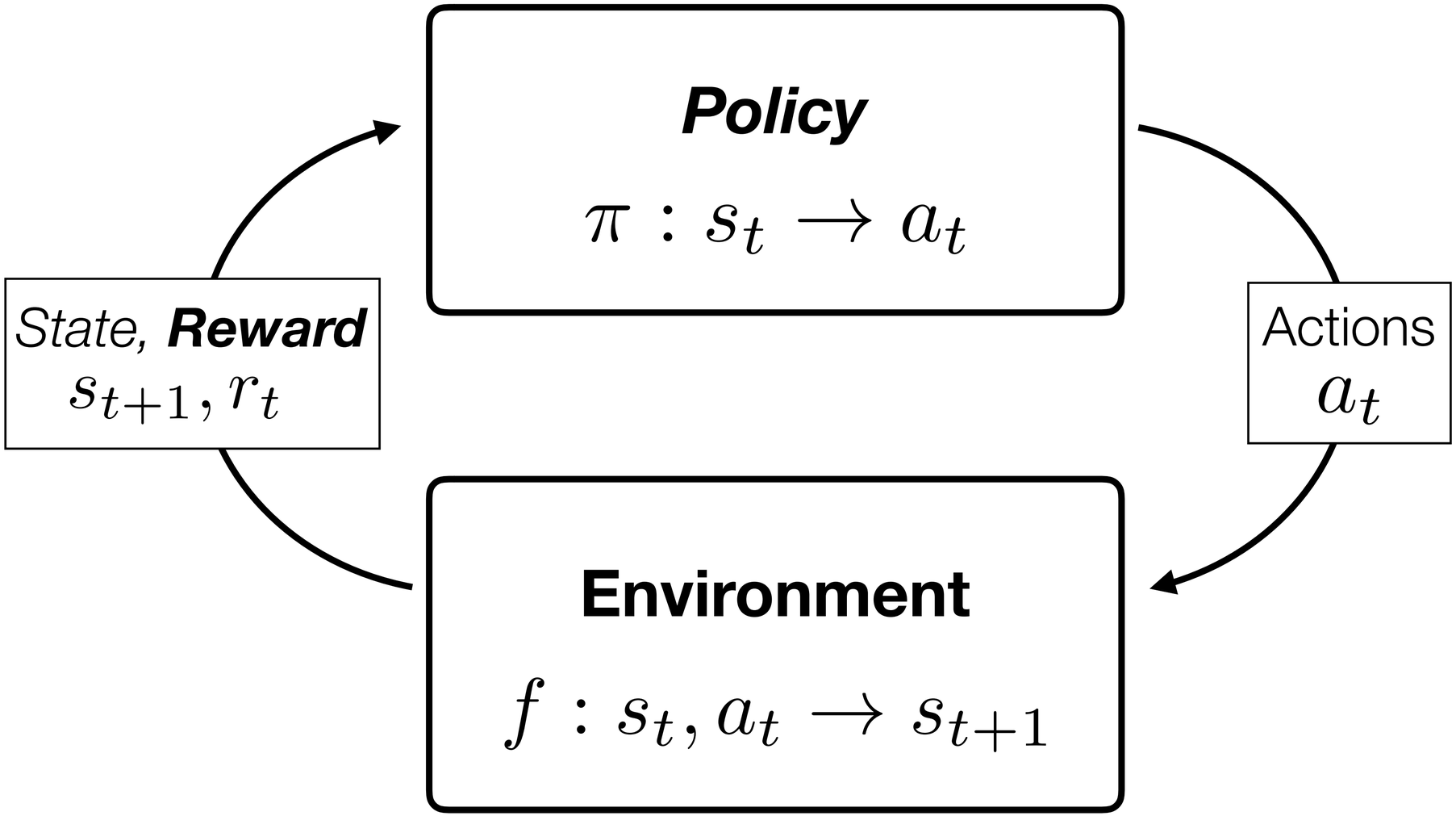
Lecture 26: More embodied learning + Ethics in computer vision (part 1)

Today

- Embodied learning
  - Q-learning
  - Playing games
  - Model-based reinforcement learning
- Ethics in computer vision (part 1)
  - Detecting fake images

