



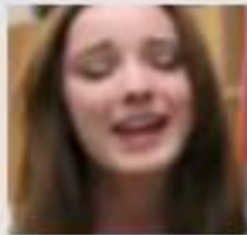
# **FEW-SHOT ADVERSARIAL LEARNING**

of

## **Realistic Neural Talking Head Models**

**Francis Bello, Mia Bragilovski, Larry Li, Mark Smith**

# CONCEPT



Source

Target → Landmarks → Result



# APPLICATIONS IN BUSINESS



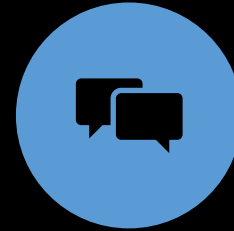
**GAMING  
INDUSTRY**



**VIRTUAL TOUR  
GUIDES**



**VIDEO  
CONFERENCING:  
INTERVIEWS /  
PUBLIC  
ENGAGEMENTS**



**PERSONAL  
ASSISTANTS /  
EMBODIED  
CONVERSATIONAL  
AGENTS**



**PERSONALIZED  
MANAGERS**

# APPLICATIONS IN HEALTH AND EDUCATION



## Health

### Model positive interactions :

- ➡ Communication for differently abled
- ➡ Alleviation of social anxiety



## Education

### Choose your tutors :

- ➡ LeBron James teaches science
- ➡ Napoleon gives history lessons
- ➡ Affordable, Anywhere, Anytime



