Conditional Generation by GAN

李宏毅

Hung-yi Lee

Text-to-Image

a dog is running 

a bird is flying 

- **Traditional supervised approach**

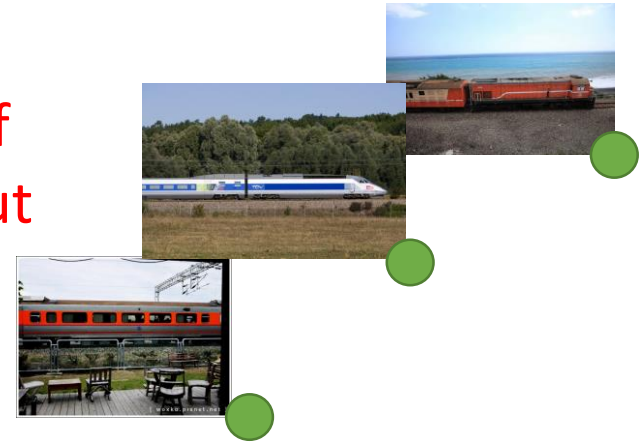
c^1: a dog is running → NN → Image ↔ 

as close as possible

Text: "train"

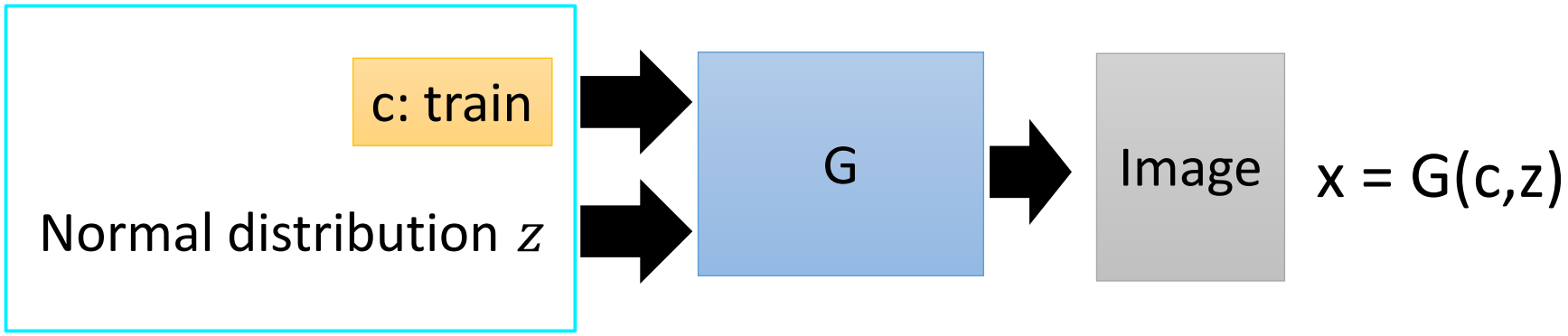
Target of NN output

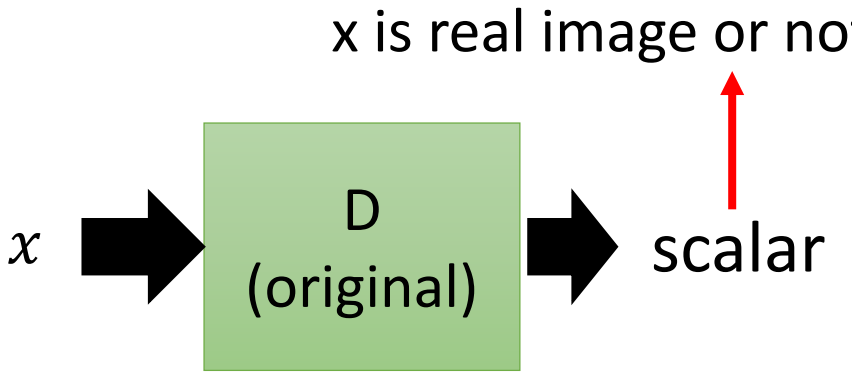
A blurry image!



[Scott Reed, et al, ICML, 2016]

Conditional GAN



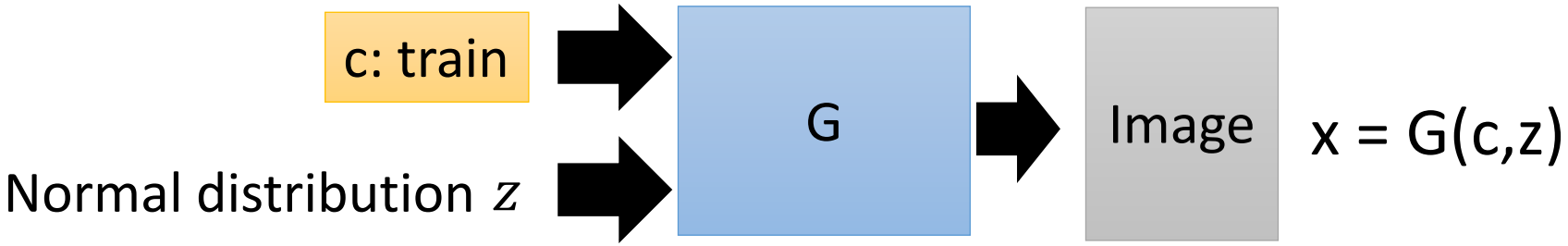


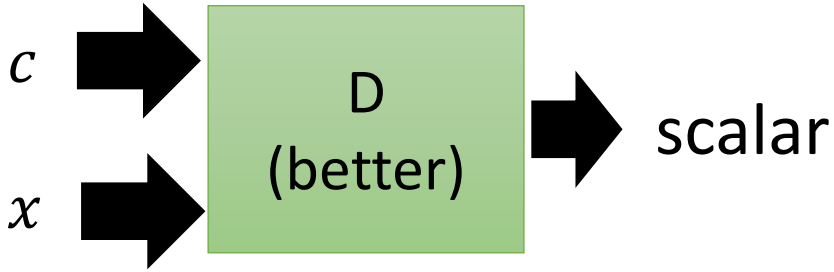
Generator will learn to generate realistic images ….