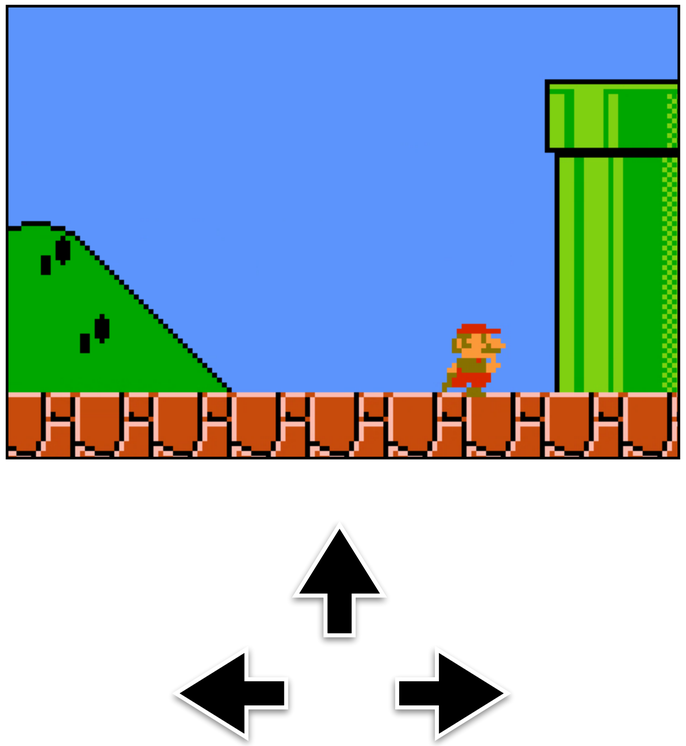
Recall: Reinforcement learning



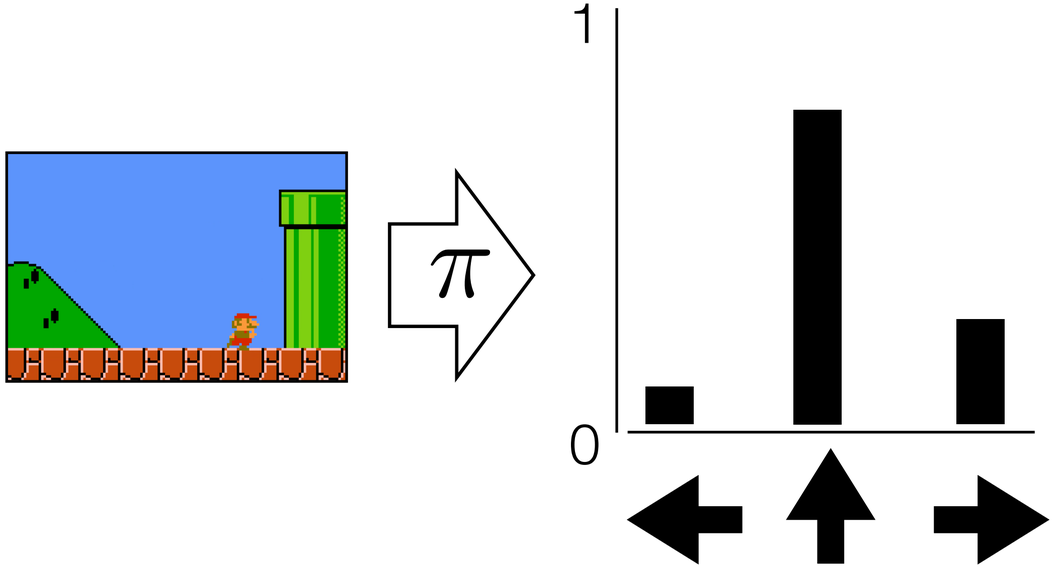
Source: Isola, Torralba, Freeman

Recall: Reinforcement learning

state: pixels!

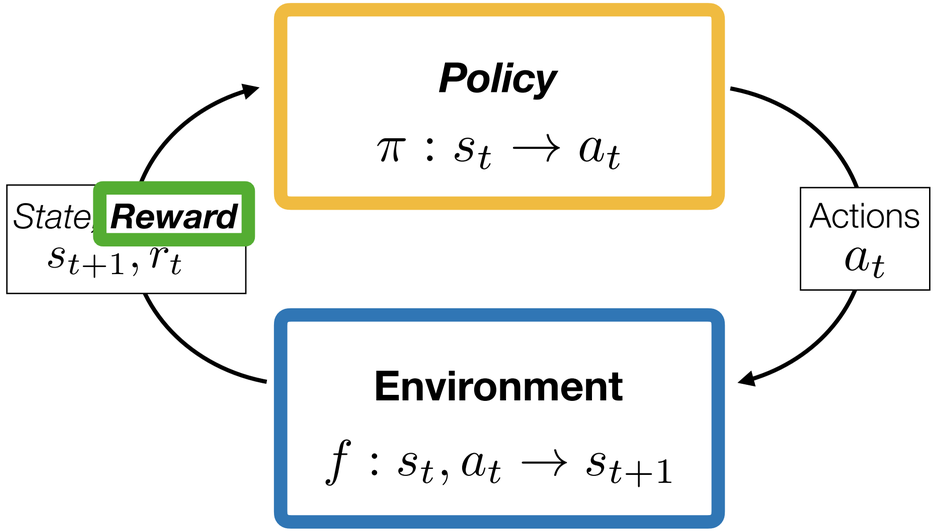


policy: action classifier



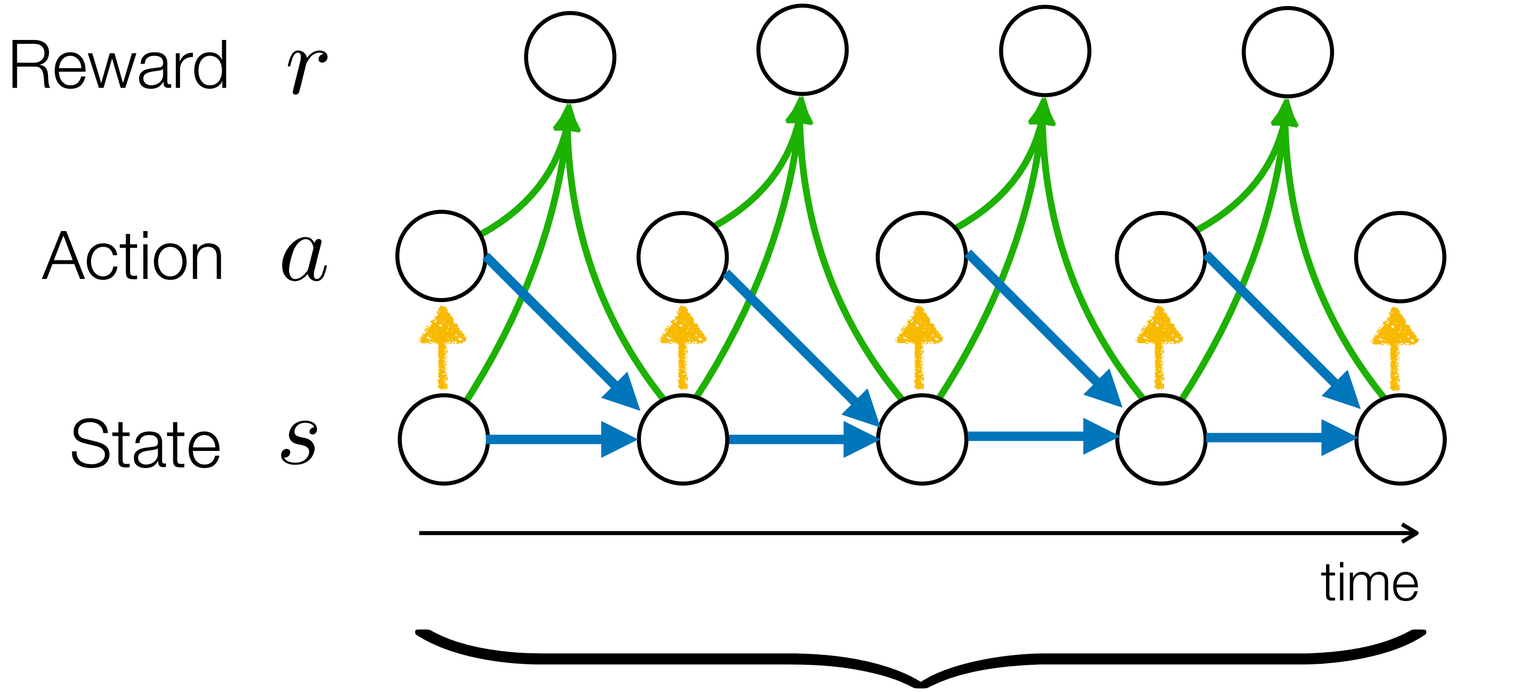
Source: Isola, Torralba, Freeman

Reinforcement learning



Markov decision process (MDP)

