BiGAN 架構 & 訓練流程

- 訓練步驟 3 – 更新鑑別器

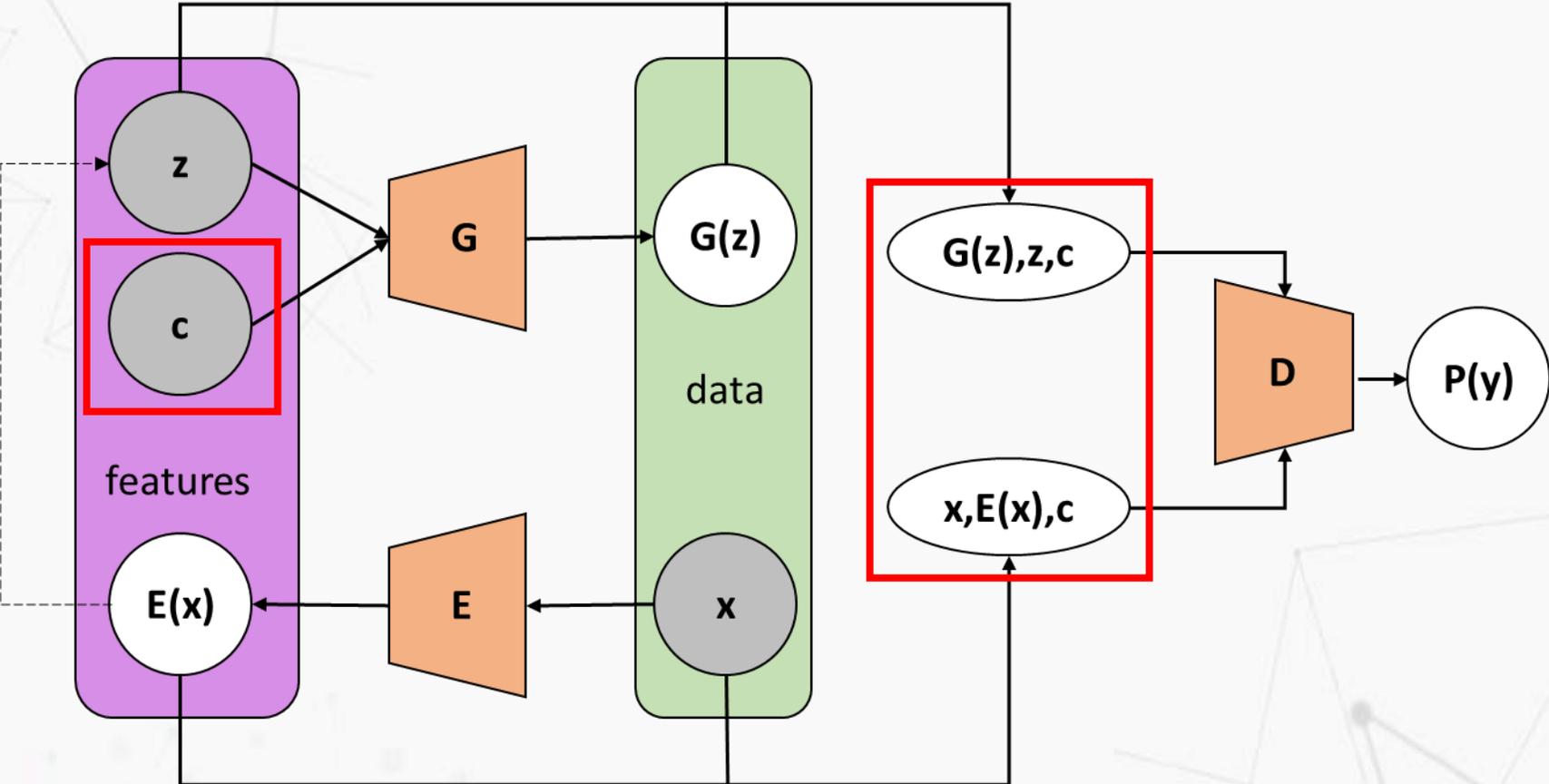


2.1 BiGAN 架構 & 訓練流程

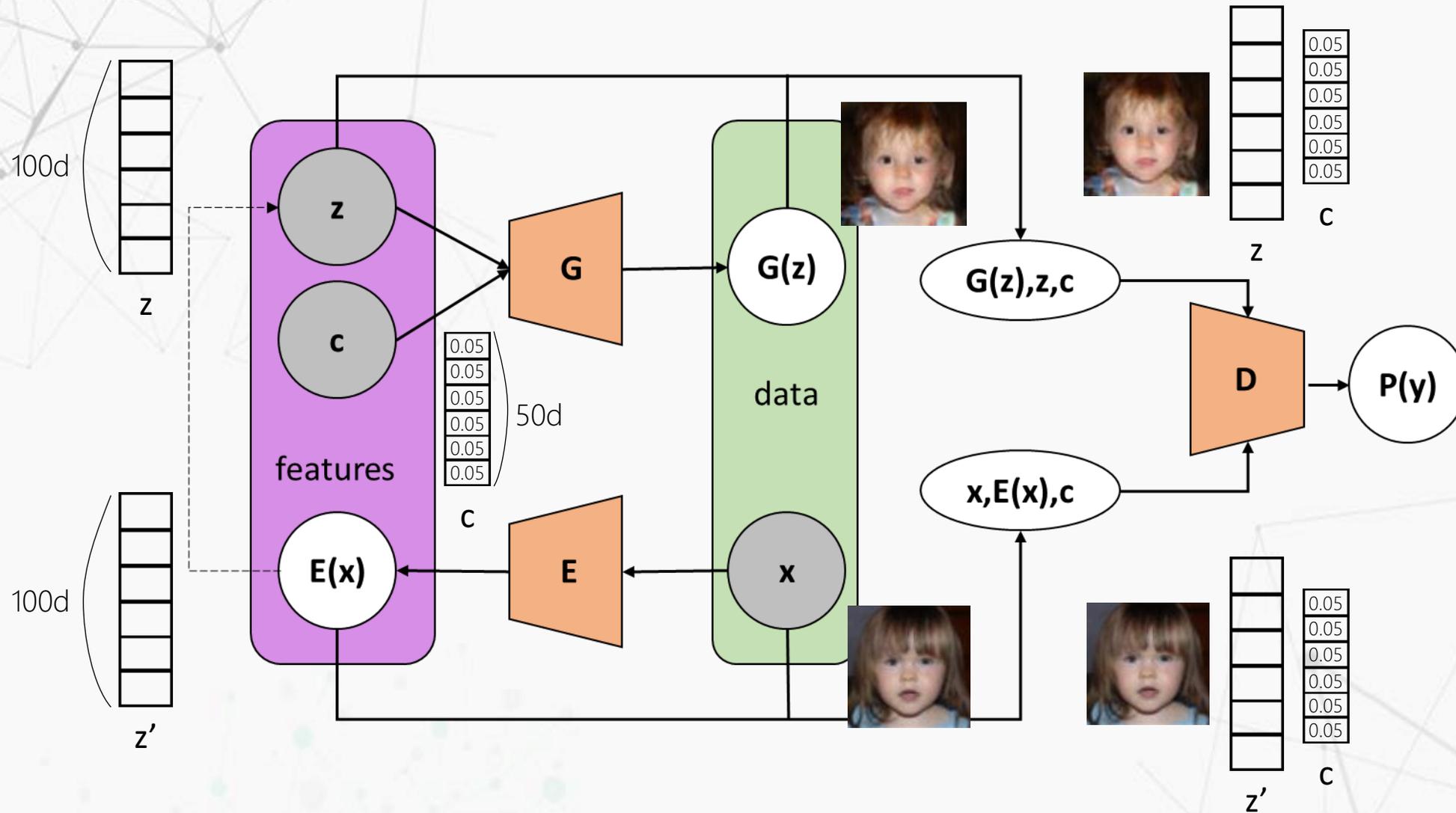
- 訓練步驟 3 – 更新生成器 & 編碼器



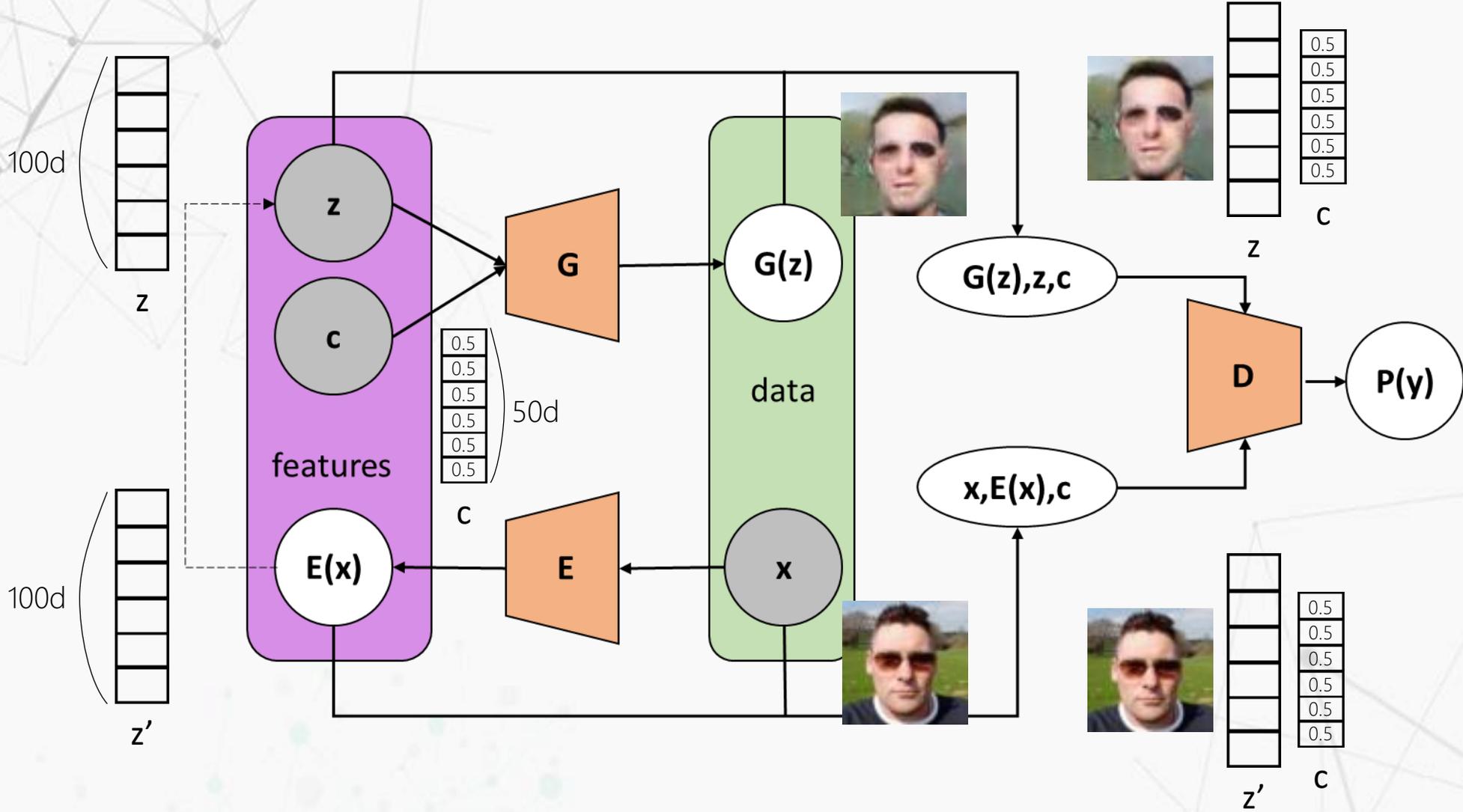
2.2 Modified BiGAN



2.2 Modified BiGAN



2.2 Modified BiGAN