But completely ignore the input conditions.

[Scott Reed, et al, ICML, 2016]

Conditional GAN





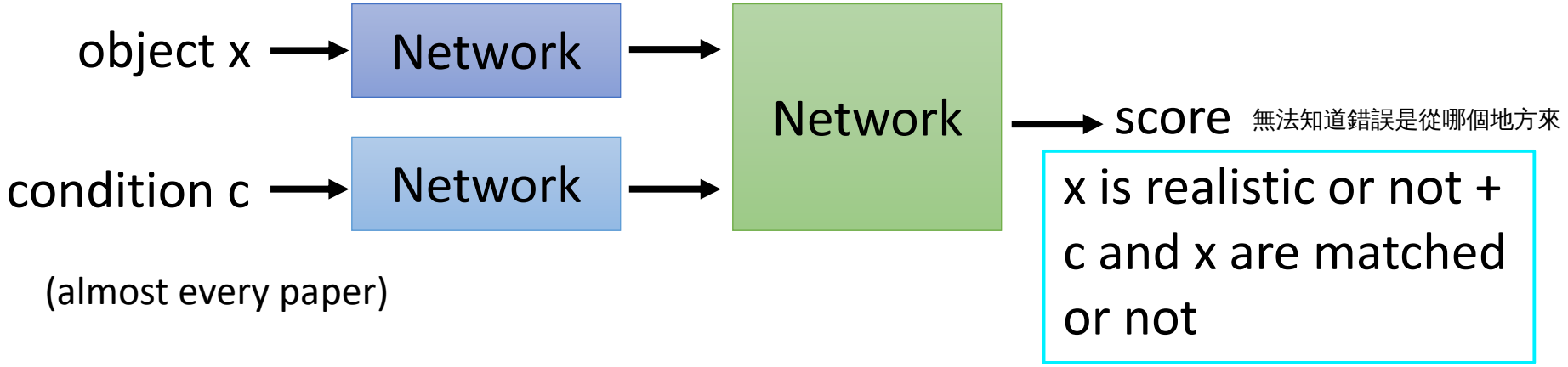
x is realistic or not +
c and x are matched or not

True text-image pairs: (train , ) 1

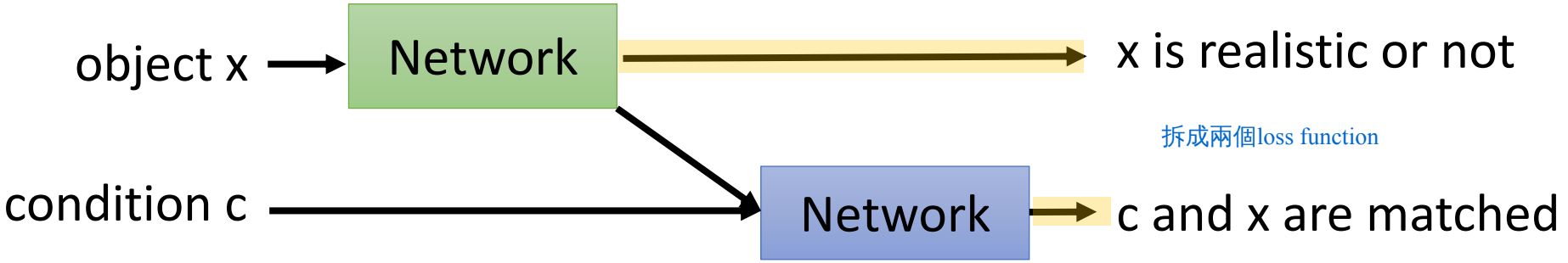
(cat , ) 0

(train , Image) 0

Conditional GAN - Discriminator



(almost every paper)



[Augustus Odena et al., ICML, 2017]

[Takeru Miyato, et al., ICLR, 2018]

[Han Zhang, et al., arXiv, 2017]

Conditional GAN

The images are generated by Yen-Hao Chen, Po-Chun Chien, Jun-Chen Xie, Tsung-Han Wu.