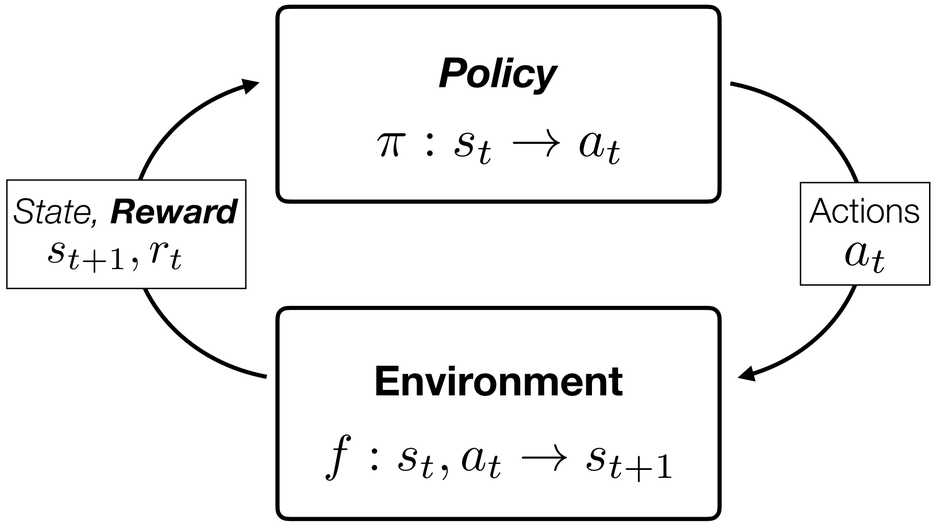
Learned



A sample from the MPD is called a **Trajectory** $\tau = (s_0, a_0, r_0, s_1, a_1, r_1, \ldots)$

Source: Isola, Torralba, Freeman

Recall: Reinforcement learning



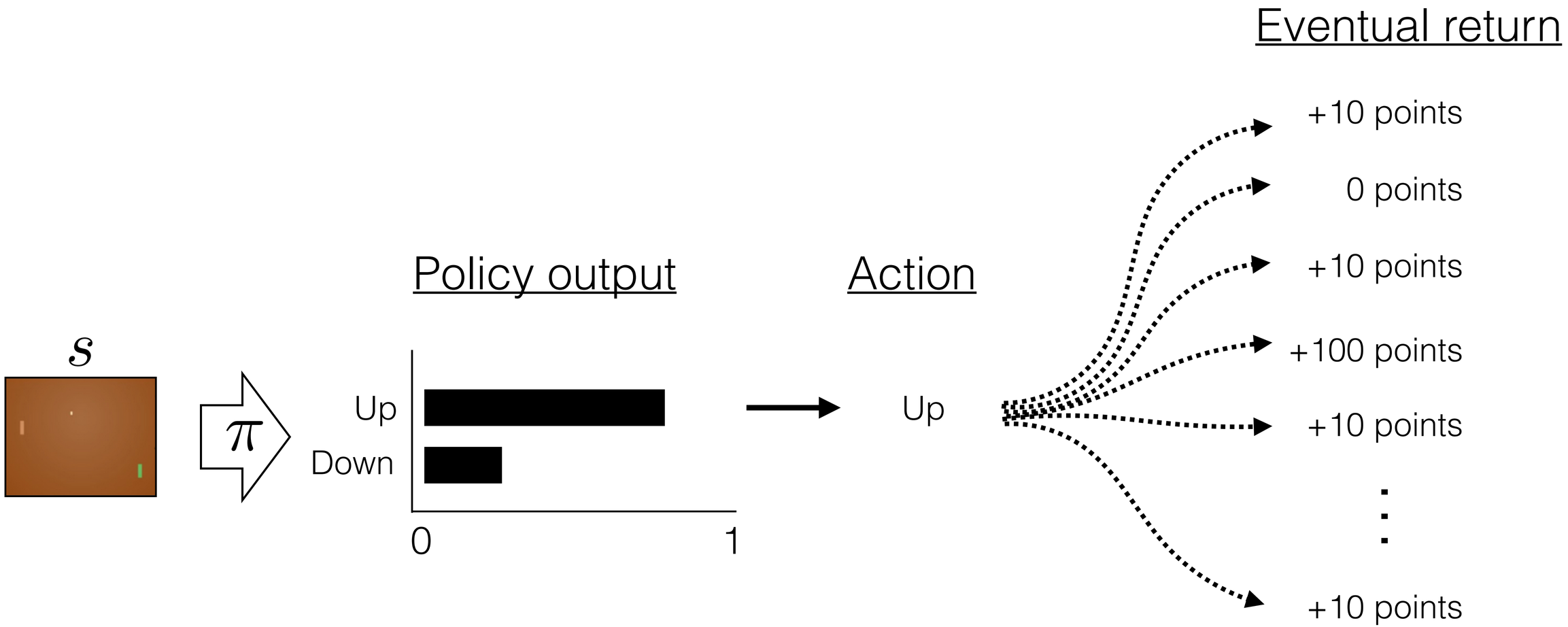
Trajectory $\tau = (s_0, a_0, r_0, s_1, a_1, r_1, \ldots)$

Discounted Returns $R(\tau) = \sum_{t=0}^{\infty} \gamma^t r_t, \quad \gamma \in (0, 1)$