實驗結果

Experiments

3.1 訓練資料集

- MNIST 手寫數字資料集

- 70000 張 28 x 28 手寫數字圖片
- 0 ~ 9 共十種數字類別



- Flickr-Faces-HQ (FFHQ) 人臉資料集

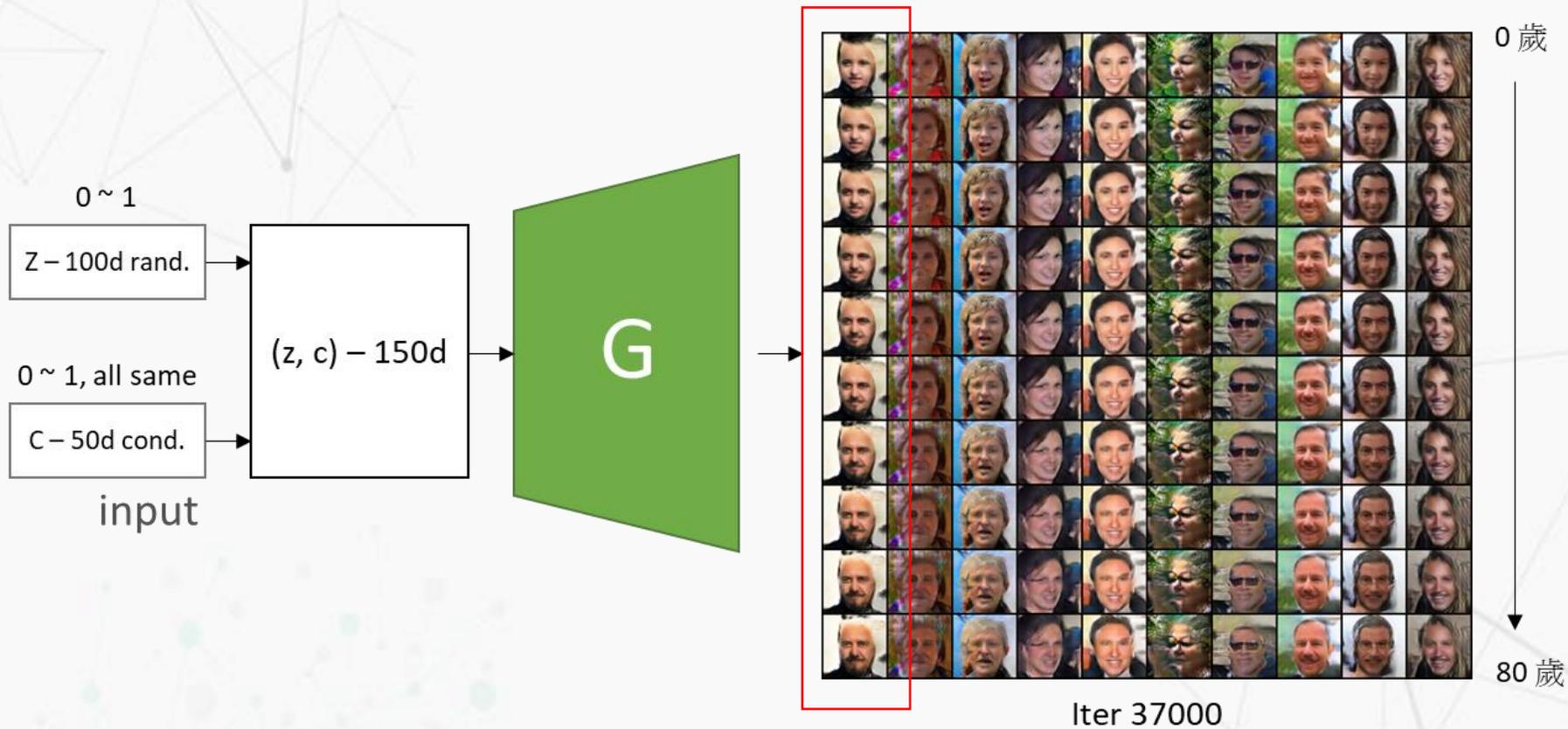
- 70000 張 1024 x 1024 高品質人臉圖片
- 每張照片均有年齡、性別、穿戴衣飾等詳細資訊，本實驗中僅使用年齡與性別作為訓練用標註資料
- 本實驗中均將圖像 resize 為 64 x 64 大小