paired data

blue eyes
red hair
short hair

Collecting anime faces and the description of its characteristics

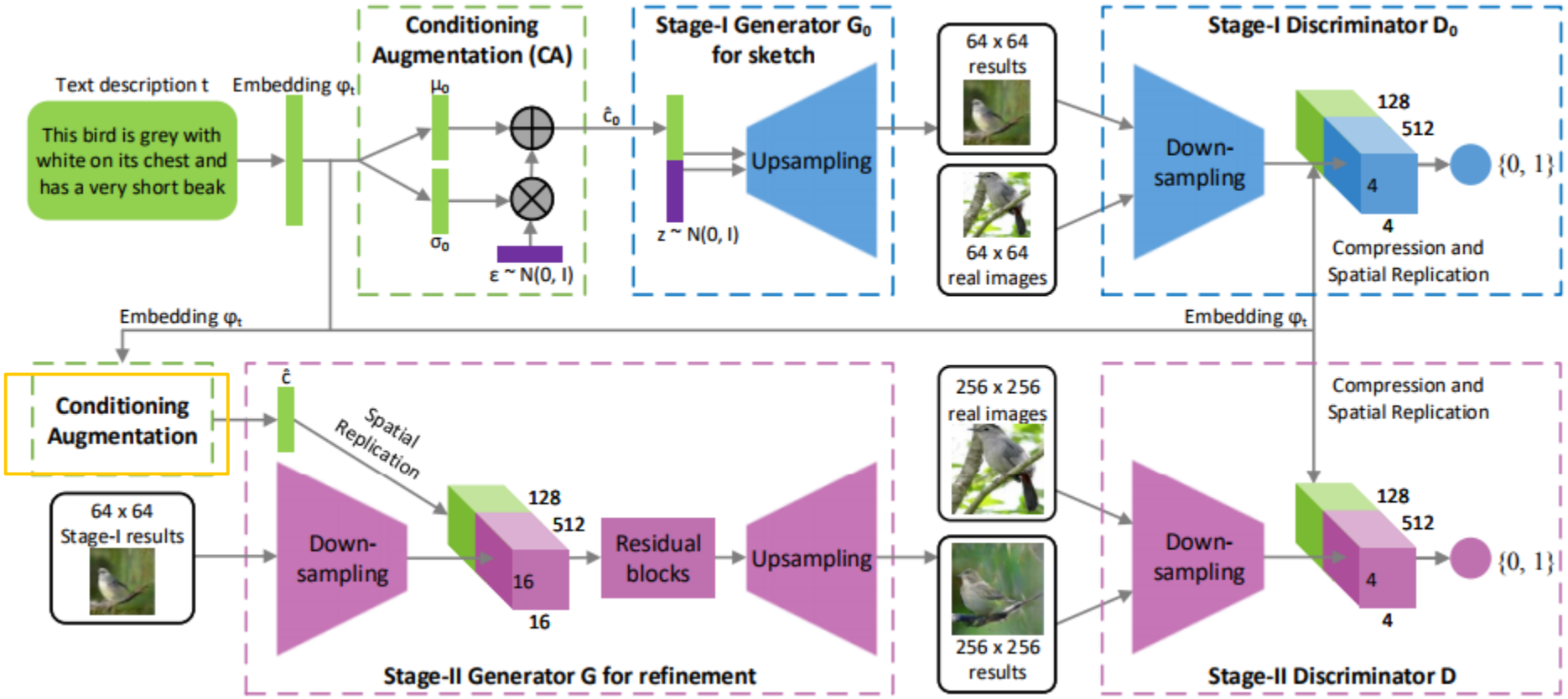
red hair, green eyes

blue hair, red eyes

Stack GAN

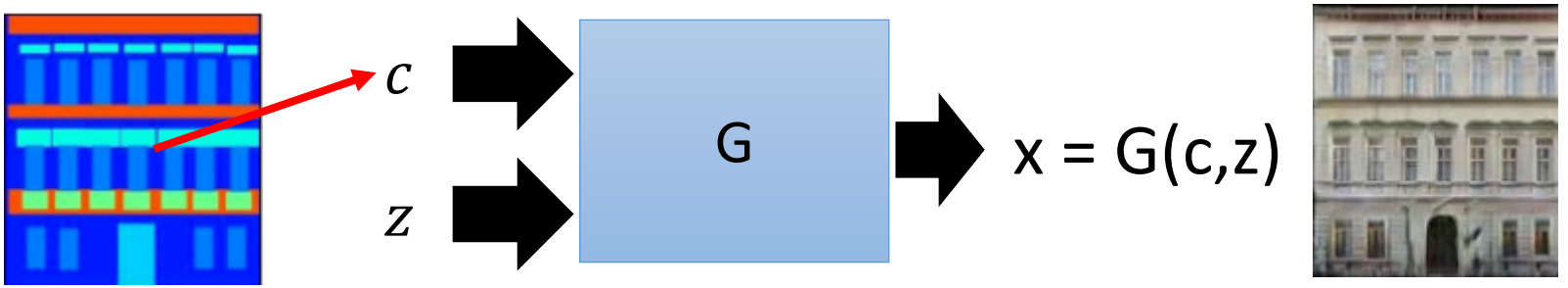
Han Zhang, Tao Xu, Hongsheng Li, Shaoting Zhang, Xiaogang Wang, Xiaolei Huang, Dimitris Metaxas, "StackGAN: Text to Photo-realistic Image Synthesis with Stacked Generative Adversarial Networks", ICCV, 2017