Learn a policy that takes actions that maximize expected reward

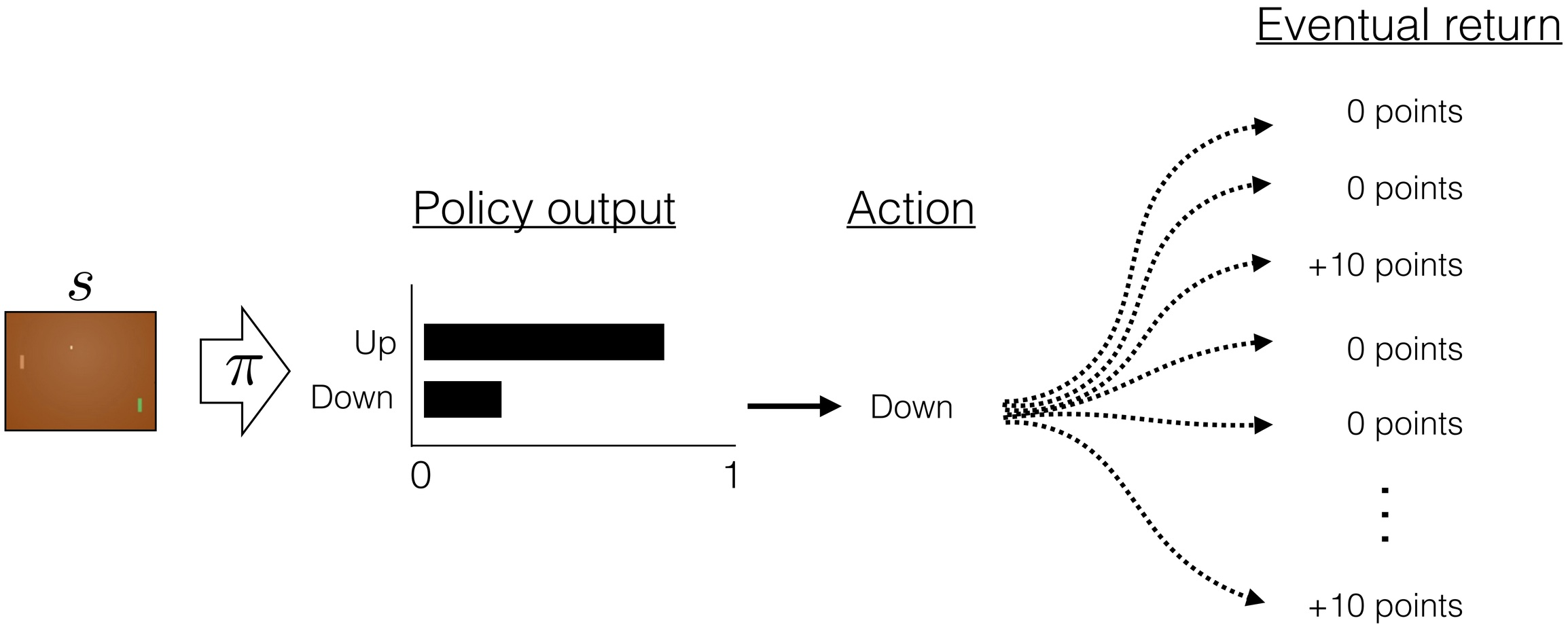
$$\pi^* = \arg\max_{\pi} \mathbb{E}_{\tau \sim \pi}[R(\tau)]$$

Source: Isola, Torralba, Freeman



$\pi(a|s)$ = probability of choosing action *a* given state *s*

Source: Isola, Torralba, Freeman

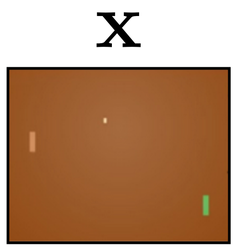


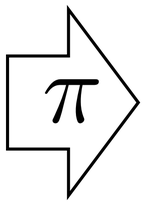
$\pi(a|s)$ = probability of choosing action *a* given state *s*

Source: Isola, Torralba, Freeman

Recall: Policy gradient

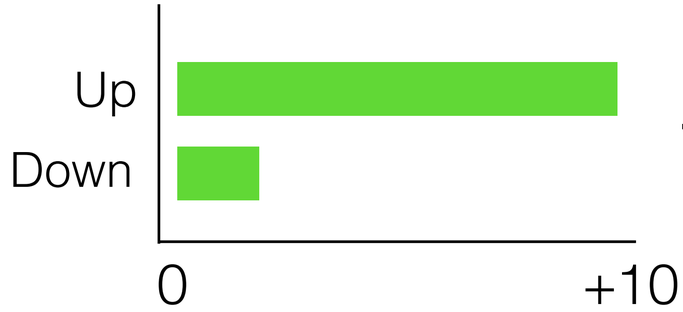
Approximated via sampling



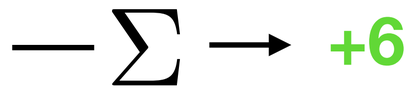


Policy output

Action conditional expected return



Expected return



$$\nabla_\theta \mathbb{E}_{\tau\sim\pi_\theta}[R(\tau)] = \mathbb{E}_{\tau\sim\pi_\theta}[R(\tau)\nabla_\theta \log \pi_\theta]$$

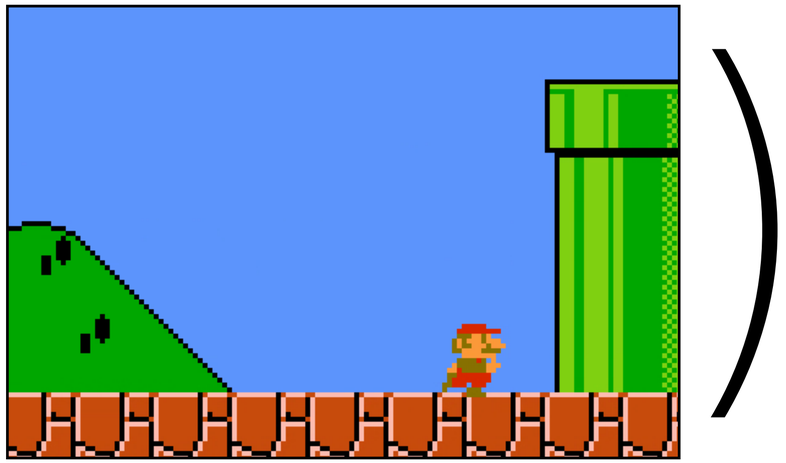
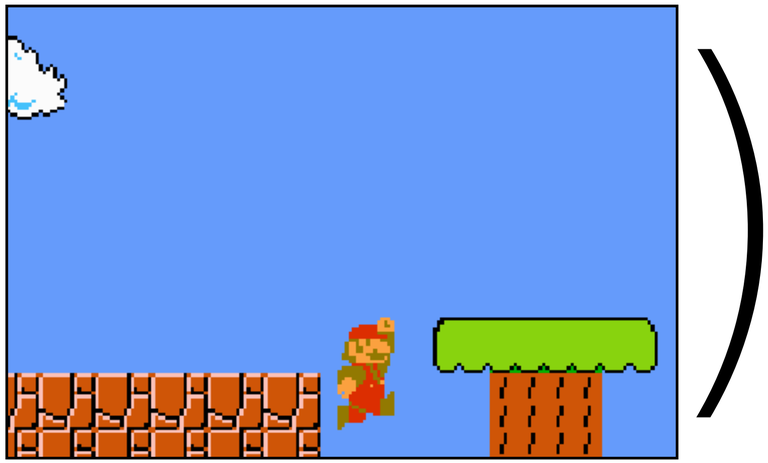
Estimate gradient using REINFORCE and do gradient descent on policy

Source: Isola, Torralba, Freeman

How good is a state?