3.2 實驗結果 - FFHQ

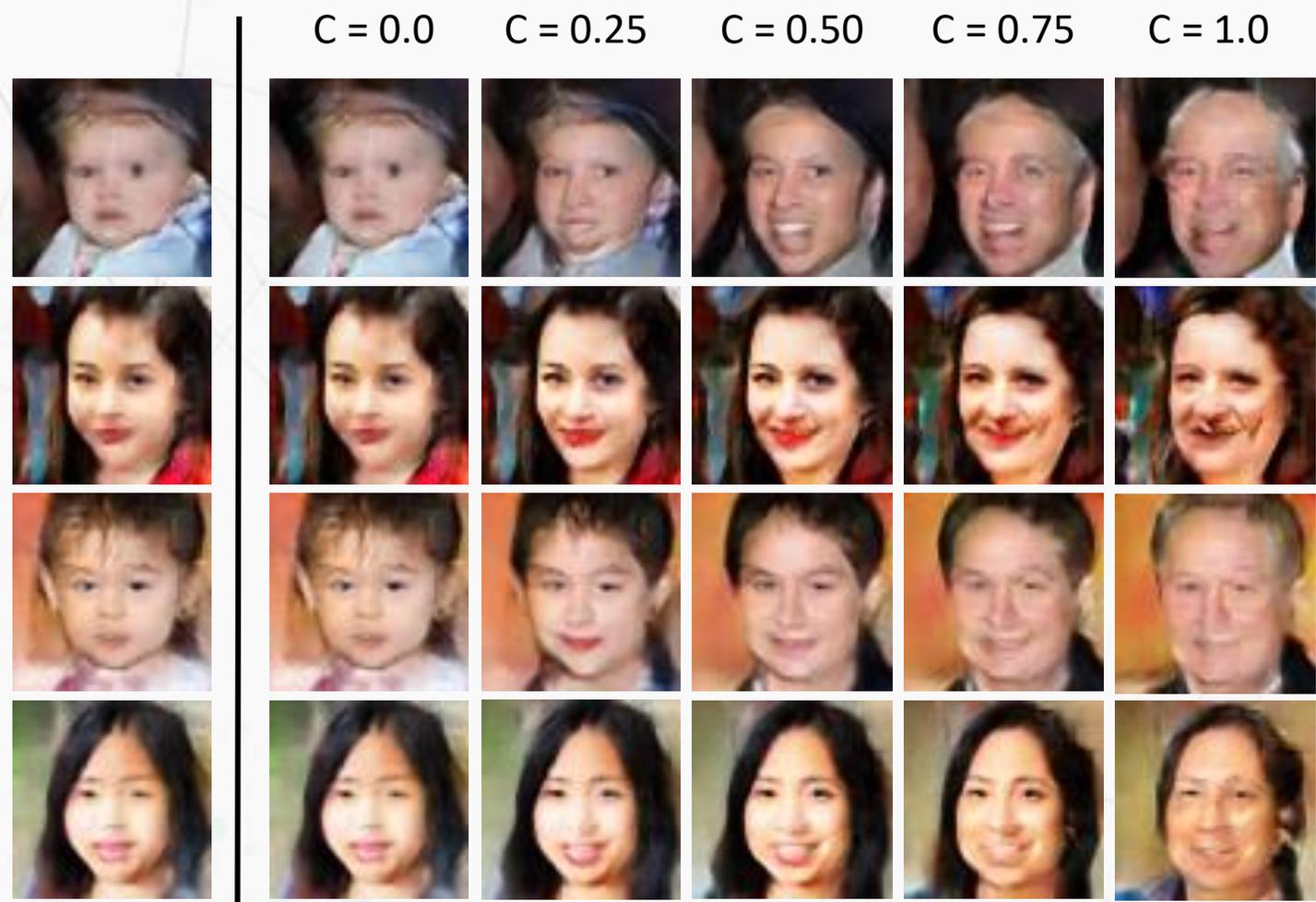
Generation

每個 column 中上至下十個圖片共用同一組隨機噪音 Z ，並逐個 row 給定值為 $0 \sim 1$ 的數字 C 代表 $0 \sim 80$ 歲，另外為了加強該數字對整體 latent code 的影響力而擴展為皆含相同數字的 50 維向量。