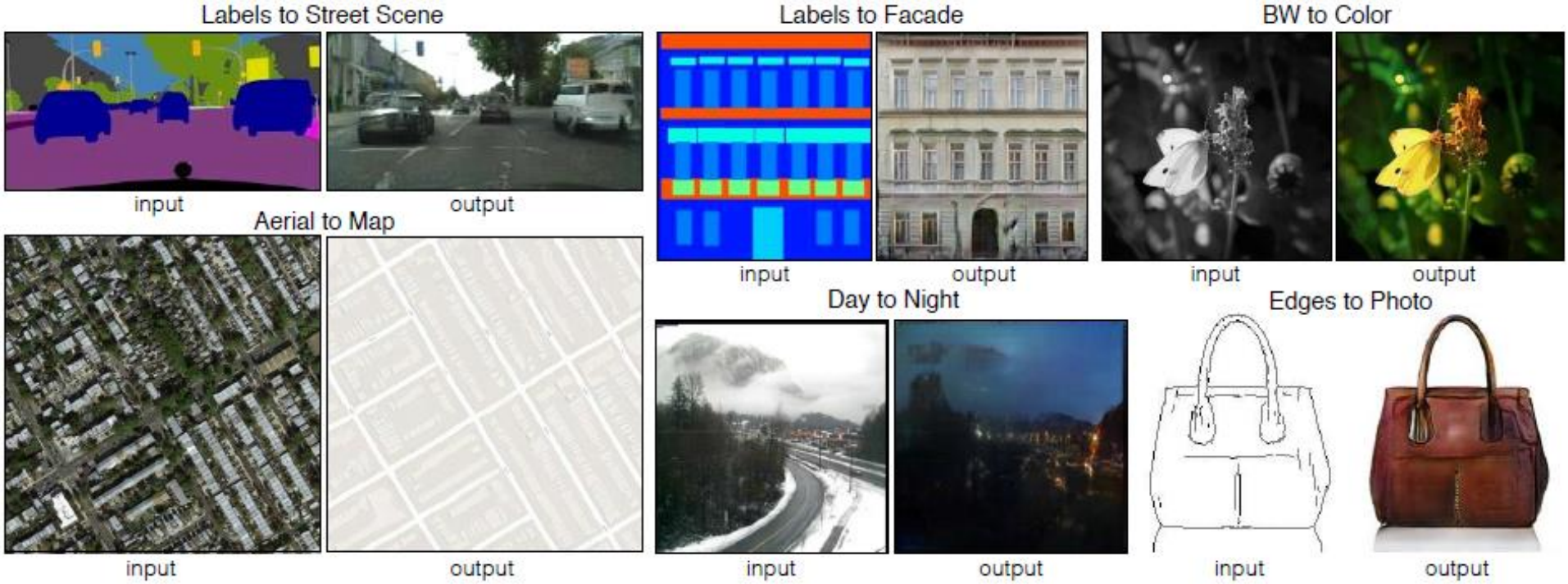
大圖整個loss function



先train小圖 再將小圖產生大圖 在把兩個一起train

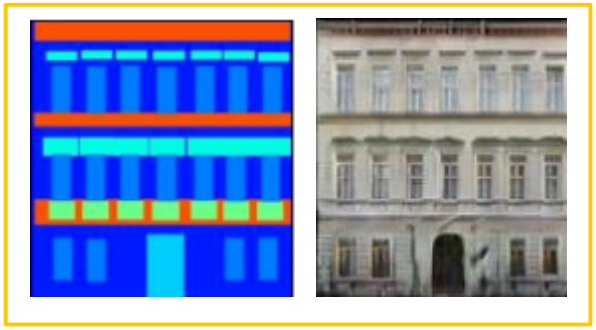
Image-to-image





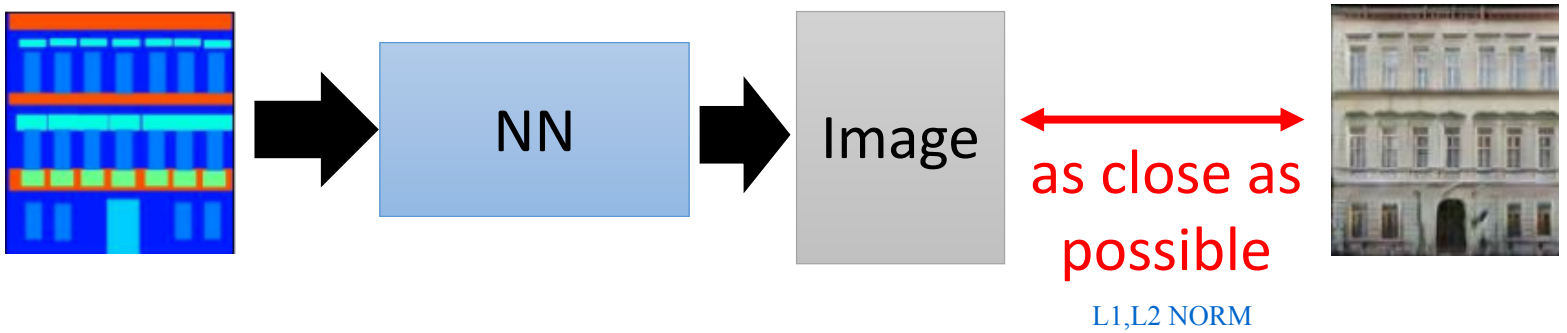
https://arxiv.org/pdf/1611.07004

Image-to-image

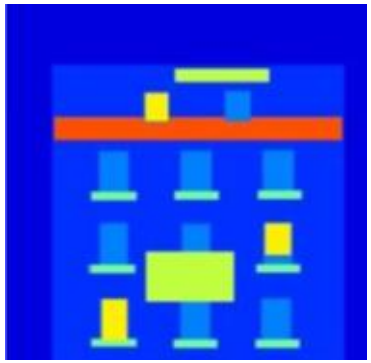


data

- Traditional supervised approach



Testing:

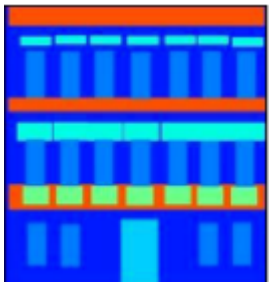


input

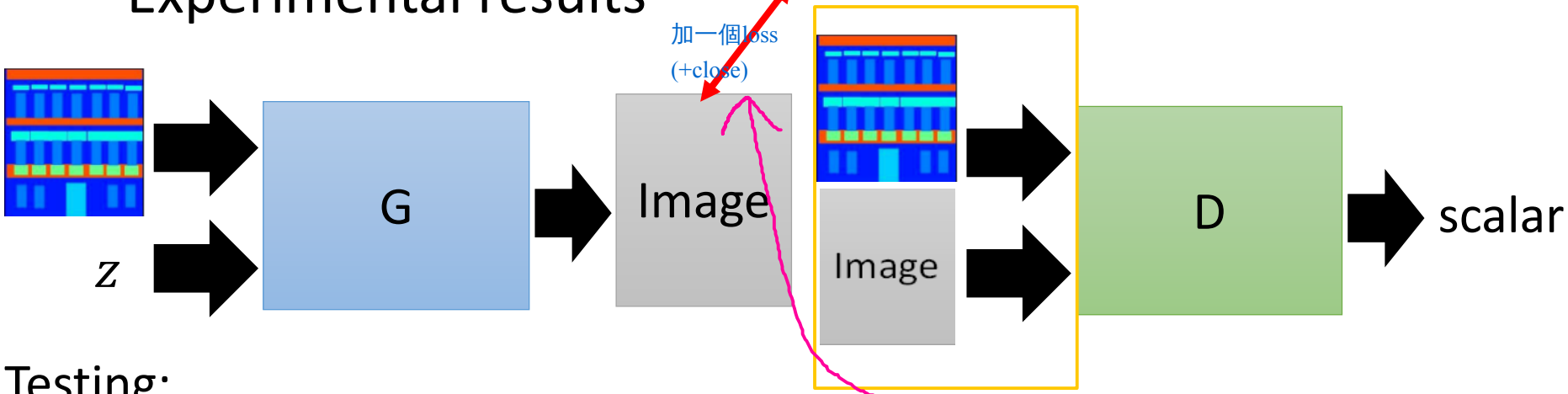
close

It is blurry because it is the average of several images.