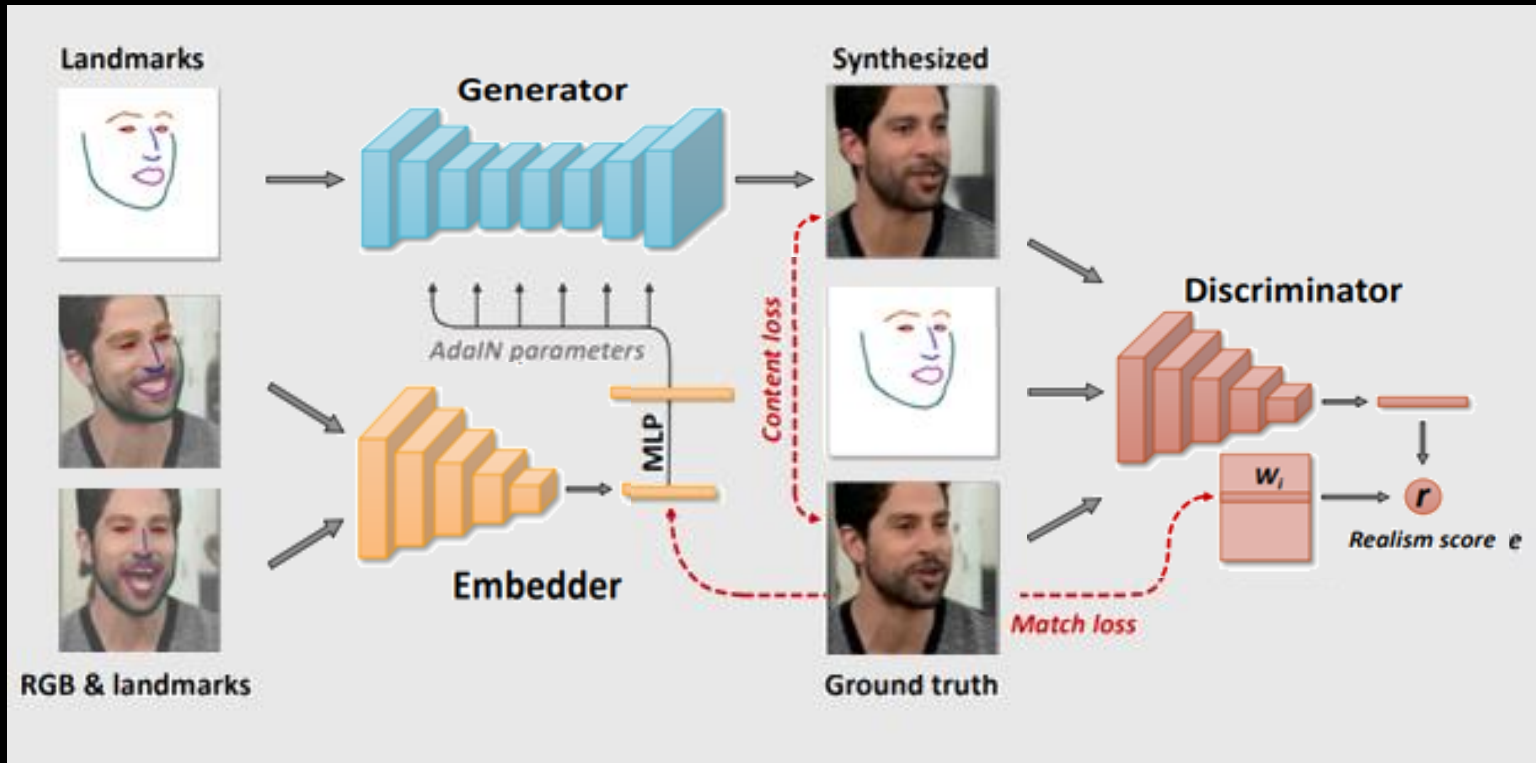
# Architecture Overview

Two training phases:

- Meta-learning
- Fine-tuning

Limitation:

- Landmarks do not represent the subject's gaze



# Architecture Overview

Two training phases:

- Meta-learning
- Fine-tuning

Networks:

- Embedder
- Generator
- Discriminator

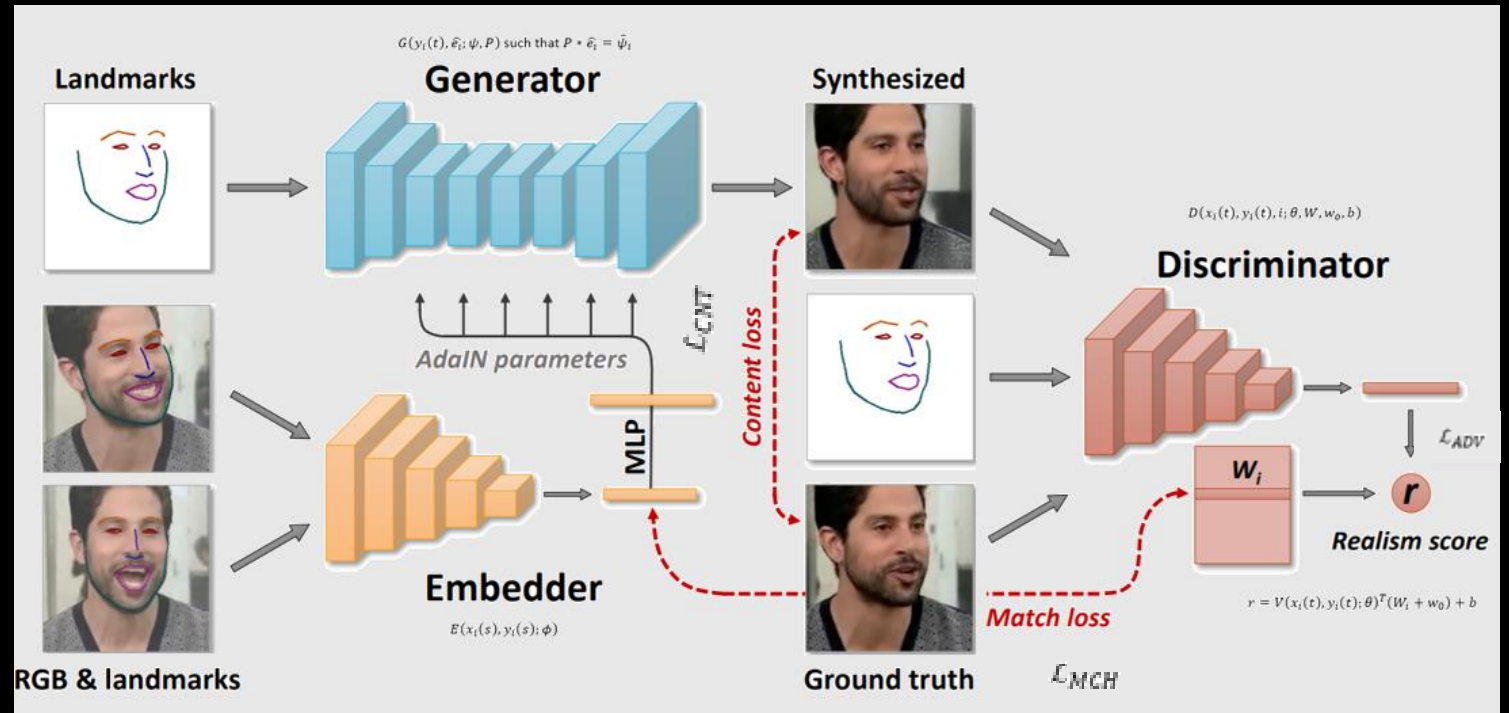


# Meta-Learning Phase

- Trains all 3 networks together in an adversarial way
- The parameters of the Embedder and Generator networks are updated to minimize the objective