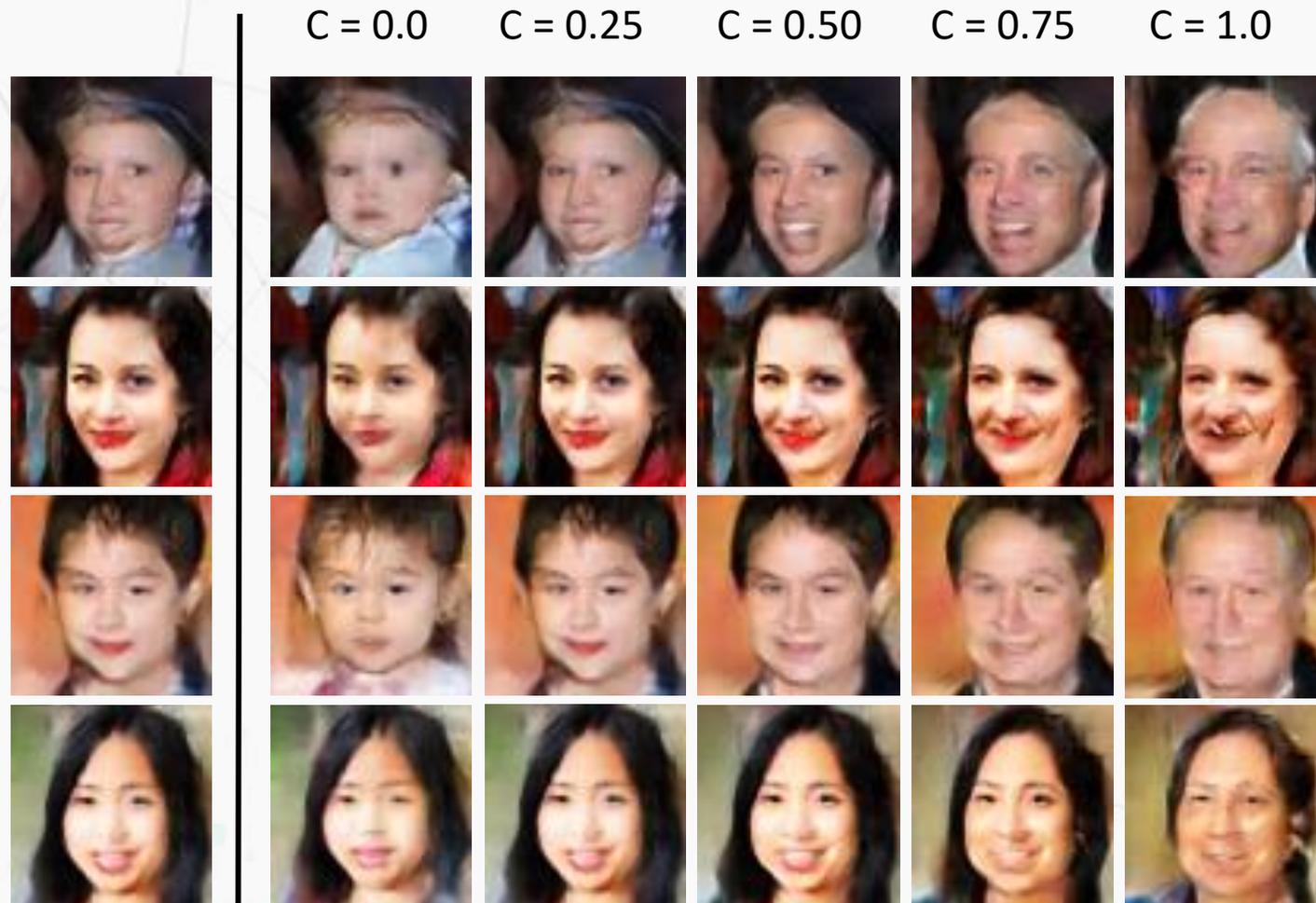
3.2 實驗結果 - FFHQ

Generation



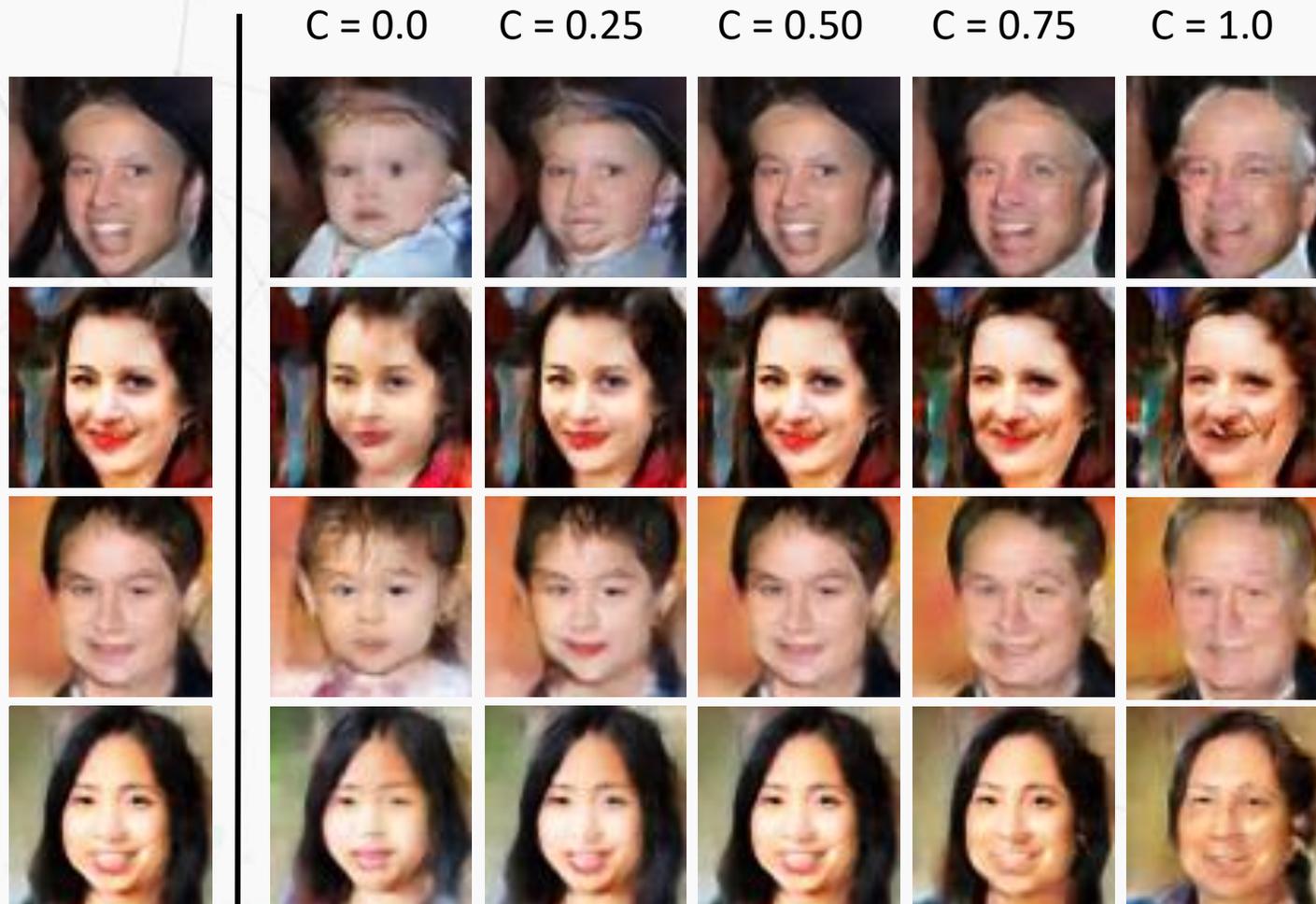
3.2 實驗結果 - FFHQ

Generation



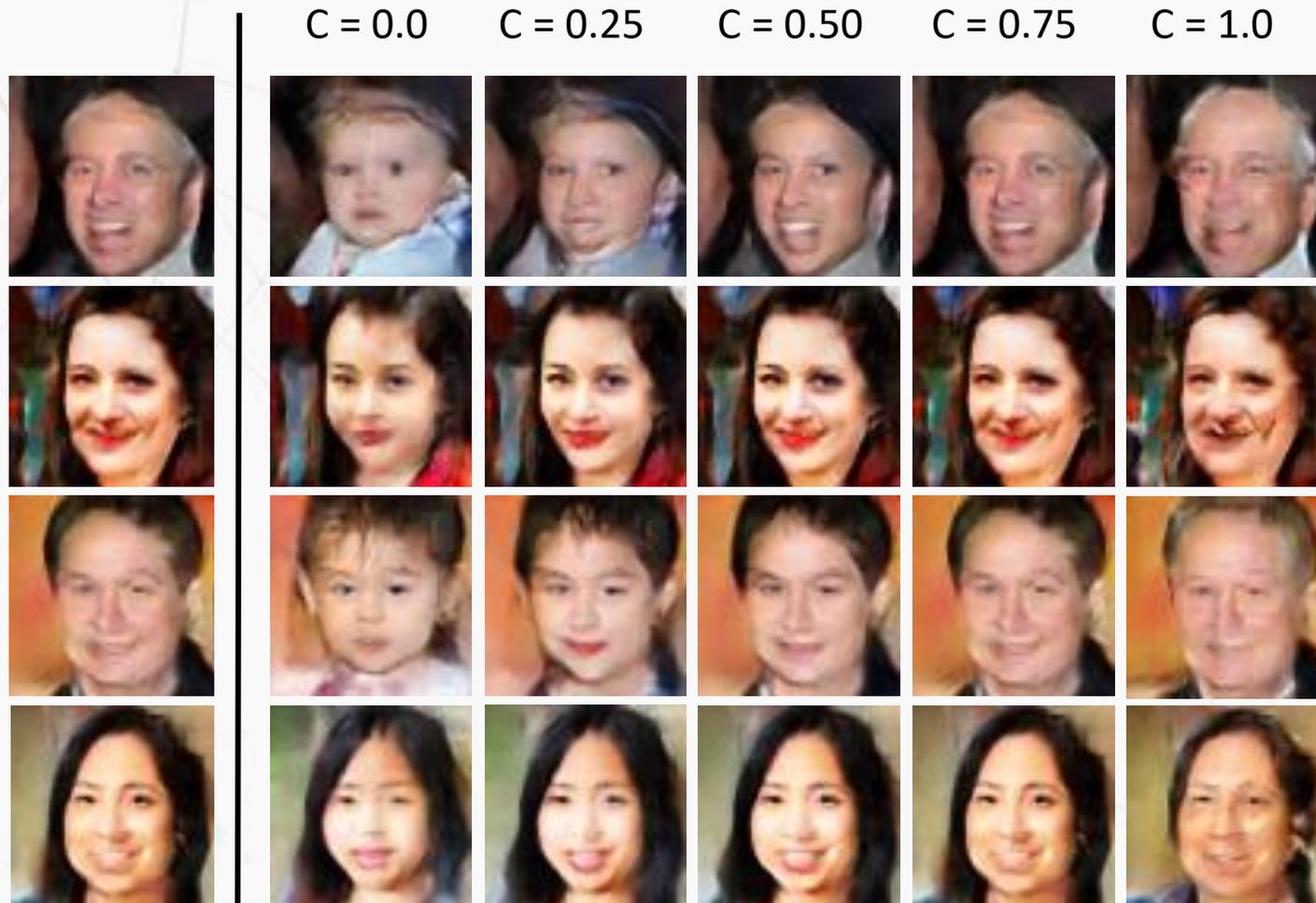
3.2 實驗結果 - FFHQ

Generation



