COCO-GAN: Generation by Parts via Conditional Coordinating

Chieh Hubert Lin, Chia-Che Chang, Yu-Sheng Chen, Da-Cheng Juan, Wei Wei, Hwann-Tzong Chen

ICCV 2019

Sanghyeon Lee

21 Mar. 2020

Example

Contribution

1. Achieve *state-of-the art* generation quality in FID score.

2. Generation image in the extrapolation coordinate patch

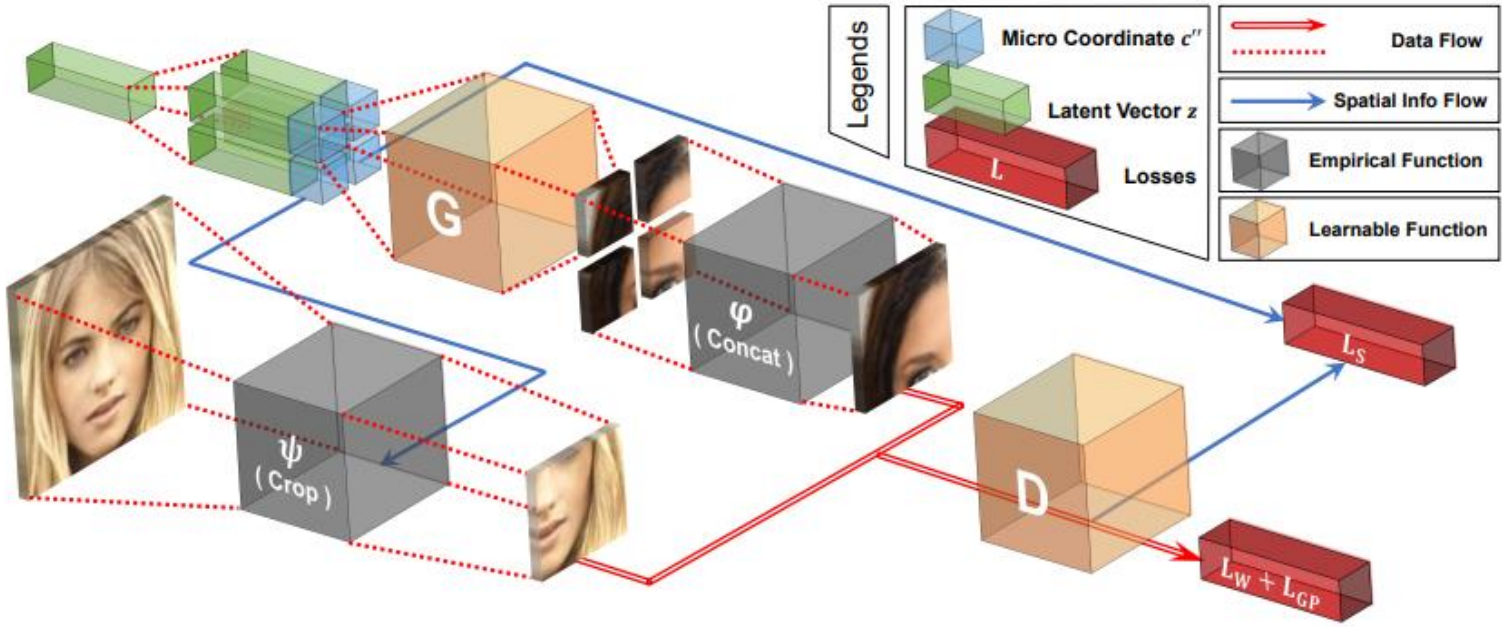
3. Saves a significant amount of memory for both training and inference in the generating image by part