Value function: expected future reward from starting in *s*.

$$V^{\pi}(s) = \mathbb{E}\left[\sum_{t \geq 0} \gamma^t r_t \mid s_0 = s, \pi\right]$$

V^{π}  vs. V^{π} 

- One advantage is *credit assignment*. We know which state/action was useful.
- Often more sample efficient, and updates have less variance.

How good is a state-action pair?

- Could we *learn* the value function and use it to choose actions?
  - Doesn't quite work. You'd also need to know the dynamics, i.e. what state you'd end up with if you took each action.

- Instead, learn **action-value function** (or **Q function**).

$$Q(s, a) = \mathbb{E}\left[\sum_{t \geq 0} R_t \mid s_0 = s, a_t = a\right]$$

- Optimal action for a state: $\operatorname*{argmax}_a Q(s, a)$

Adapted from R. Grosse and J. Ba

Finding a good Q function

- Good Q function should satisfy a recurrence relation called the Optimal Bellman Equation:

Quality of state/action pair

Where will I end up?

What if I take the *very best* next action?

$$Q^*(s,a) = r(s,a) + \gamma \mathbb{E}_{p(s'|s,a)} \left[\max_{a'} Q^*(s_{t+1}, a') \mid s_t = s, a_t = a \right]$$

Finding a good Q function

- Measuring the Bellman error for Q:

$$r(s_t, a_t) + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t)$$

- Approximate Q with a neural net Q(s, a; θ). For each episode *i*:

1. Do the policy induced by Q and get a trajectory:

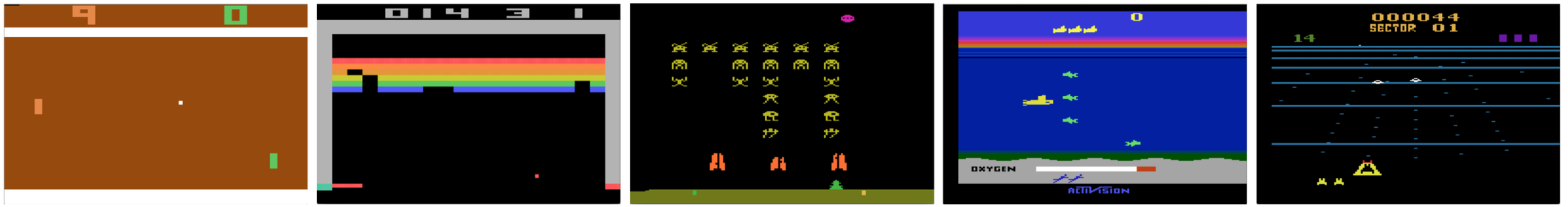
$$\tau = (s_0, a_0, r_0, s_1, a_1, r_1, \ldots)$$

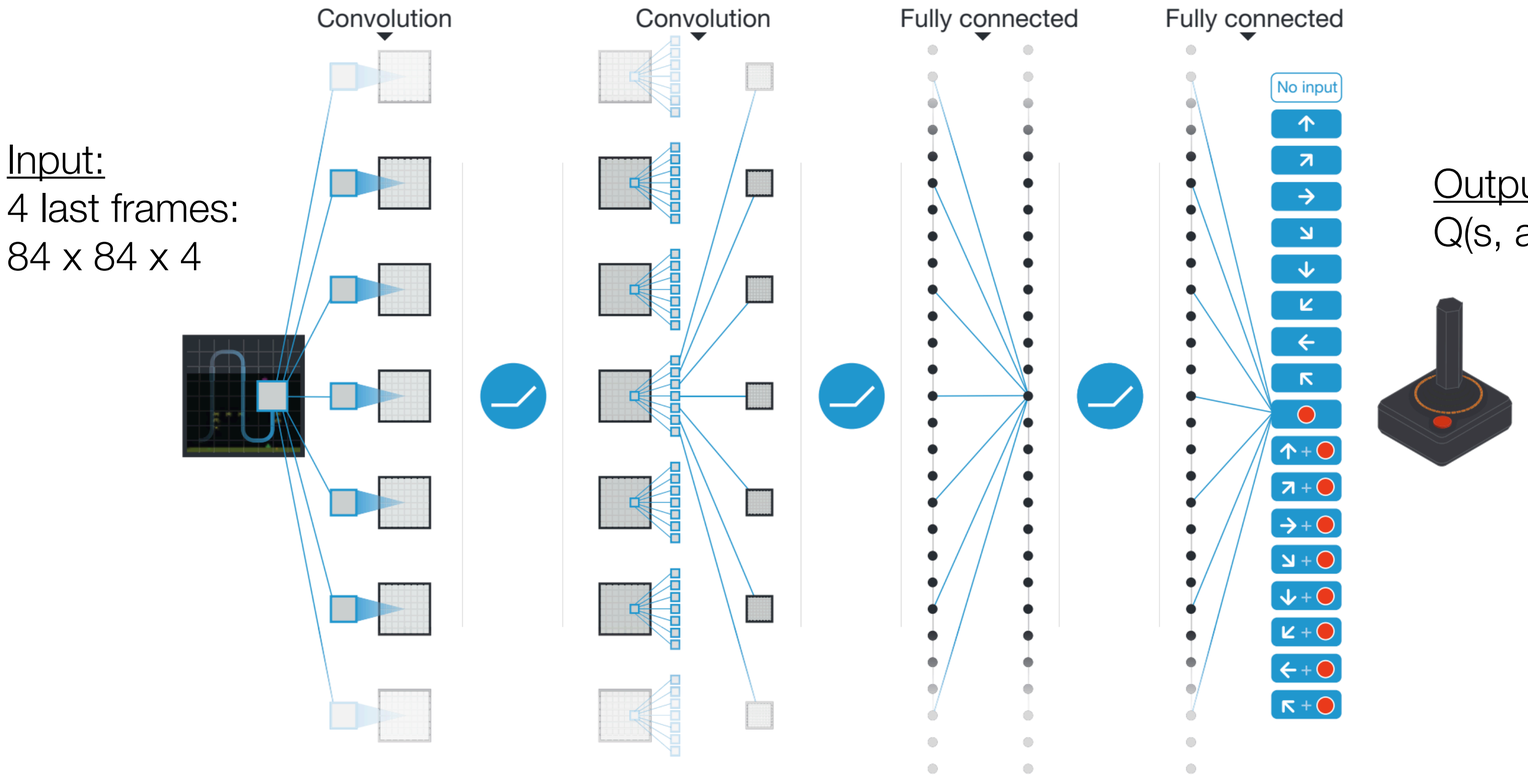
2. Update the parameters using backprop, minimizing approximation error:

$$t_i = r(s_t, a_t) + \gamma \max_a Q(s_{t+1}, a; \theta_{i-1})$$

$$L(\theta_i) = (t_i - Q(s_t, a_t; \theta_i))^2$$

Playing Atari games



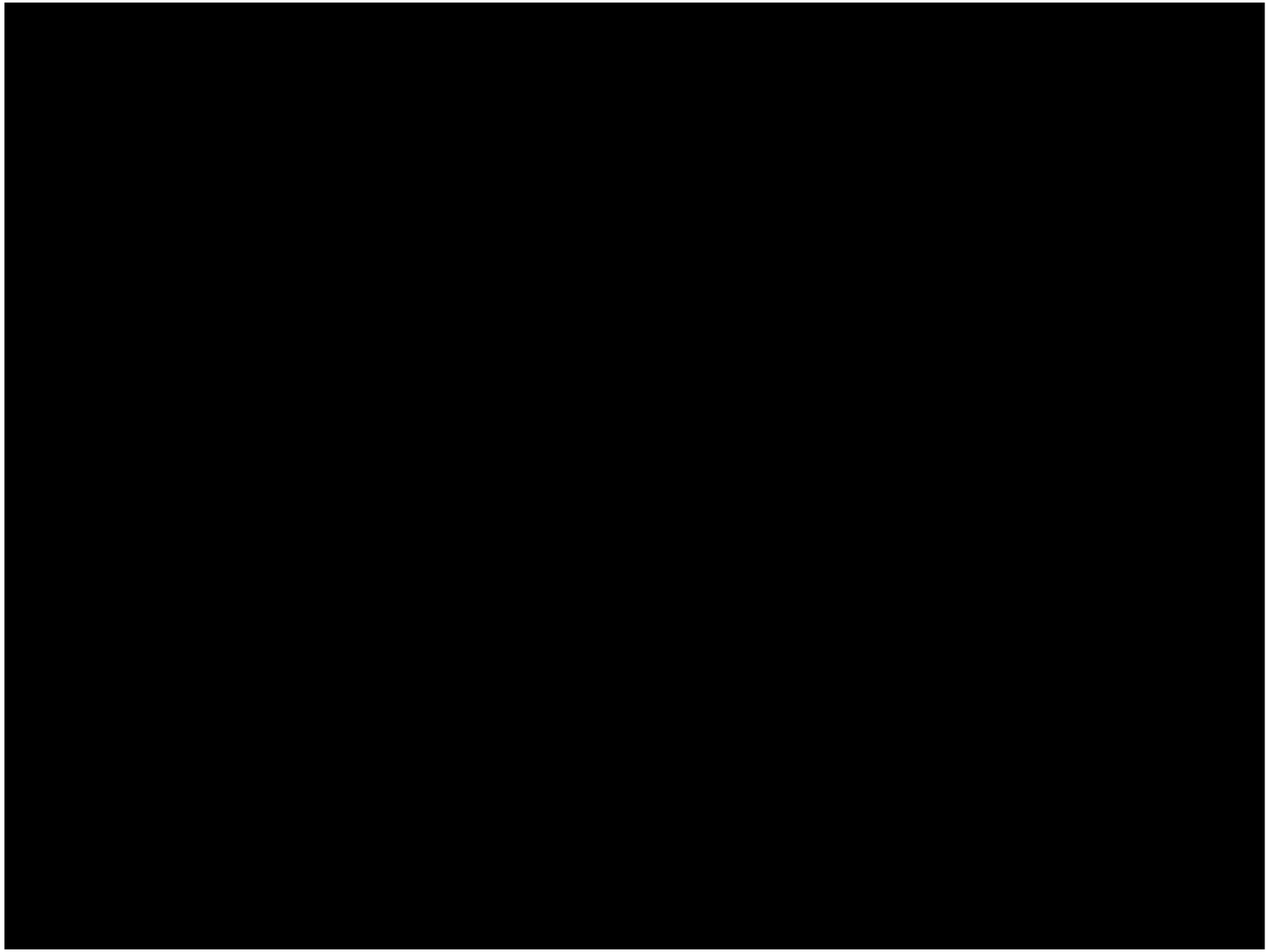


Input:
4 last frames:
84 x 84 x 4

Output:
Q(s, a) for each action *a*

5-layer CNN

Playing Atari games

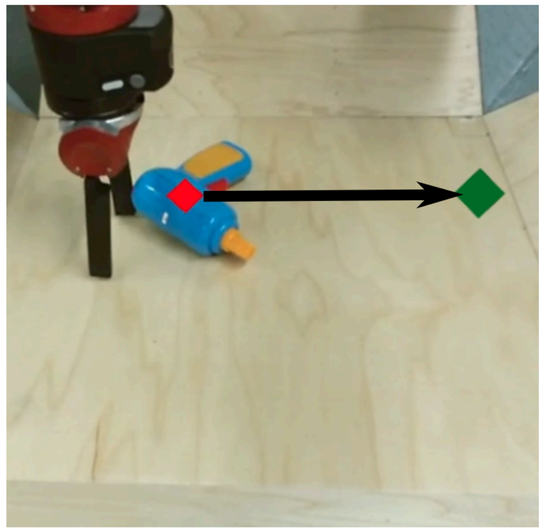


Model-based control

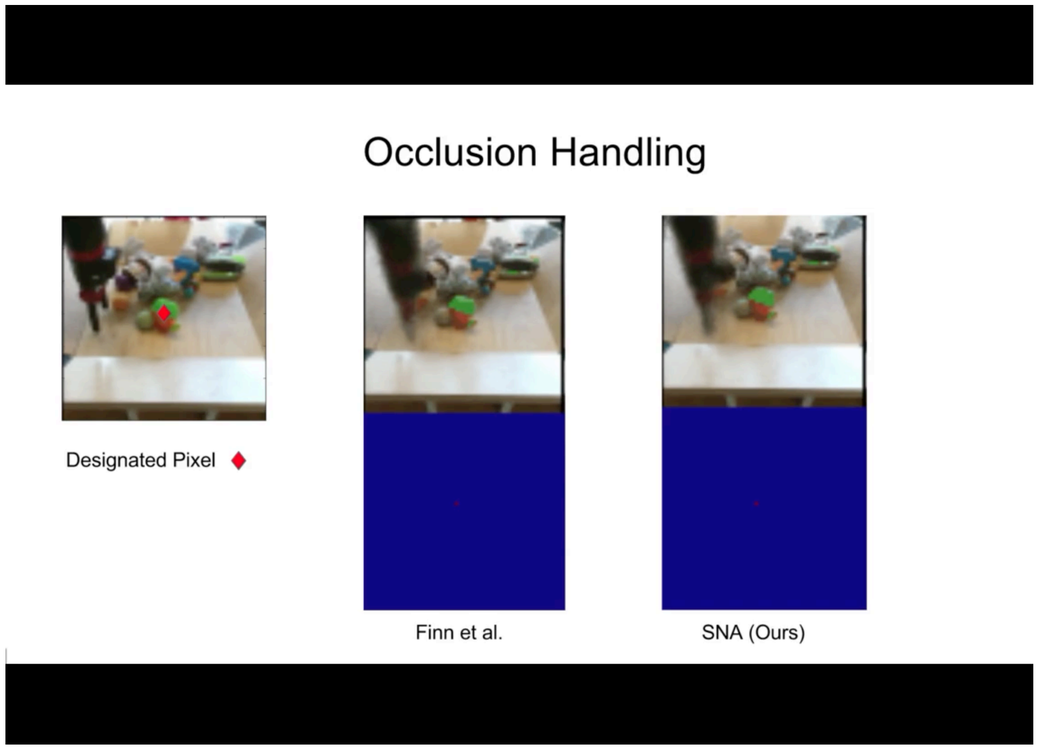
- Learn the **dynamics** of the environment: $p(s_{t+1} \mid s_t, a_t)$
- You can learn that through exploration, without an explicit reward.
- If states = images, we want to **predict the future** after you do an action

Model-based control

- Learn the **dynamics** of the environment: $p(s_{t+1} \mid s_t, a_t)$
- You can learn that through exploration, without an explicit reward.
- If states = images, we want to **predict the future** after you do an action



Pushing task



Video prediction

From [Ebert et al., "Self-Supervised Visual Planning with Temporal Skip Connections" 2017]