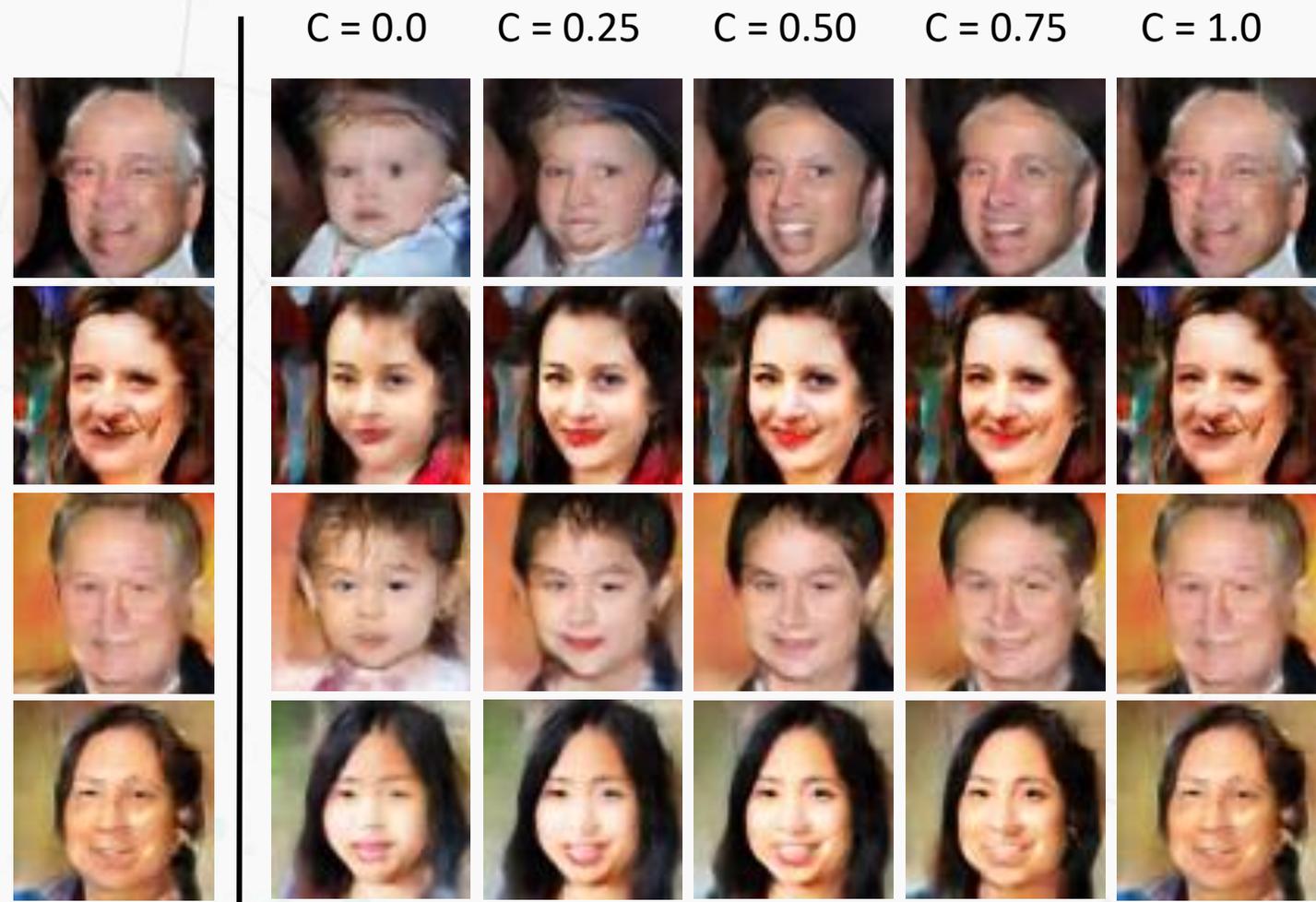
3.2 實驗結果 - FFHQ

Generation



3.2 實驗結果 - FFHQ

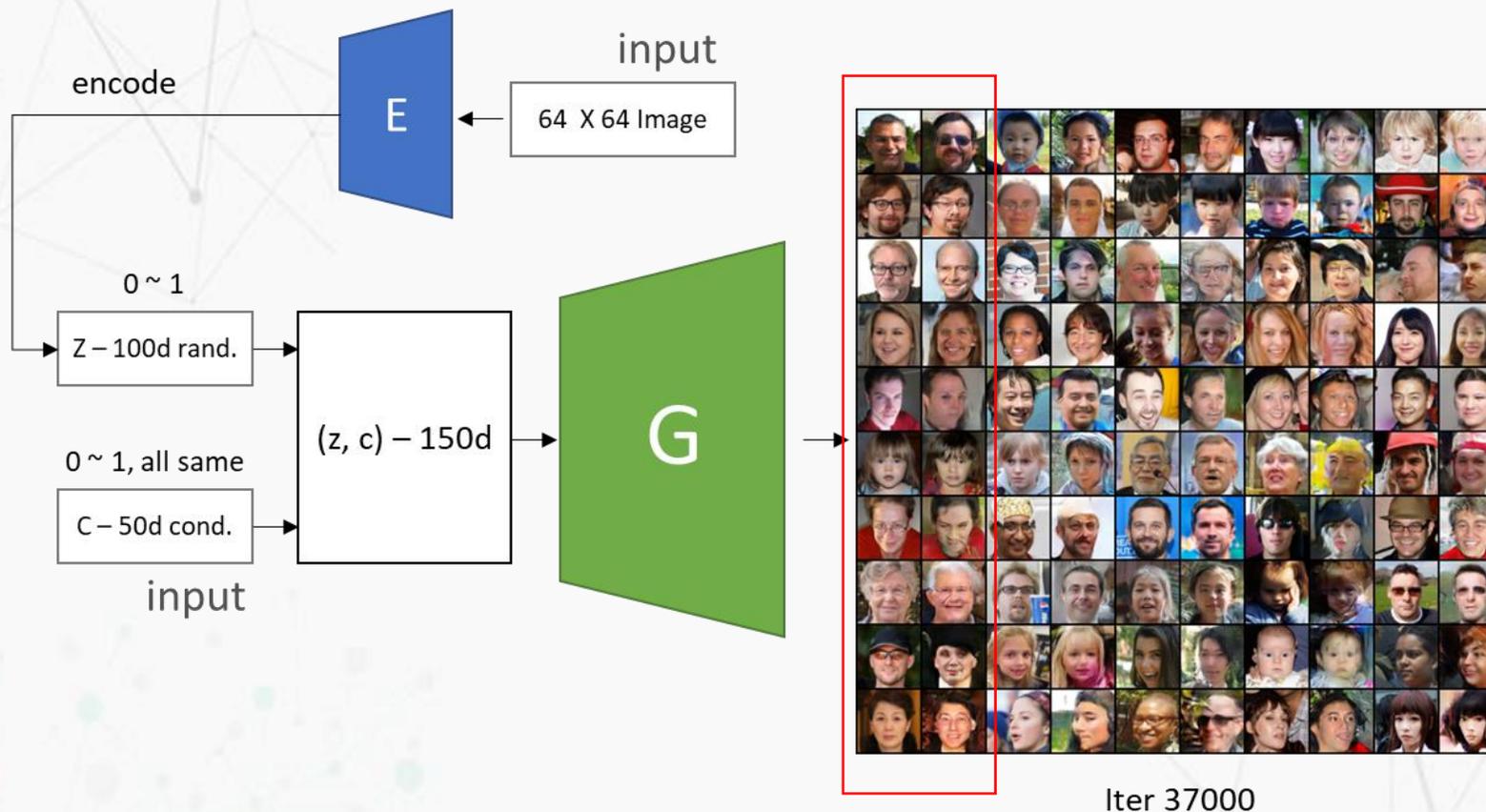
Generation



3.2 實驗結果 - FFHQ

Reconstruction

兩個 column 為一組，左為原圖，右為經編碼後生成向量 Z ，並與原圖對應年齡擴展向量 C ，經生成器重新生成的圖片。