4. Generate panorama image without any special treatment

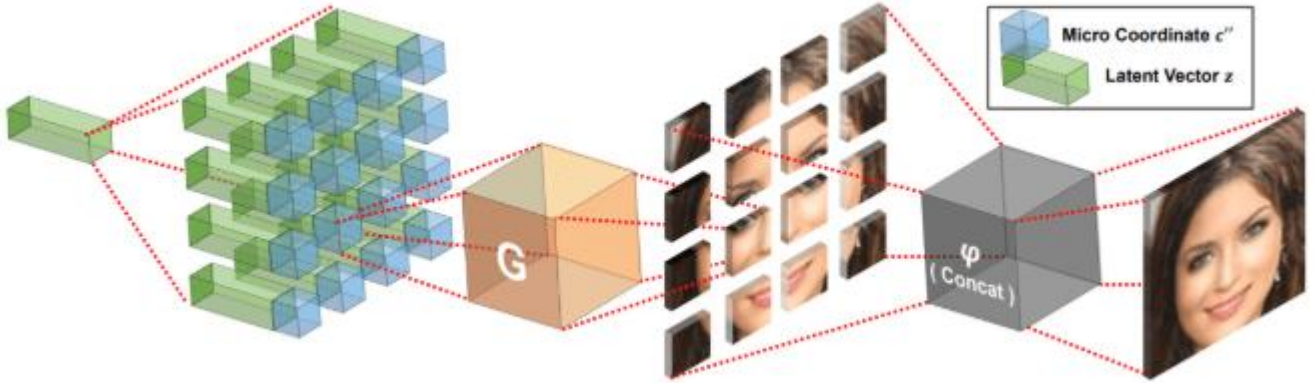
Method

Overview of COCO-GAN training and inference

Training

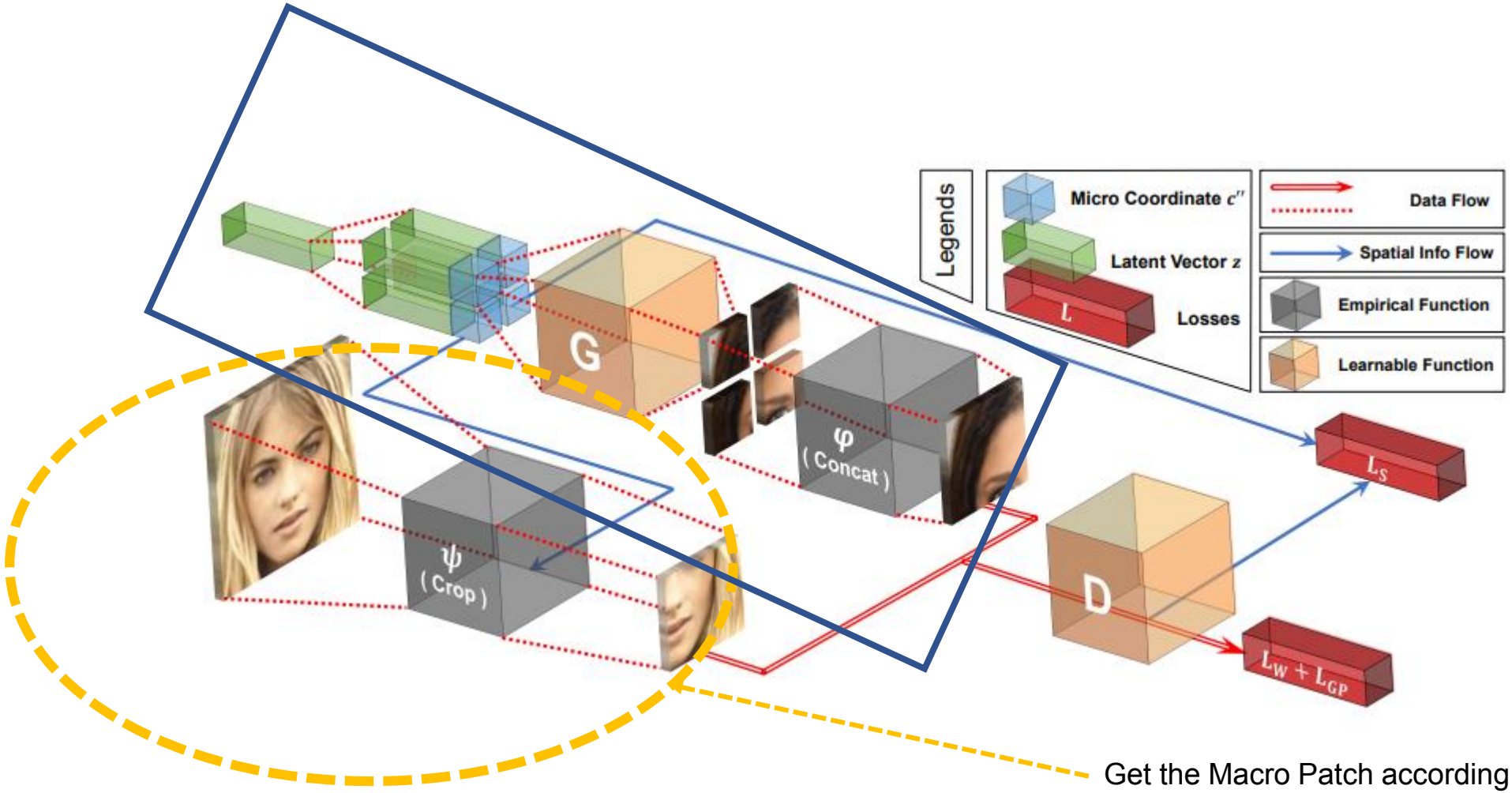


Inference



Method

Training



Get the Macro Patch according to coordinate

Method

Loss

WGAN loss

