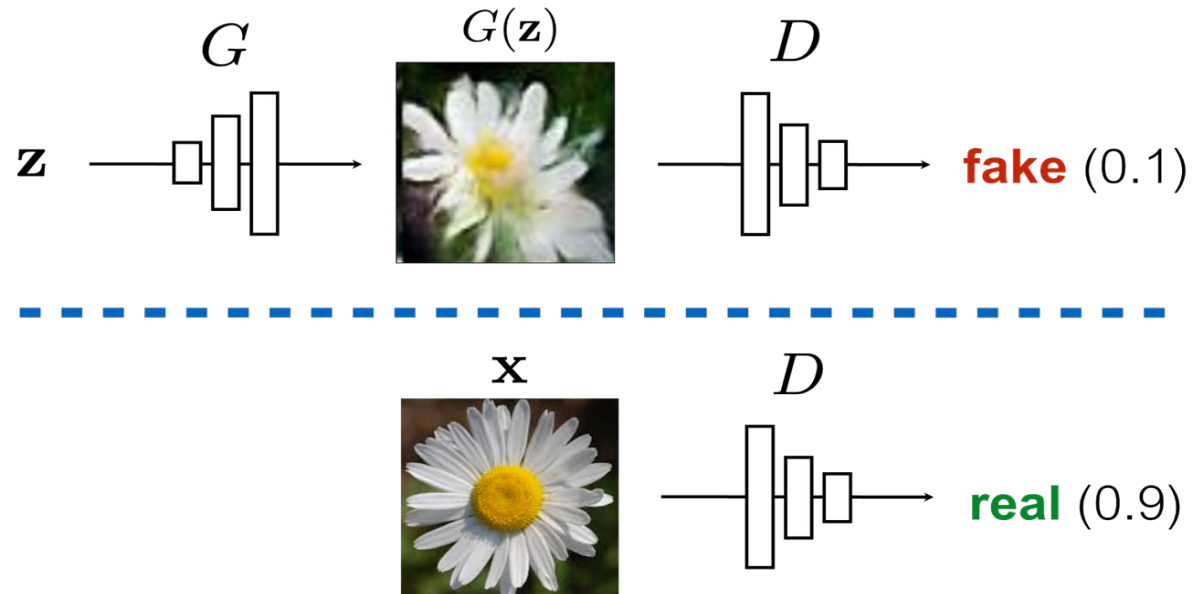


Extra Slides for Generative Models

by Margrit Betke, CS 585, April 4, 2024

Hao on GANs last time:

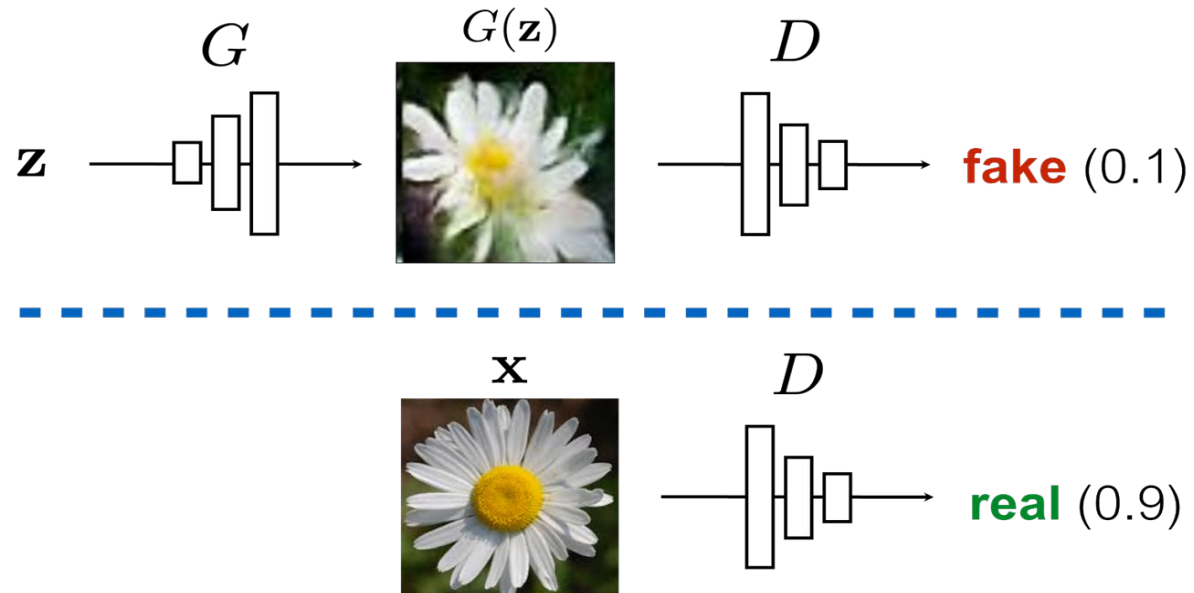


$$\arg \max_D \mathbb{E}_{\mathbf{x} \sim p_{\text{data}}(\mathbf{x})} [\log D(\mathbf{x})] + \mathbb{E}_{\mathbf{z} \sim p_z(\mathbf{z})} [\log(1 - D(G(\mathbf{z})))]$$