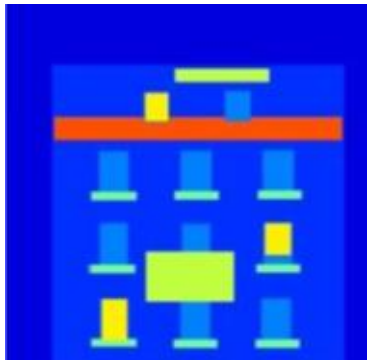
Image-to-image



- Experimental results



Testing:



input

close

GAN

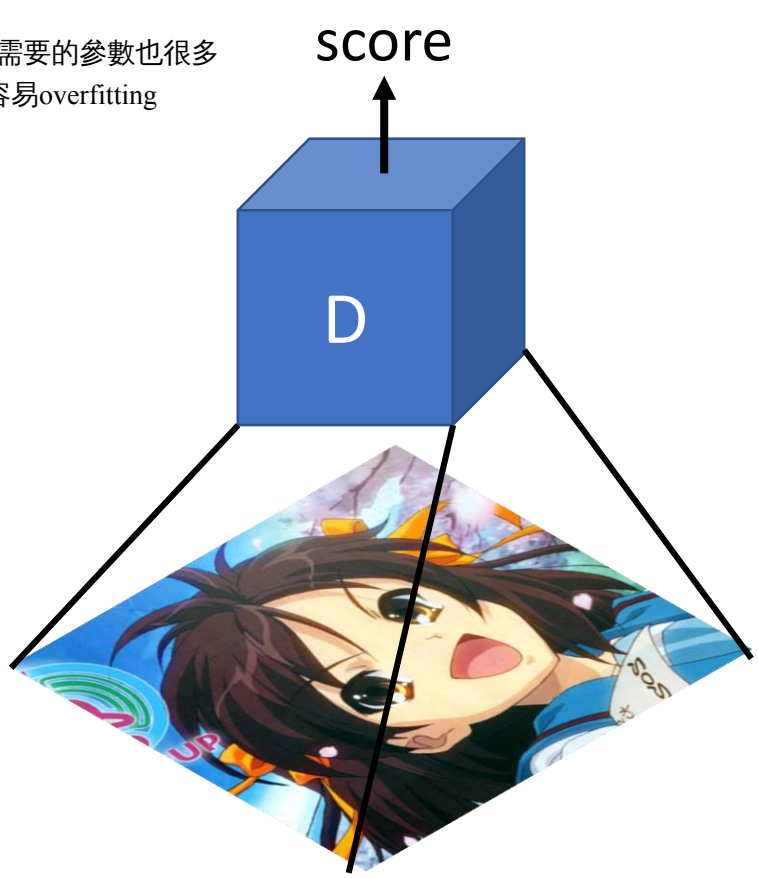
GAN + close

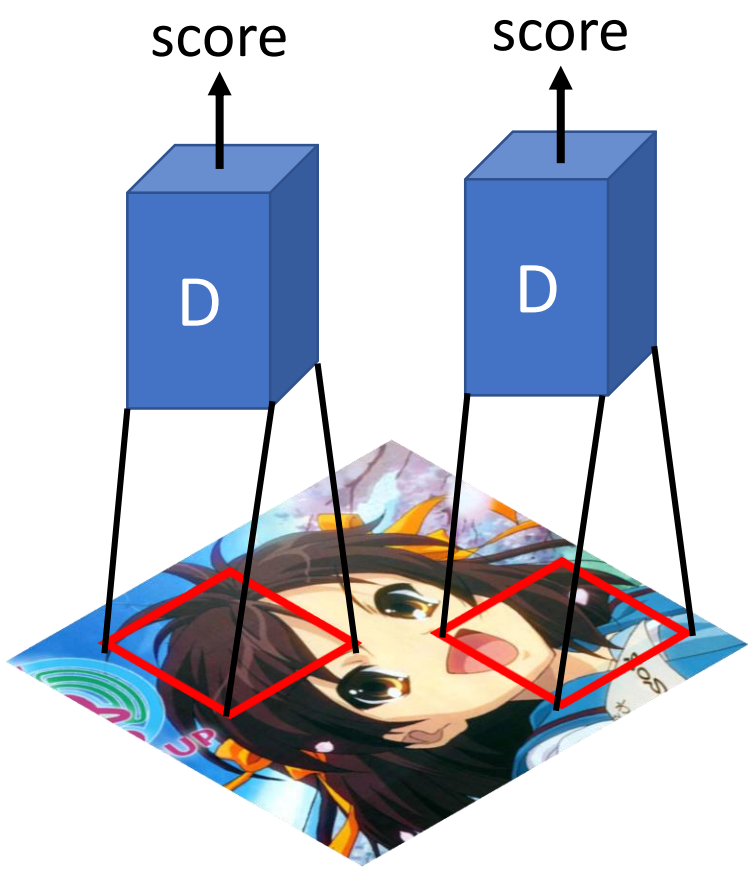