3.2 實驗結果 - FFHQ

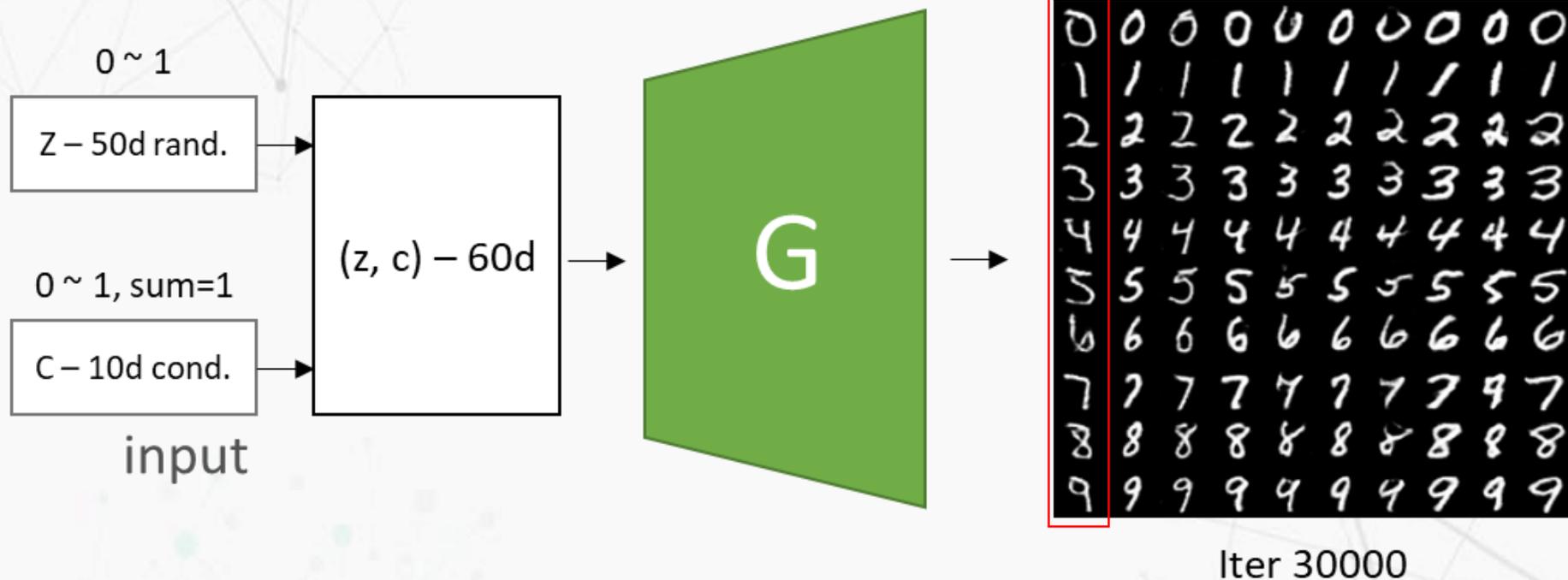
Reconstruction



3.2 實驗結果 - MNIST

Generation

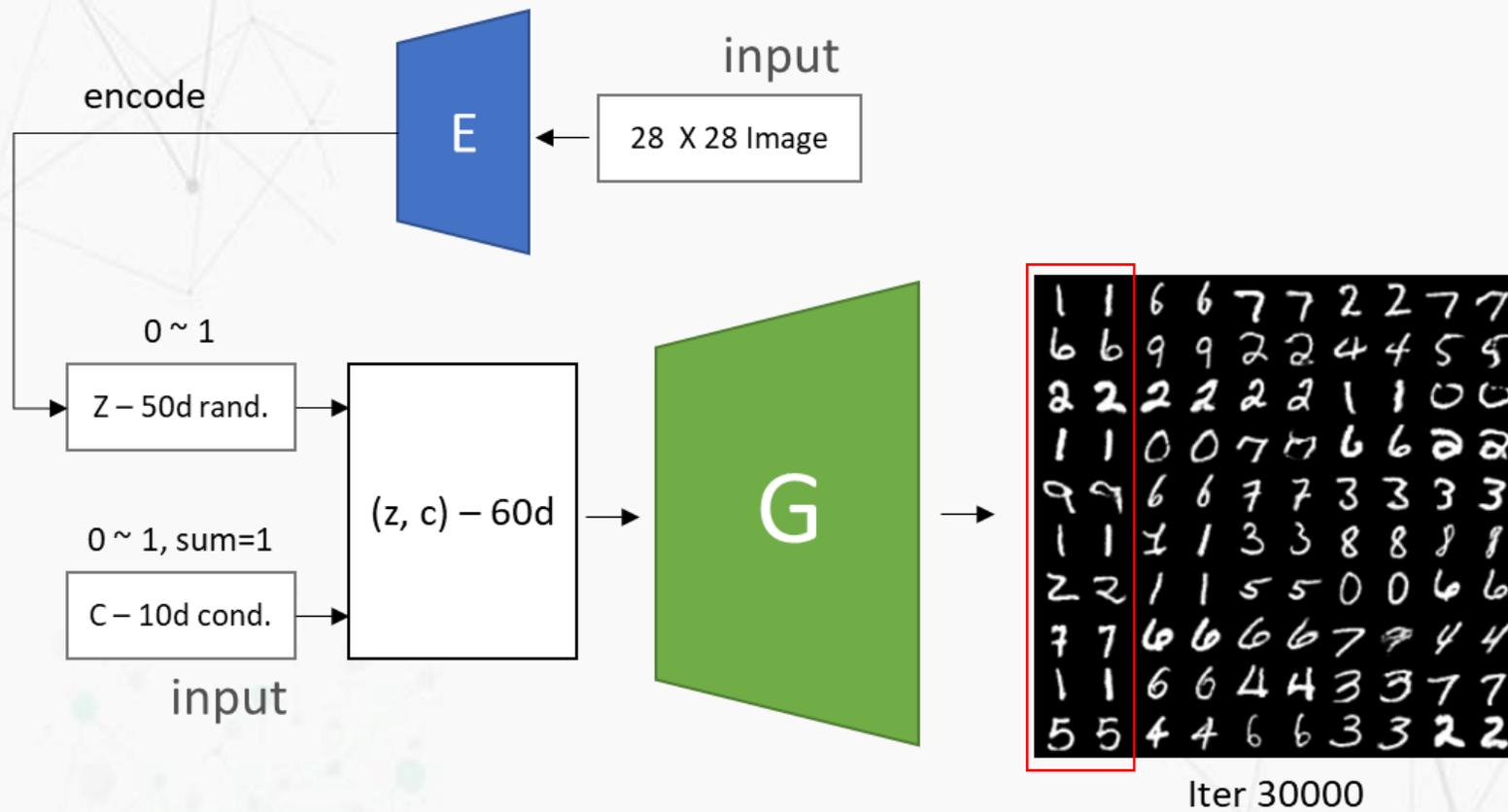
每個 column 上至下十個圖片共用同一組隨機噪音 Z，並逐個 row 給定 0 ~ 9 的限制生成條件離散向量 C，生成 28 x 28 的手寫數字圖片。