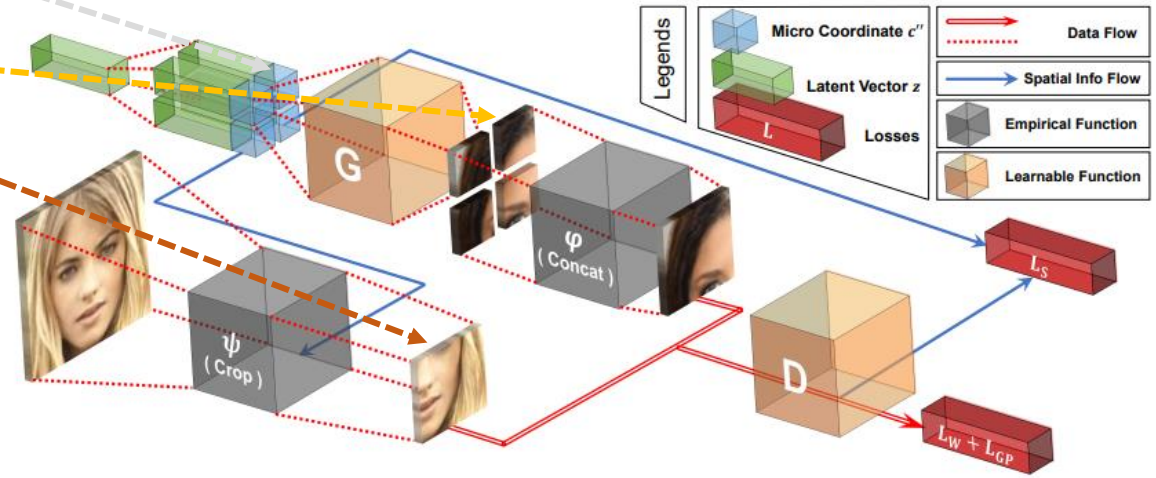
$$L_W = \mathop{\mathbb{E}}_{x,c'}[D(\psi(x,c'))] - \mathop{\mathbb{E}}_{z,C''}[D(\varphi(G(z,C'')))] .$$

Gradient penalty loss

$$L_{GP} = \mathop{\mathbb{E}}_{\hat{s}'}\left[(\|\nabla_{\hat{s}'} D(\hat{s}')\|_2 - 1)^2\right] , \quad \hat{s}' = \epsilon s' + (1-\epsilon)\, x'$$