$$\mathcal{L} = \mathcal{L}_{CNT} + \mathcal{L}_{ADV} + \mathcal{L}_{MCH}:$$

- $\mathcal{L}_{CNT}$  = content loss
- $\mathcal{L}_{ADV}$  = adversarial loss
- $\mathcal{L}_{MCH}$  = match loss



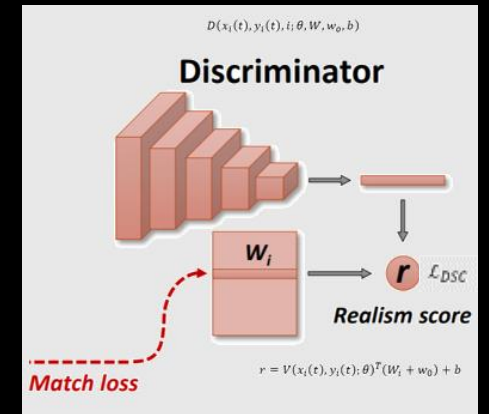
# Meta-Learning Phase

- Trains all 3 networks together in an adversarial way
- The parameters of the Embedder and Generator networks are updated to minimize the objective

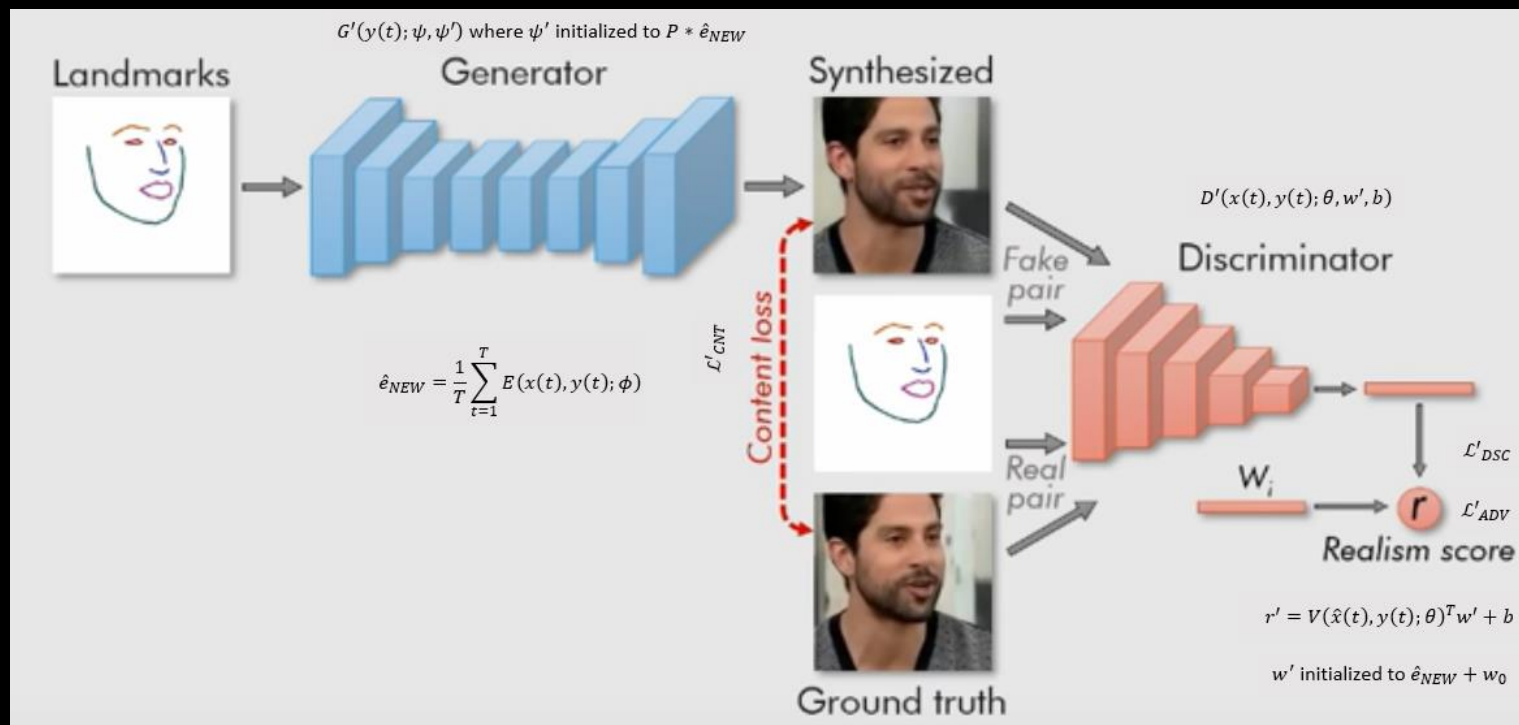
$$\mathcal{L} = \mathcal{L}_{CNT} + \mathcal{L}_{ADV} + \mathcal{L}_{MCH}:$$