3.2 實驗結果 - MNIST

Reconstruction

兩個 column 為一組，左為原圖，右為原圖經編碼後生成向量 Z，並與原圖對應數字離散向量 C，經生成器重新生成的圖片。

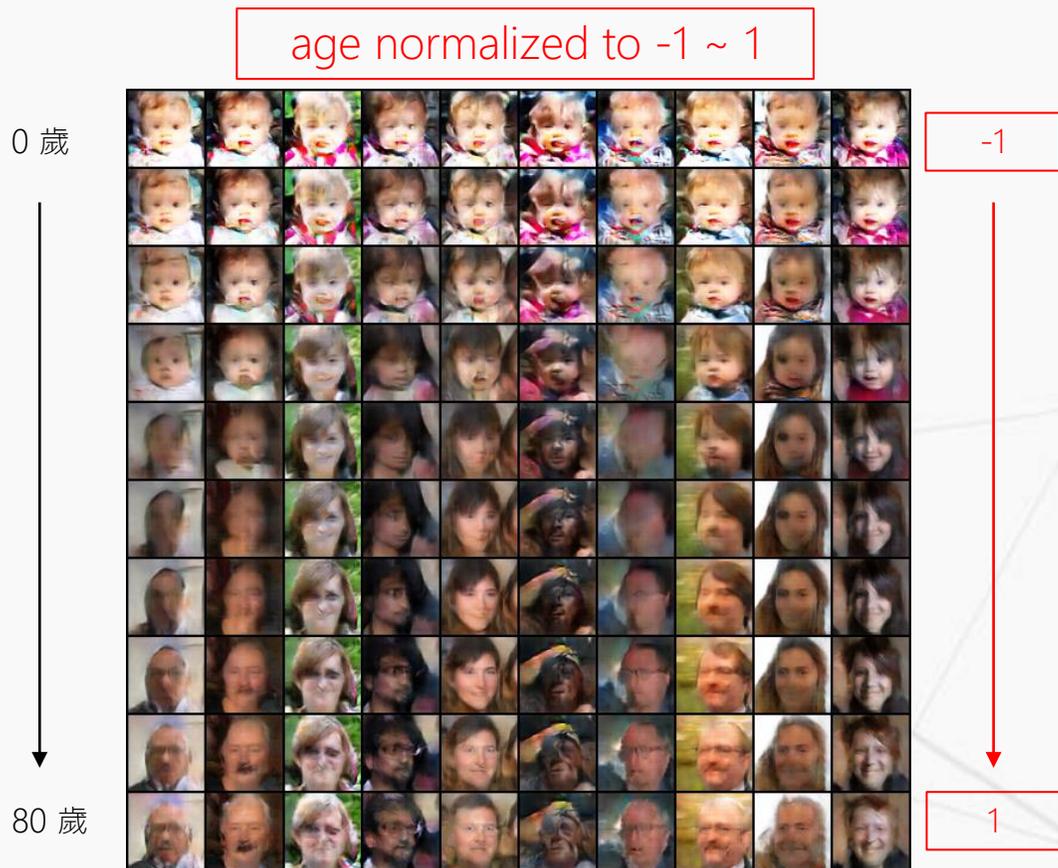
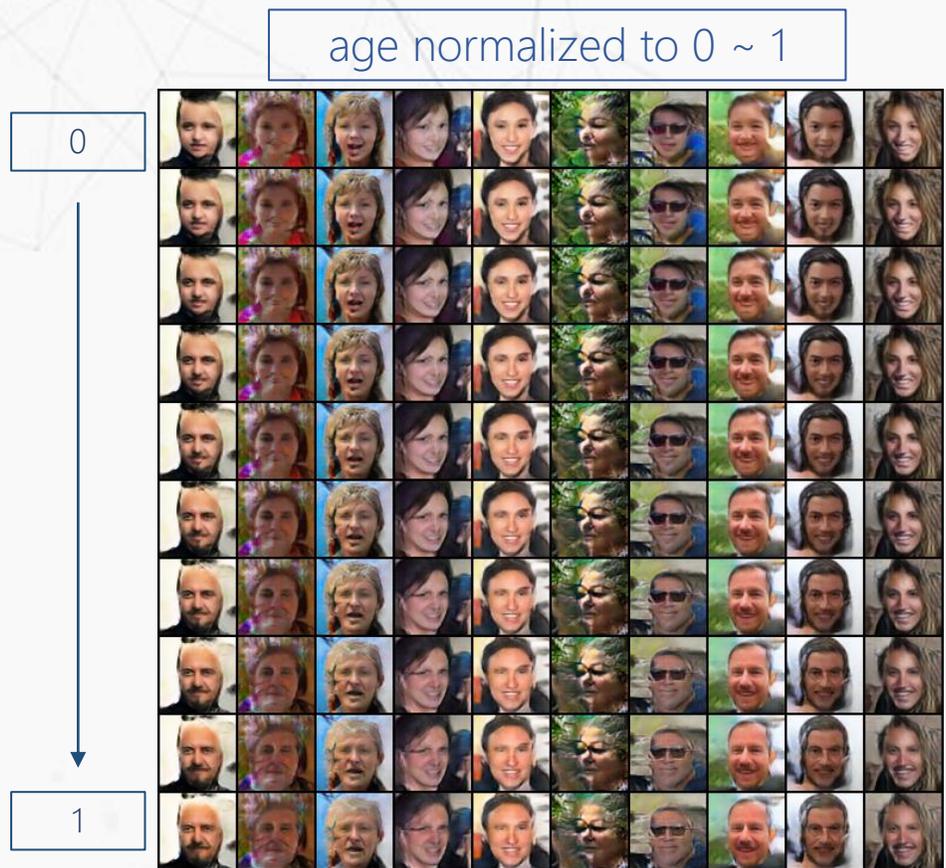