Spatial consistency loss

$$L_S = \mathop{\mathbb{E}}_{c'}[\|c' - A(x')\|_2] .$$

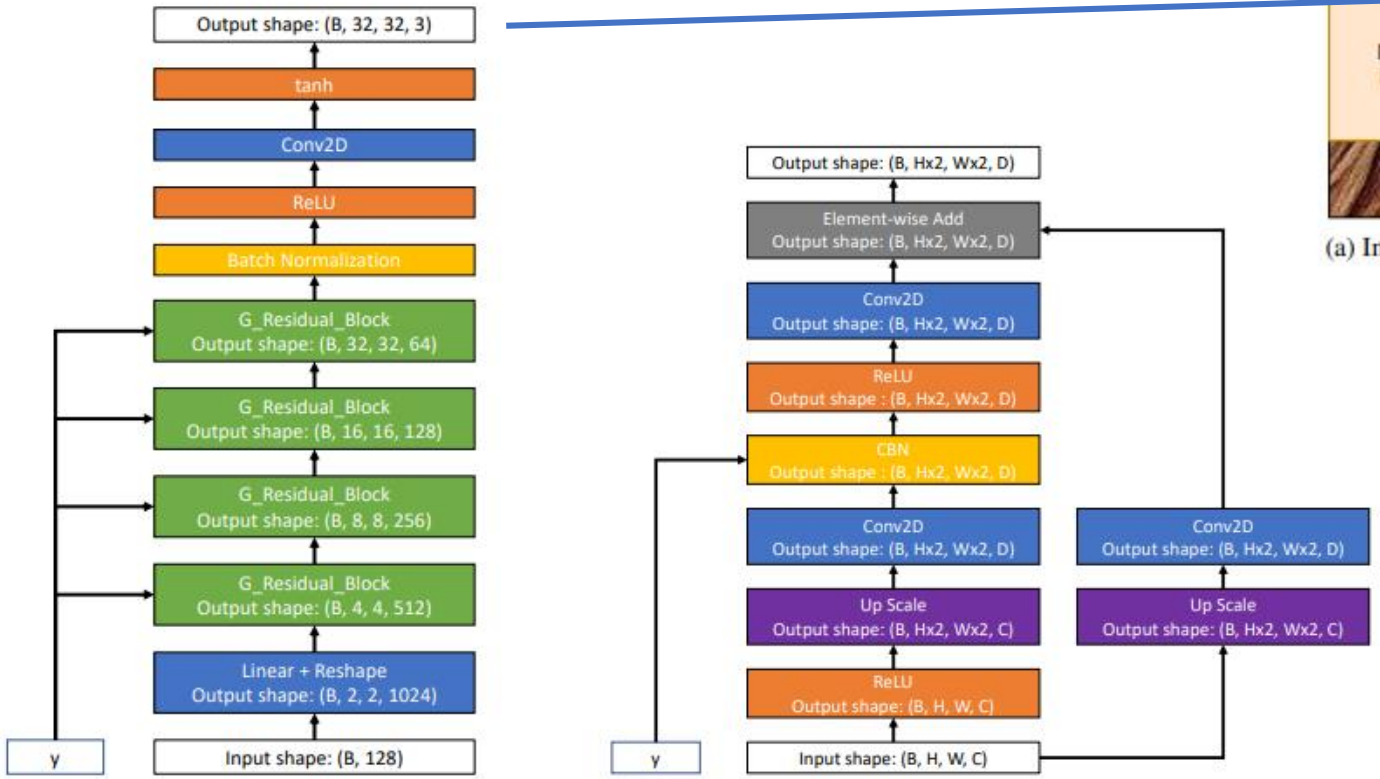


s': Generated macro patch
x': Real macro patch
c': Macro coordinate (from macro patch, (x,y)
C'': Micro coordinate (Matrix)

$$\begin{cases} L_W + \lambda L_{GP} + \alpha L_S, & \text{for } D, \\ -L_W + \alpha L_S, & \text{for } G. \end{cases}$$