- $\mathcal{L}_{CNT} = \text{content loss}$
- $\mathcal{L}_{ADV} = \text{adversarial loss}$
- $\mathcal{L}_{MCH} = \text{match loss}$

- The parameters of the Discriminator are updated to minimize the objective  $\mathcal{L} = \mathcal{L}_{DSC}$ :
  - $\mathcal{L}_{DSC} = \text{hinge loss}$
- Generator network gets better at producing realistic images while the Discriminator network gets better at determining which images are real and which are synthesized
  - *Parameter updates alternate between the Embedder / Generator and the Discriminator*







## Fine-Tuning Phase

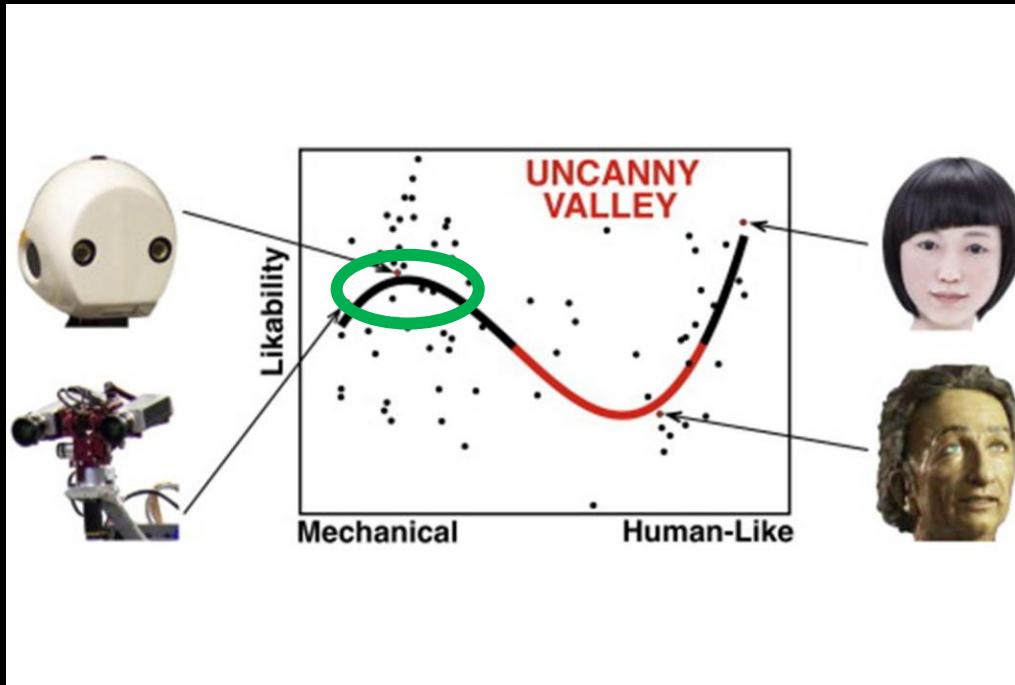
- After meta-learning has converged
- Only the Generator and Discriminator are trained
- Trains on  $T$  new images (could be as few as one!)
- Training is done in an analogous way to the meta-learning phase
- After training has converged, you can pass in a sequence of landmarks and synthesize a deep fake video
- With 32 images, the model achieves perfect realism