希望G產生的output

Patch GAN

https://arxiv.org/pdf/1611.07004.pdf

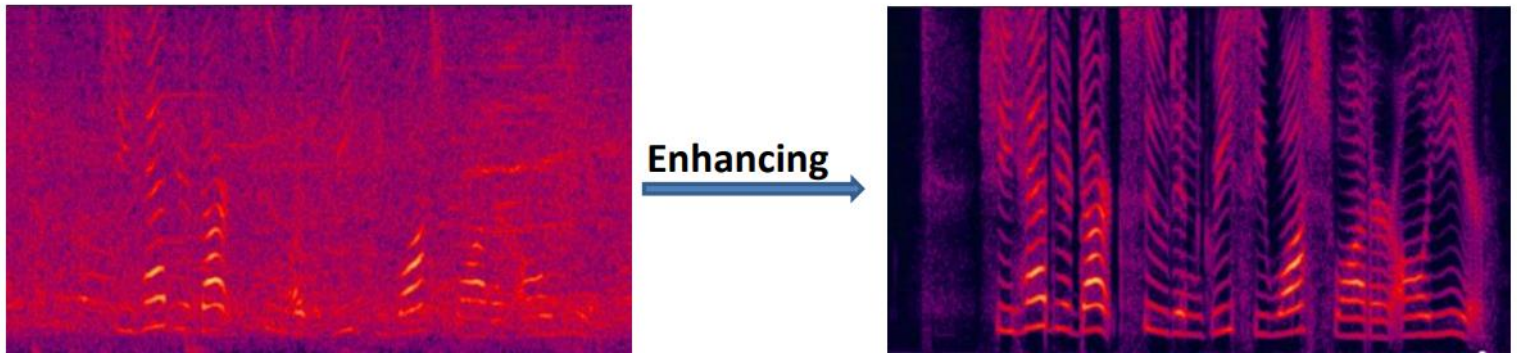
image大 需要的參數也很多
會太多 容易overfitting



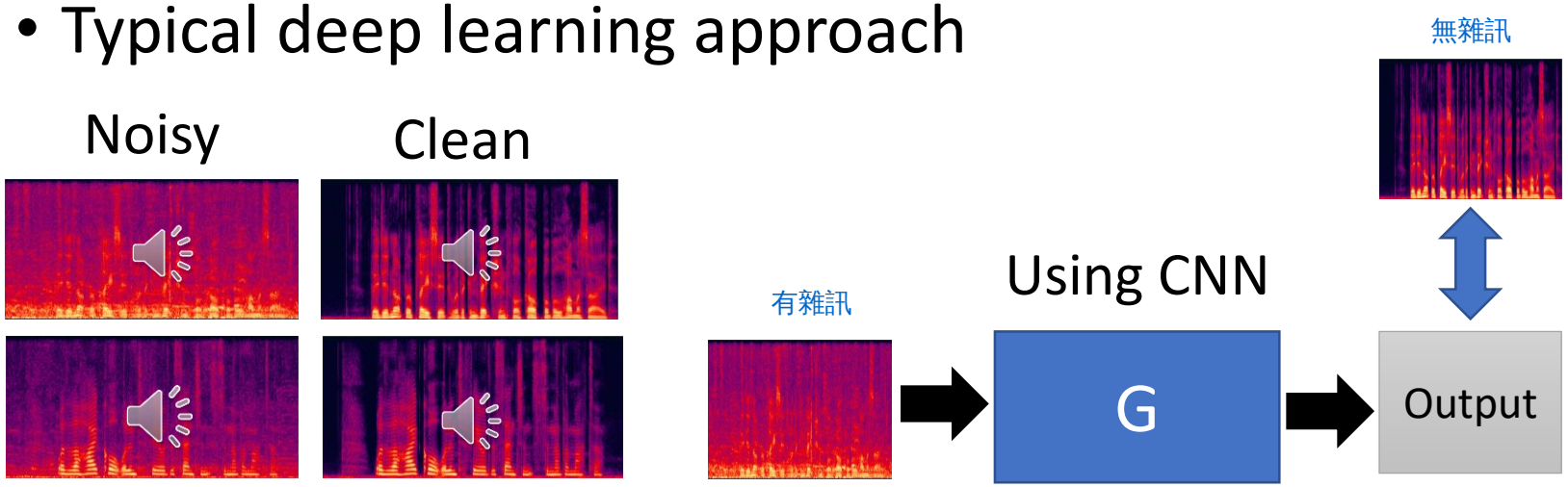