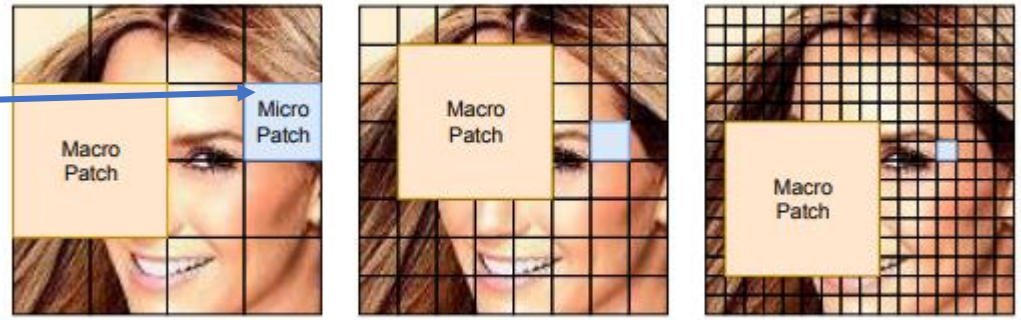
Method

Generator



(a) Generator Overall Architecture

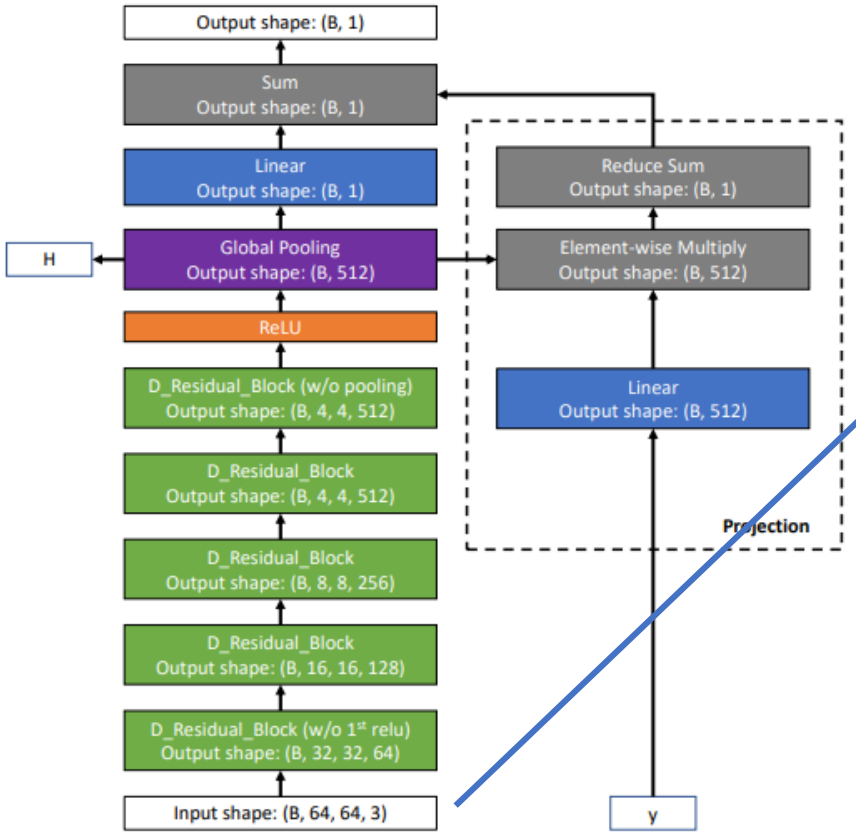
(b) Generator Residual Block



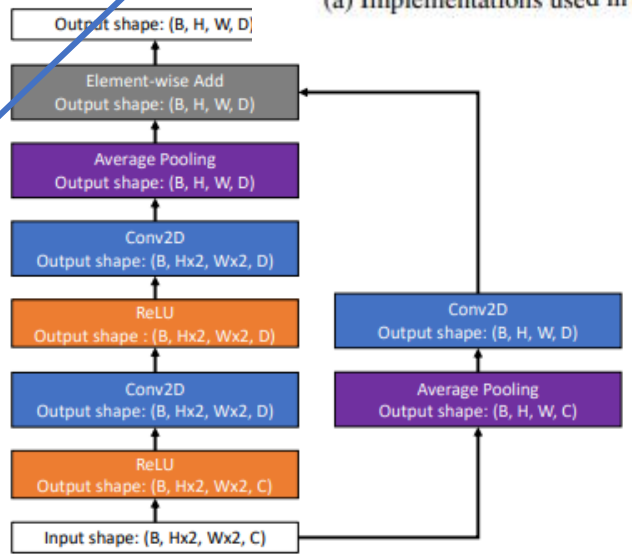
(a) Implementations used in this paper with (Left) P4x4, (Middle) P8x8 and (Right) P16x16.