## How digital avatars are made

- Identify the target
- Collect data
- Reconstruct
- Manual Enhancement & detailing

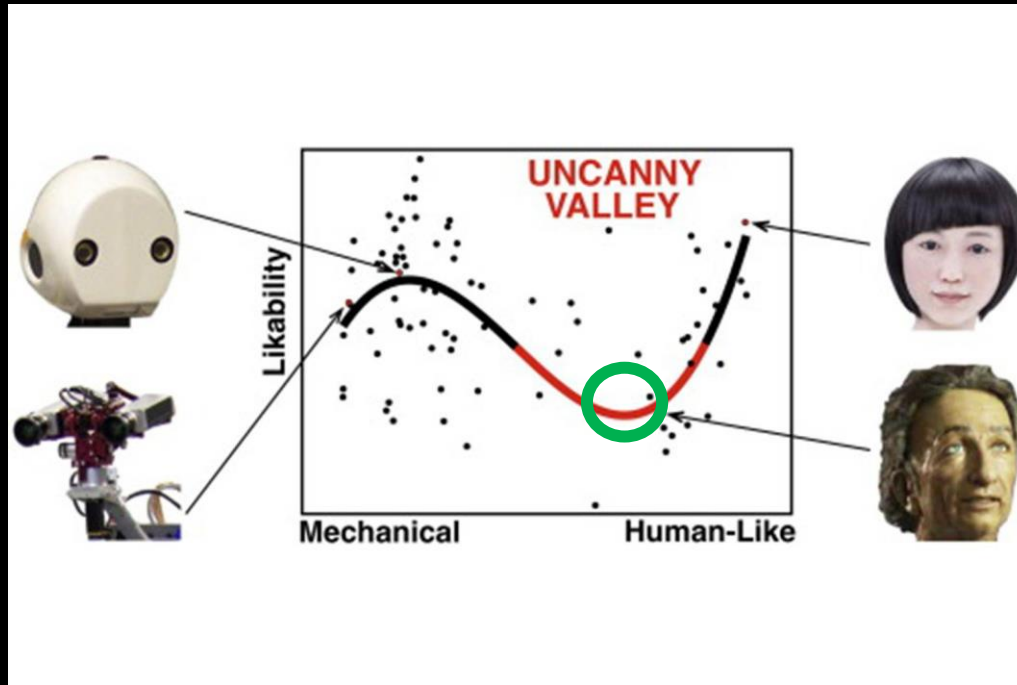




# Uncanny Valley

- **Mechanical/Cartoonish**

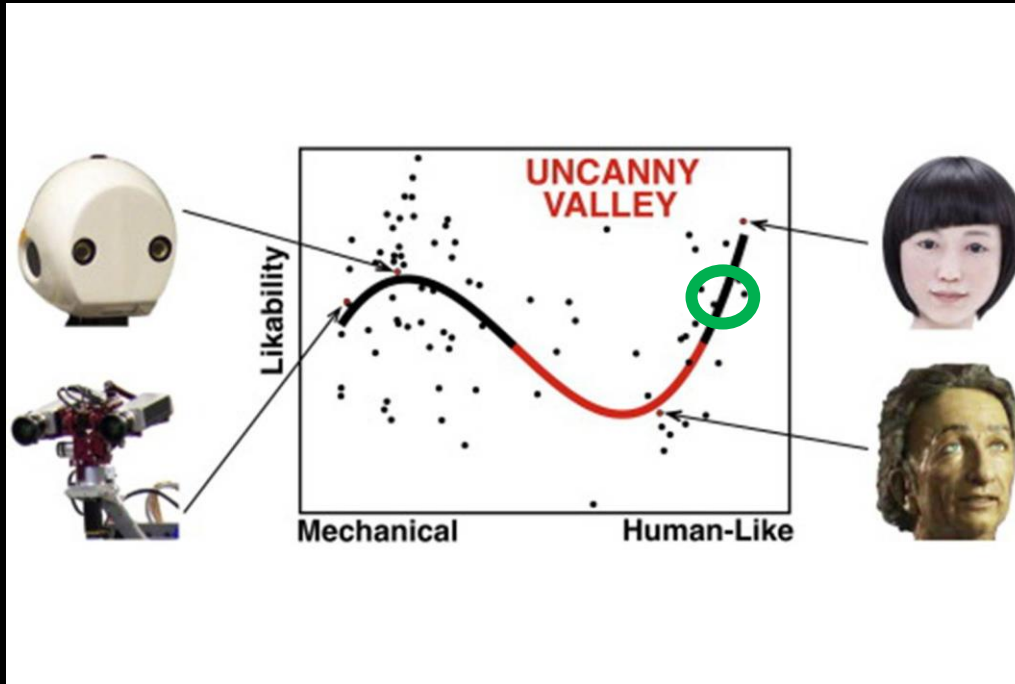




# Uncanny Valley

- **Mechanical/Cartoonish**
- **Try to be Human-Like**



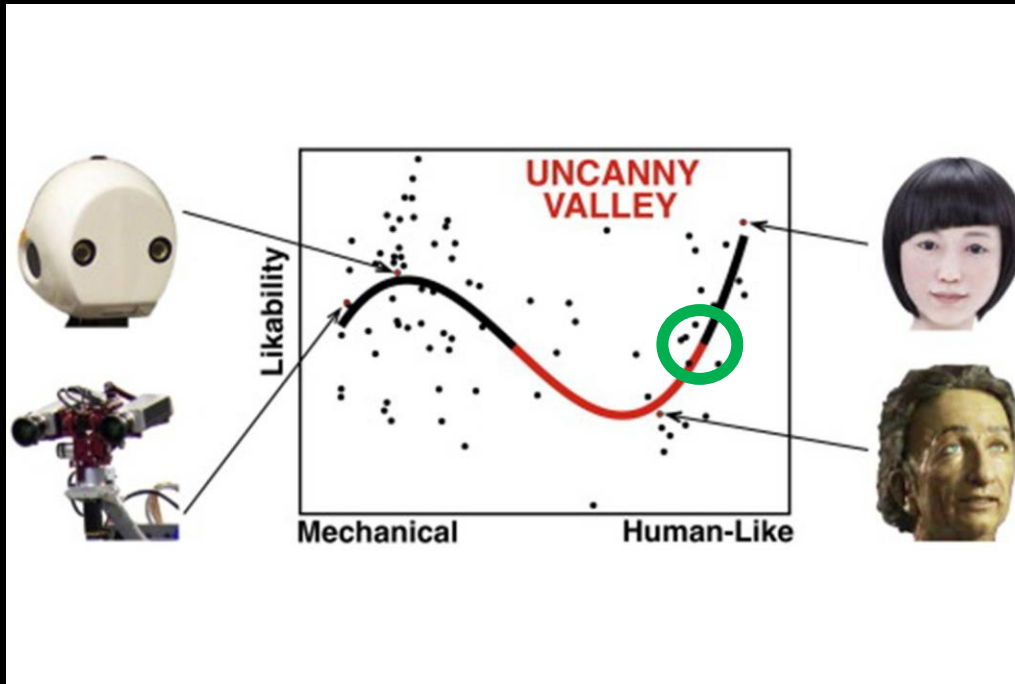


# Uncanny Valley

- **Mechanical/Cartoonish**
- **Try to be Human-Like**
- **Photo realistic**



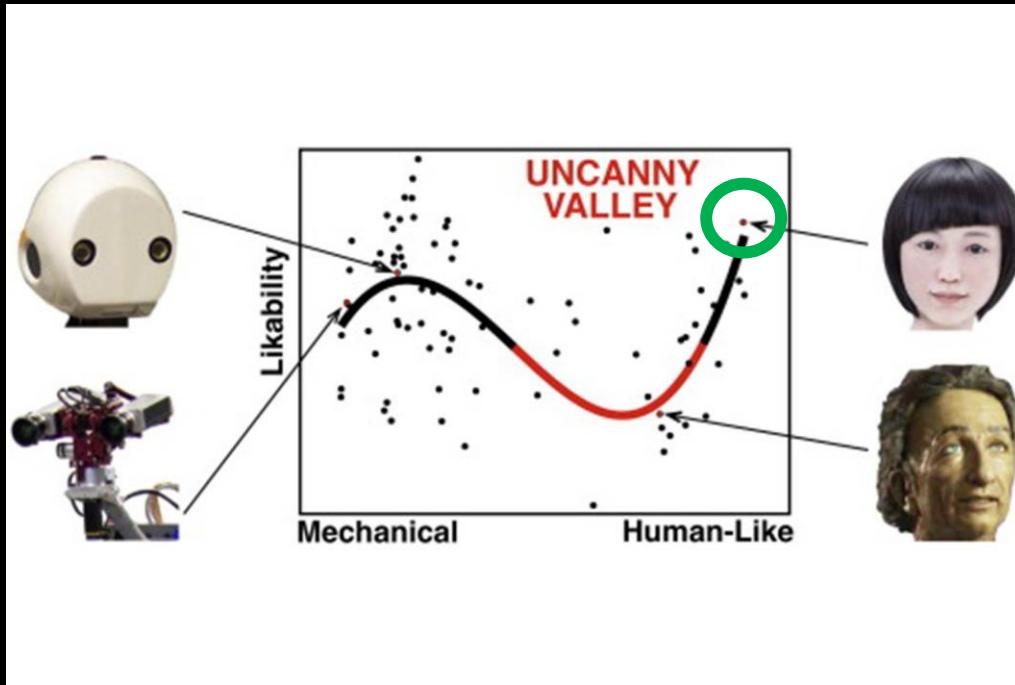