Image forensics

Image manipulation

From Forrest Gump, 1994

Malicious image manipulation

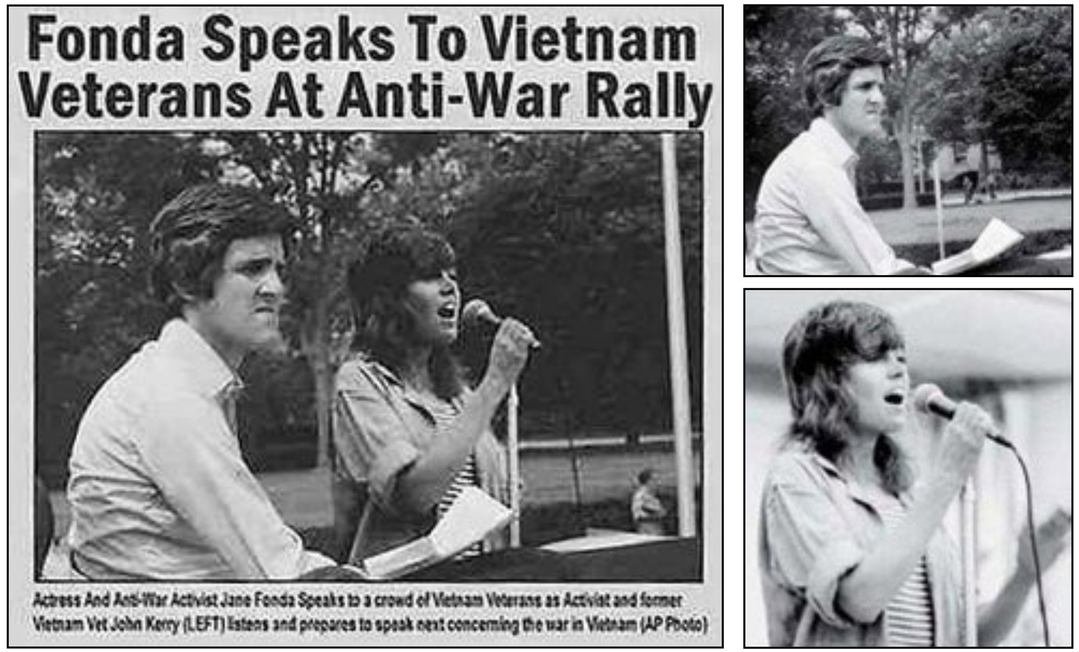
Malicious image manipulation

Malicious image manipulation

Fonda Speaks To Vietnam Veterans At Anti-War Rally

Actress And Anti-War Activist Jane Fonda Speaks to a crowd of Vietnam Veterans as Activist and former Vietnam Vet John Kerry (LEFT) listens and prepares to speak next concerning the war in Vietnam (AP Photo)

Malicious image manipulation



his associates simply found photos of athletes on the Internet and either used those photos or used software such as PhotoShop to insert the applicants' faces onto the bodies of legitimate athletes. For example, as set forth in greater detail below, CW-1 explained to McGLASHAN that he would create a falsified athletic profile for McGLASHAN's son, something he told McGLASHAN he had "already done … a million times," and which would involve him using "Photoshop and stuff" to deceive university admissions officers.