[\[Goodfellow et al., 2014\]](#)

Hao on GANs last time:

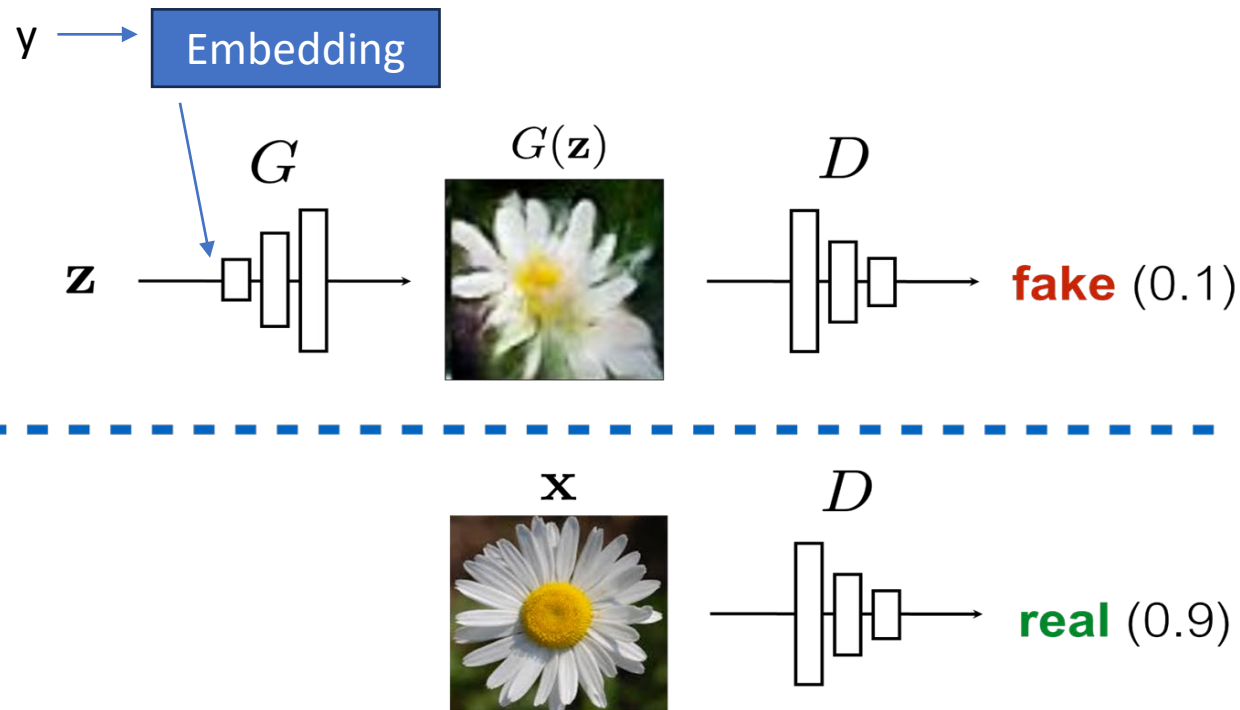
How can you make the model create an image of a daisy flower?



$$\arg \max_D \mathbb{E}_{\mathbf{x} \sim p_{\text{data}}(\mathbf{x})} [\log D(\mathbf{x})] + \mathbb{E}_{\mathbf{z} \sim p_z(\mathbf{z})} [\log(1 - D(G(\mathbf{z})))]$$

[Goodfellow et al., 2014]

Hao on GANs last time:



How can you make the model create an image of a daisy flower?

Use a conditional GAN:
[Mirza and Osindero, 2014](#)

y ="daisy flower"