Method

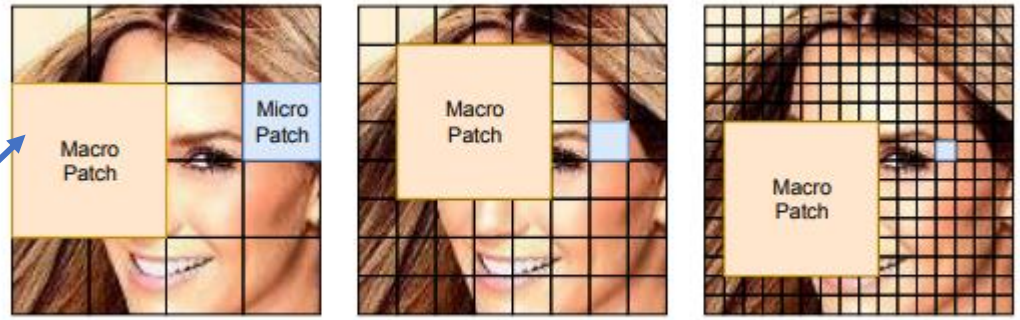
Discriminator



(a) Discriminator Overall Architecture

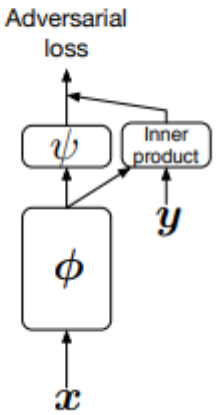


(b) Discriminator Residual Block



(a) Implementations used in this paper with (Left) P4x4, (Middle) P8x8 and (Right) P16x16.

(d) (ours) Projection



Experiments

1. Quality of Generation by Parts

(a) CelebA (N2,M2,S32) (full image: 128×128).

Original image size: 128*128
Macro patch size: 64*64 (1/4 of original image)
Micro patch size: 32*32 (1/16 of original image)

(a) CelebA (N4,M4,S16) (full image: 128×128, FID: 10.82).

(b) CelebA (N8,M8,S8) (full image: 128×128, FID: 15.99).

(c) CelebA (N16,M16,S4) (full image: 128×128, FID: 23.90).

Experiments

2. Latent Space Continuity

Figure 5: The results of full-images interpolation between two latent vectors show that all micro patches are changed synchronously in response to the change of the latent vector. More interpolation results are available in Appendix G.

Latent vector interpolation

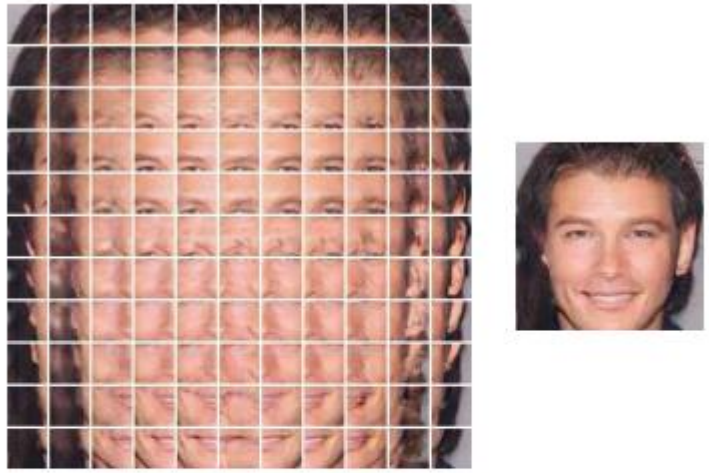


Figure 6: An example of spatial coordinates interpolation showing the spatial continuity of the micro patches. The spatial coordinates are interpolated between range $[-1, 1]$ of the micro coordinate with a fixed latent vector. More examples are shown in Appendix I.

Coordinate interpolation

Experiments

3. Beyond-Boundary Generation

Training coordinate : [-1,1]
Augmented coordinate: [-1.66, 1.66]

Post training output

Figure 7: "Beyond-Boundary Generation" generates additional contents by extrapolating the learned coordinate manifold. Note that the generated samples are 384×384 pixels, whereas ***all*** of the training samples are of a smaller 256×256 resolution. The red box annotates the 256×256 region for regular generation without extrapolation.

Thank you