檢查一小塊圖片 就好: 決定一次要檢查(loss)多大????只看一個pixel也是不行的~~~會太過模糊

Speech Enhancement

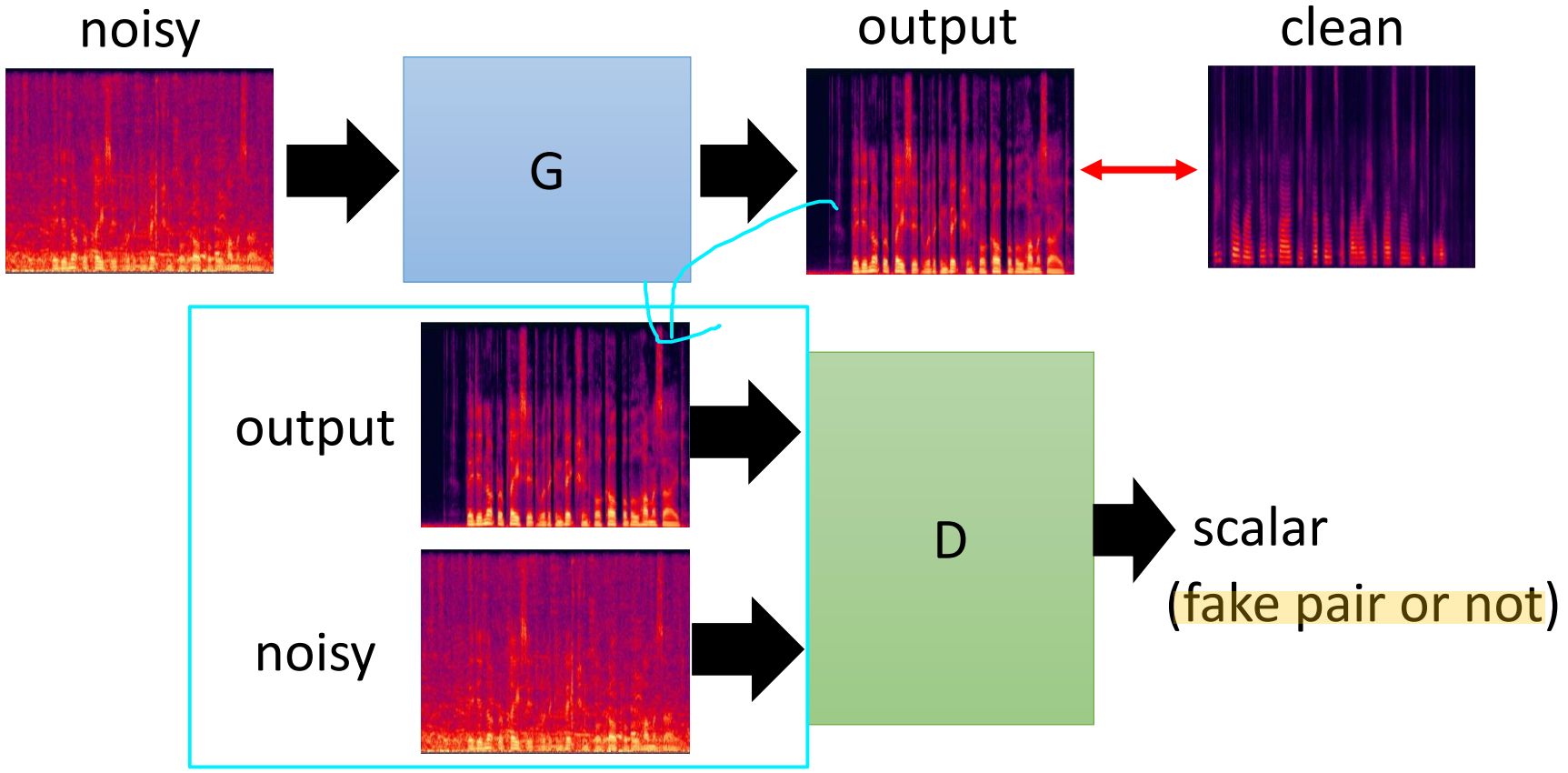


- Typical deep learning approach

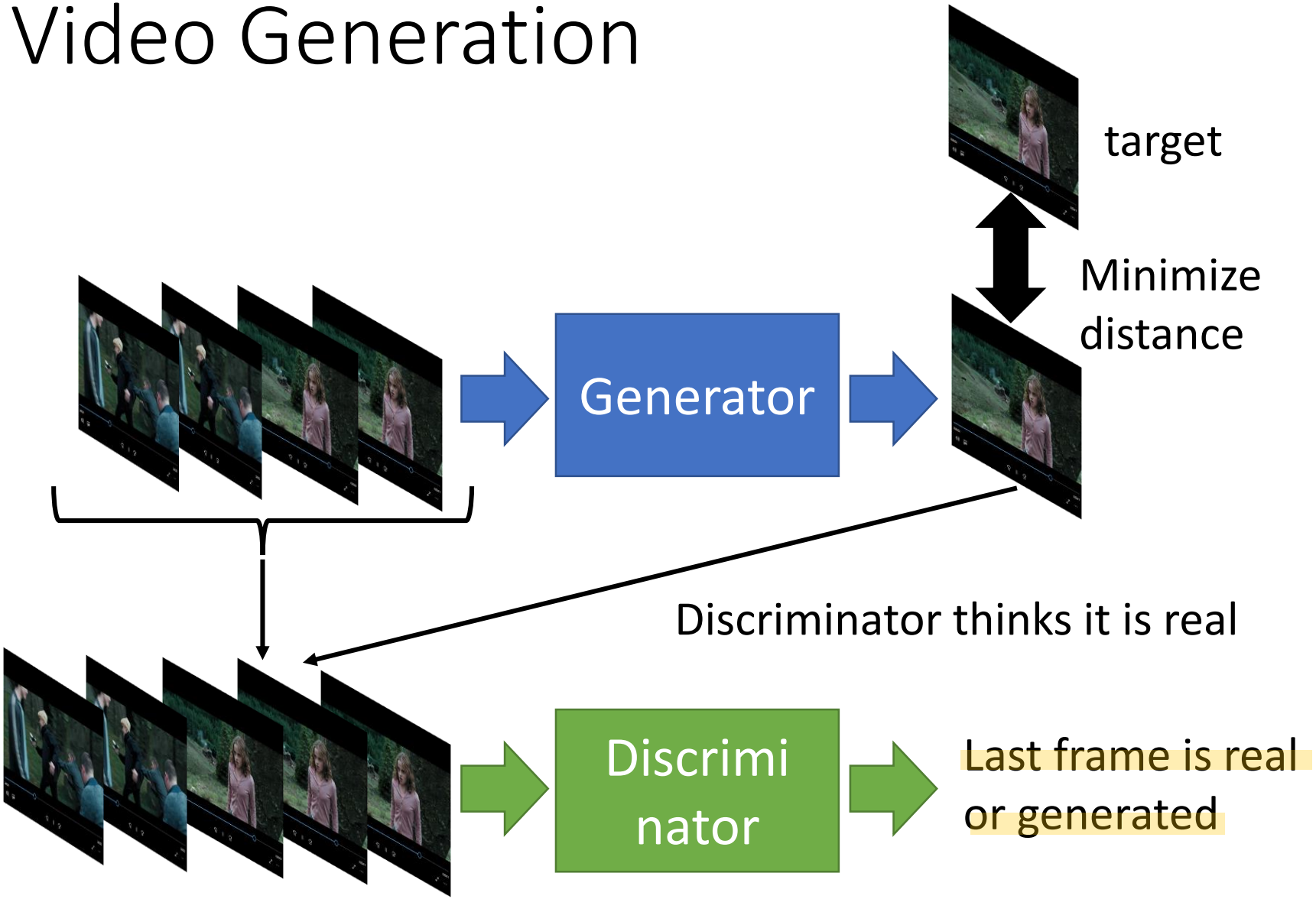