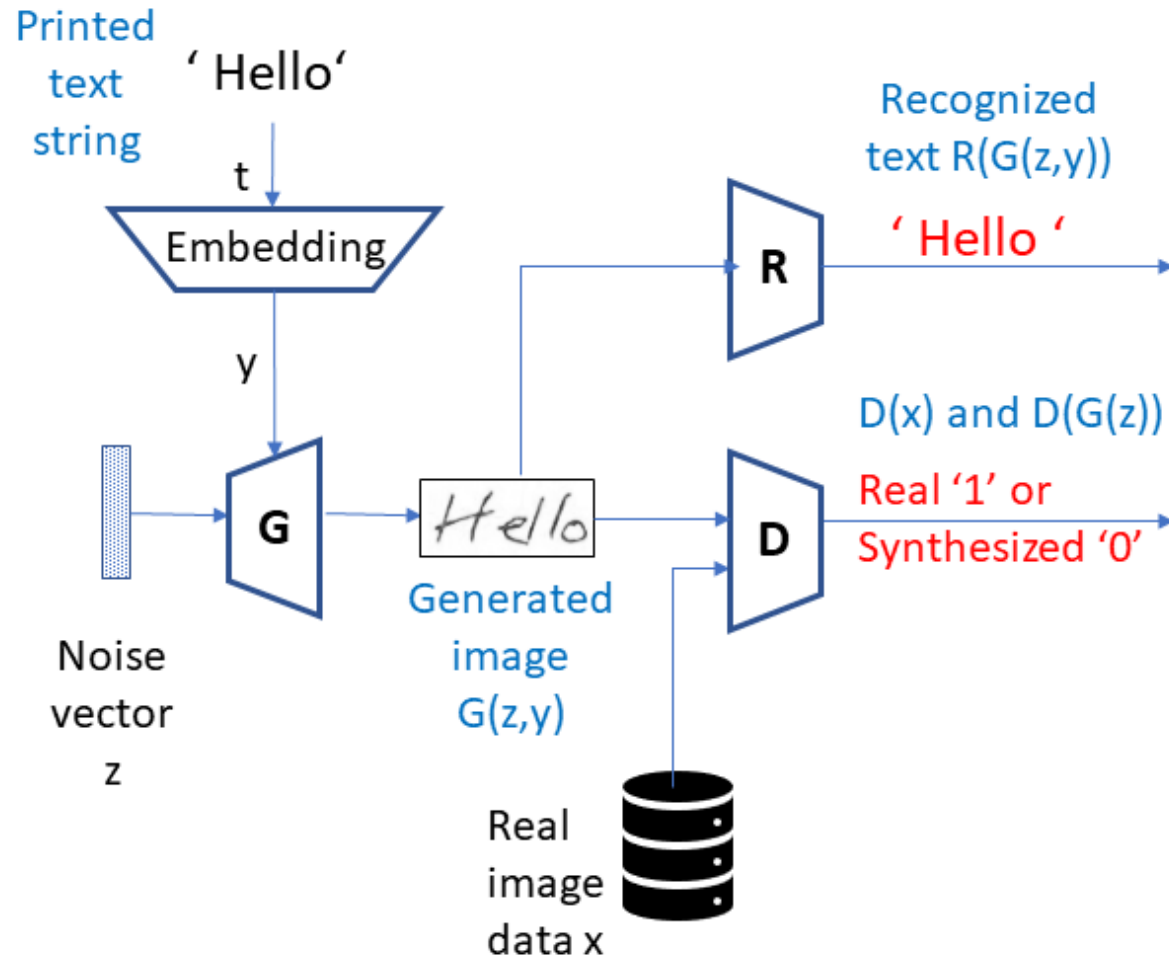
$$\arg \max_D \mathbb{E}_{\mathbf{x} \sim p_{\text{data}}(\mathbf{x})} [\log D(\mathbf{x})] + \mathbb{E}_{\mathbf{z} \sim p_z(\mathbf{z})} [\log(1 - D(G(\mathbf{z})))]$$
$$: \mathbb{E}_{\mathbf{x} \sim p_{\text{data}}(\mathbf{x})} [\log D(\mathbf{x}|\mathbf{y})] + \mathbb{E}_{\mathbf{z} \sim p_z(\mathbf{z})} [\log(1 - D(G(\mathbf{z}|\mathbf{y})))]$$

[\[Goodfellow et al., 2014\]](#)

Where are GANs still better than
Diffusion Models
(as of early 2024)?

Handwriting Generation

Alonso et al.'s Seminal GAN for Handwriting Synthesis, 2019



Add a **text recognition network R**

Here: gated convolutional recurrent network (CRNN), consisting of an encoder of five layers, with tanh activations and convolutional gates, followed by a max pooling layer, and a decoder made up of two stacked bidirectional LSTM layers

$$l_{D,G} = \mathbb{E}_x [\log(D(x|y))] + \mathbb{E}_z [\log(1 - D(G(z|y)))]$$

$$l_R = \mathbb{E}_{(z,t)} [CTC(t, R(G(z, y)))],$$

CTC: Connectionist Temporal Classification algorithm

Works extending Alonso et al.'s 2019 model:

- Text
- Writing style
- Variable length words
- Paragraphs

	Single word	Arbitrary text
SCRABBLEGAN		
HiGAN		
Davis et al.		
JokerGAN		
HiGAN+		
SLOGAN		

Diffusion Model for Handwriting?