3.3 實驗優化

- 優化 1 – 正規化域值調整

本實驗中原先將 0~80 歲正規化至 0~1 的範圍內，結果上雖然可實現臉部細節隨年齡增長的變化，但準確對應年齡的能力有限。

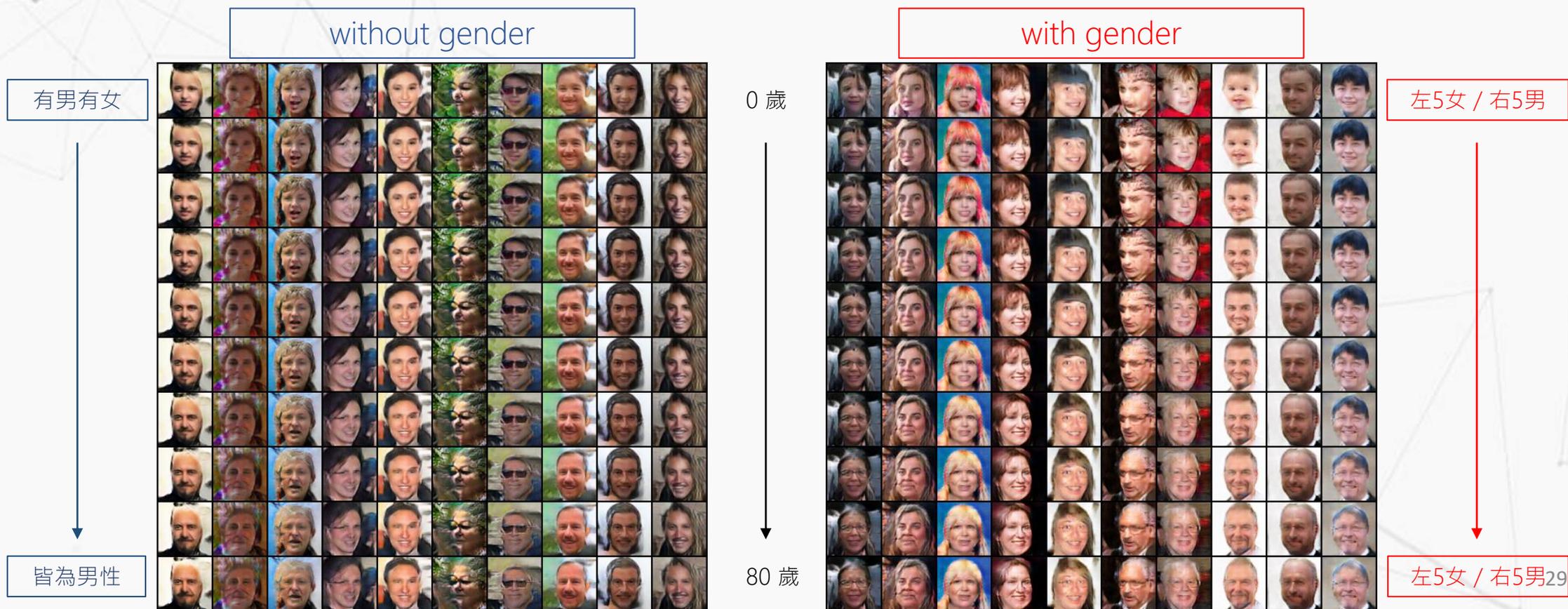
因此我們延展正規化目標範圍為 -1 ~ 1，經測試後發現可有效增強年齡匹配的效果，但細部輪廓仍需再經模型調整優化。

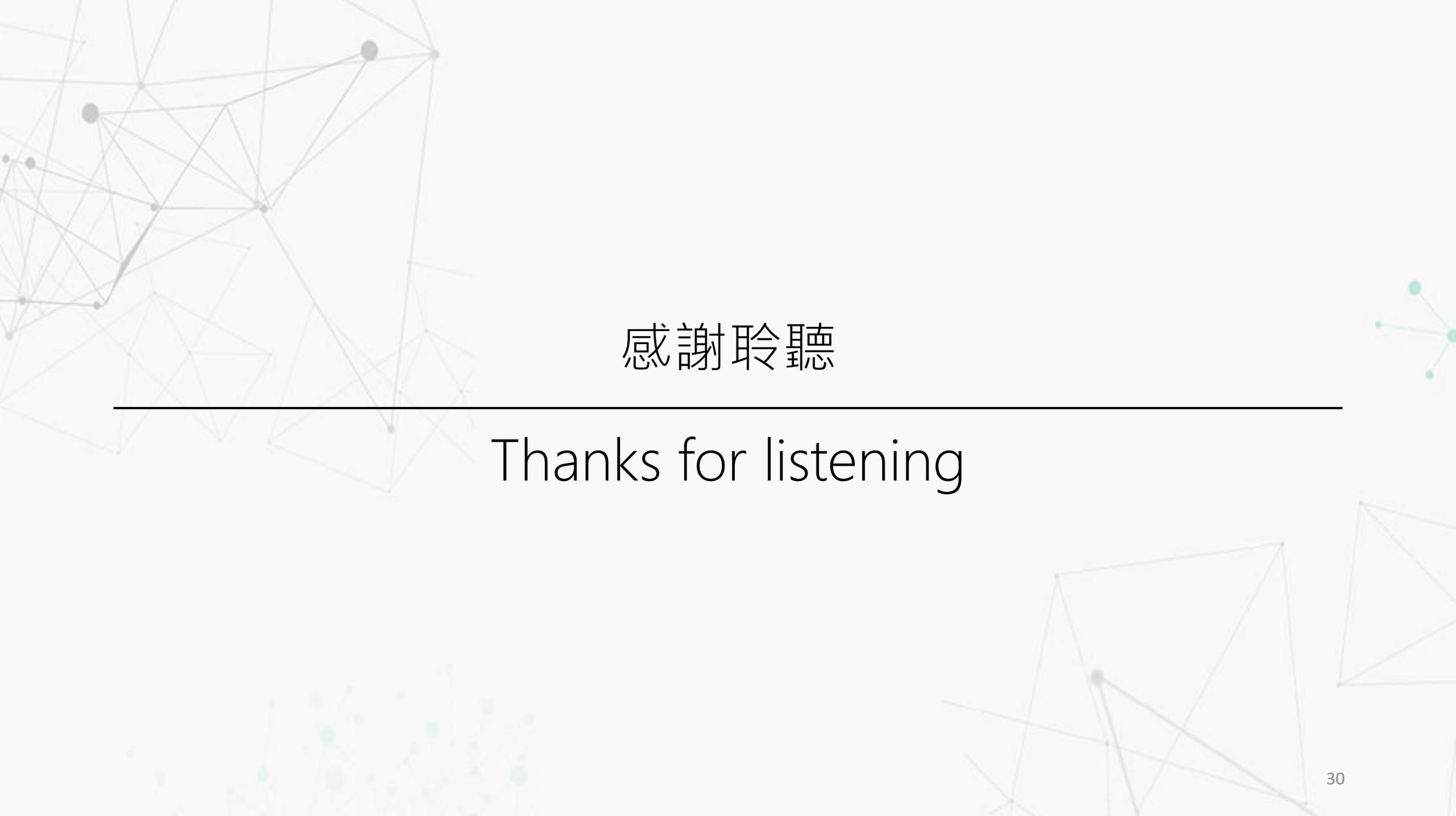


3.3 實驗優化

- 優化 2 – 增加性別控制項

在原先實驗結果中，亦可發現無論一人臉在年齡控制值較小時性別為何，一旦加大該控制項至接近 80 歲，皆會轉變為男性臉孔，推測是因為資料集中高歲數照片以男性佔比較高導致。因此我們增加了性別控制項進行訓練，結果顯示此法可成功克服該問題。





感謝聆聽

Thanks for listening