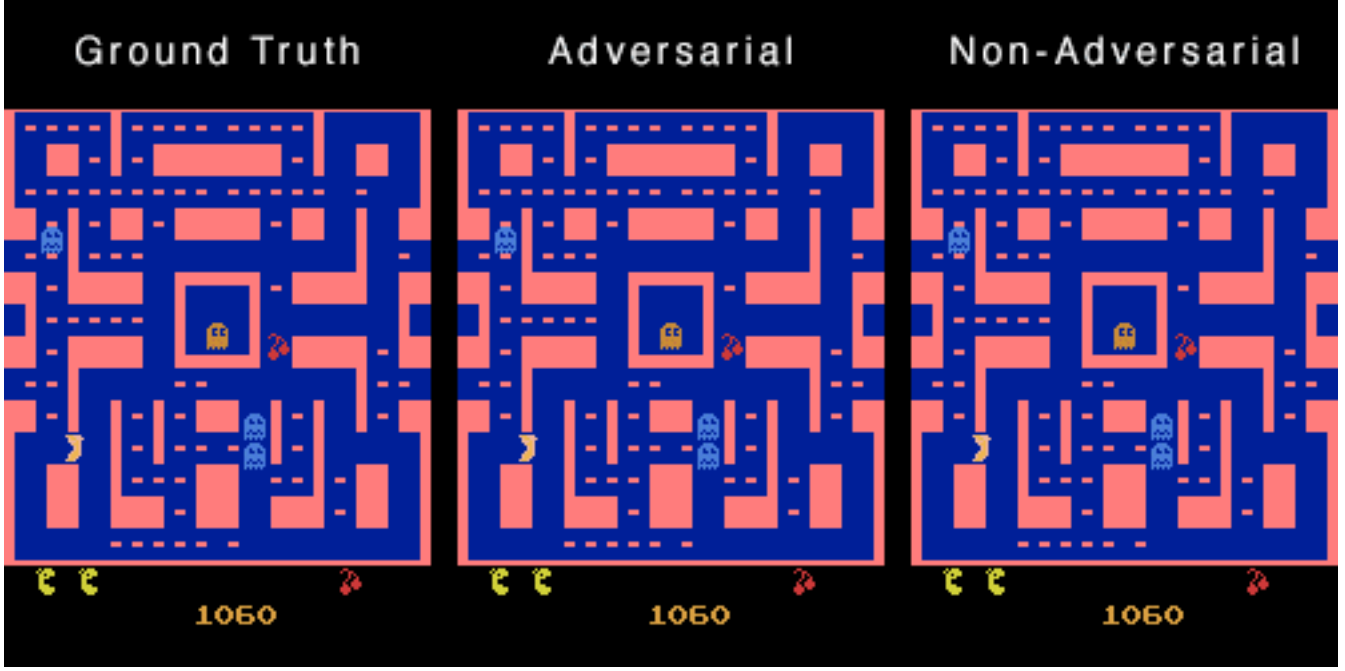
Speech Enhancement

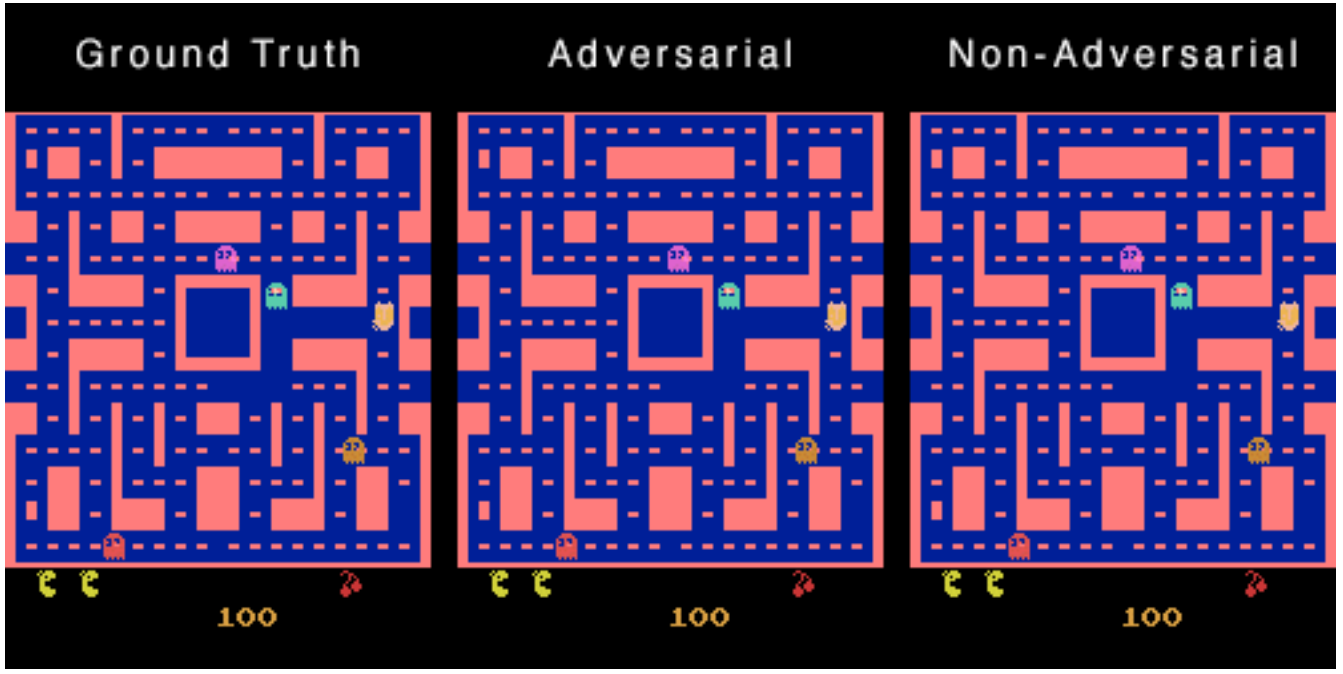
- Conditional GAN



Video Generation







https://github.com/dyelax/Adversarial_Video_Generation