# Uncanny Valley

- **Mechanical/Cartoonish**
- **Try to be Human-Like**
- **Photo realistic**
- **Where is Few-Shot ?**





# Uncanny Valley

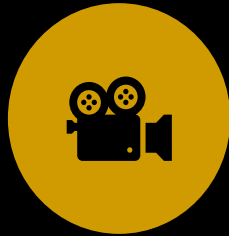
- **Mechanical/Cartoonish**
  - **Try to be Human-Like**
  - **Photo realistic**
- 
- **Where is Few-Shot ?**
  - **Climbing up**



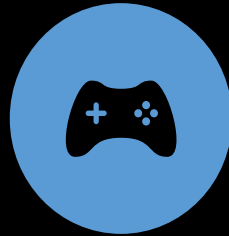
## BENEFITS: RECAP



**PERSONAL  
MANAGERS /  
ASSISTANTS**



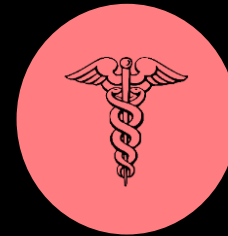
**VIDEO  
CONFERENCING**



**GAMING  
INDUSTRY**



**VIRTUAL TOUR  
GUIDES**



**HEALTH**



**EDUCATION**



**HAVE WE ACHIEVED DIGITAL IMMORTALITY?**

**THERE ARE RISKS!**

# RISKS AND ISSUES



**Raises Ethical Issues**



**Discredit**



**Proliferation of Fake News**



**Incriminate**



**Impact on Reputation**



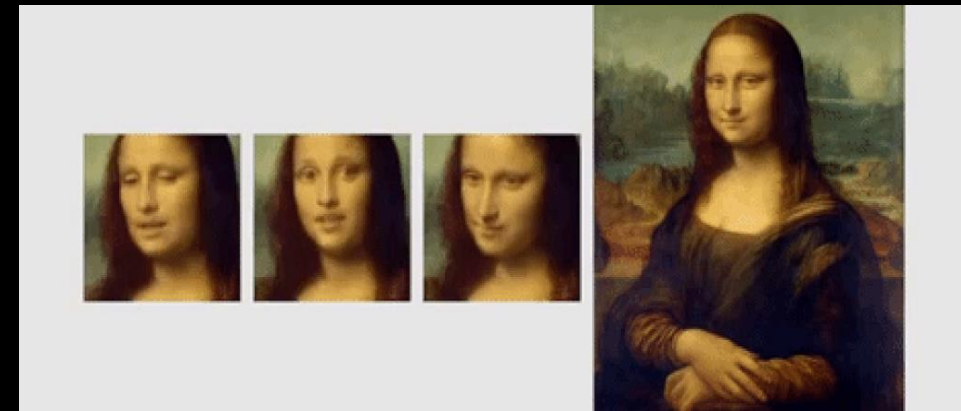
**Change History/"Lost Tapes"**

# WITH ONLY A FEW SHOTS

Living portraits



Living portraits



**WE GARNER TOTAL REPRESENTATIONAL POWER OVER ANOTHER PERSON.  
THE LINE BETWEEN FANTASY AND REALITY HAS BECOME BLURRIER.**