FBI affidavit on college admissions scandal

Malicious image manipulation

Malicious image manipulation

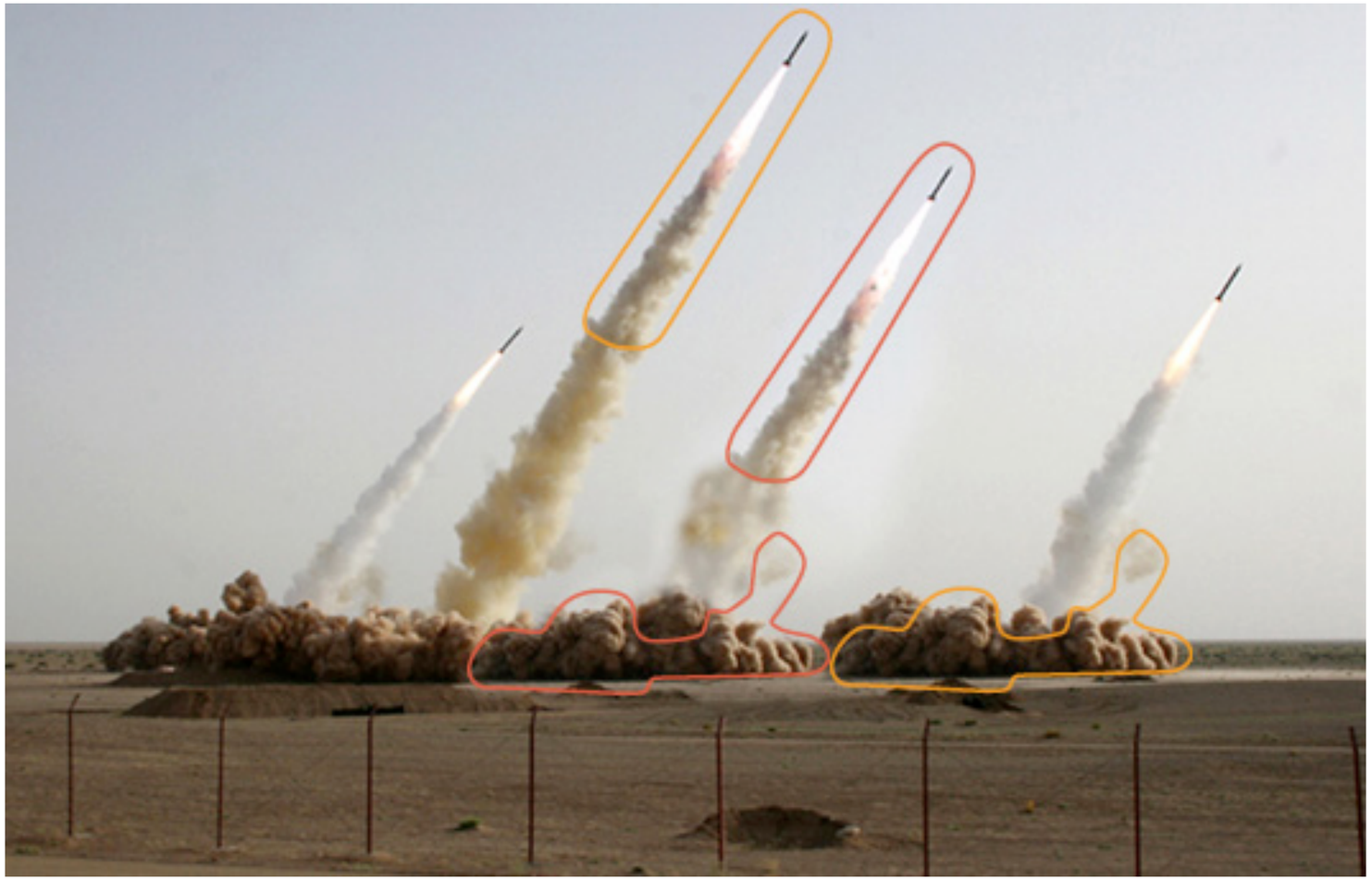
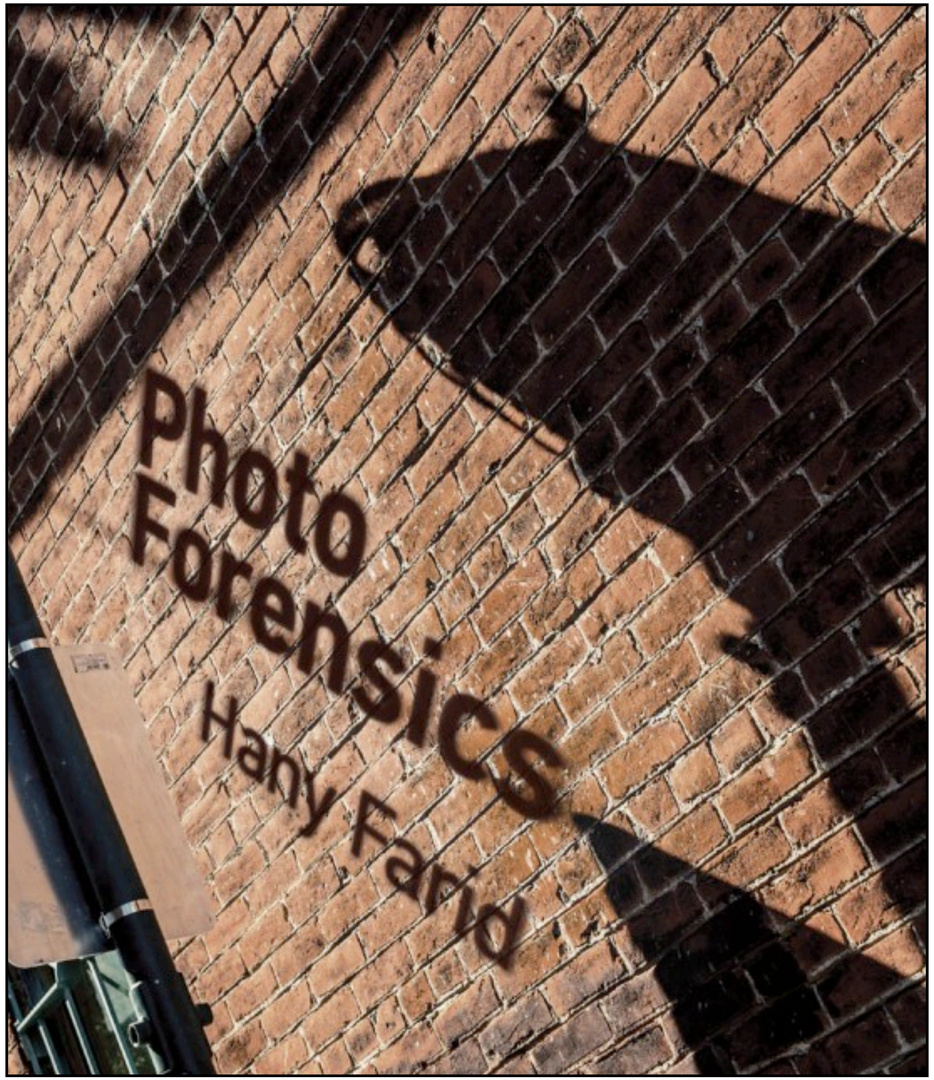
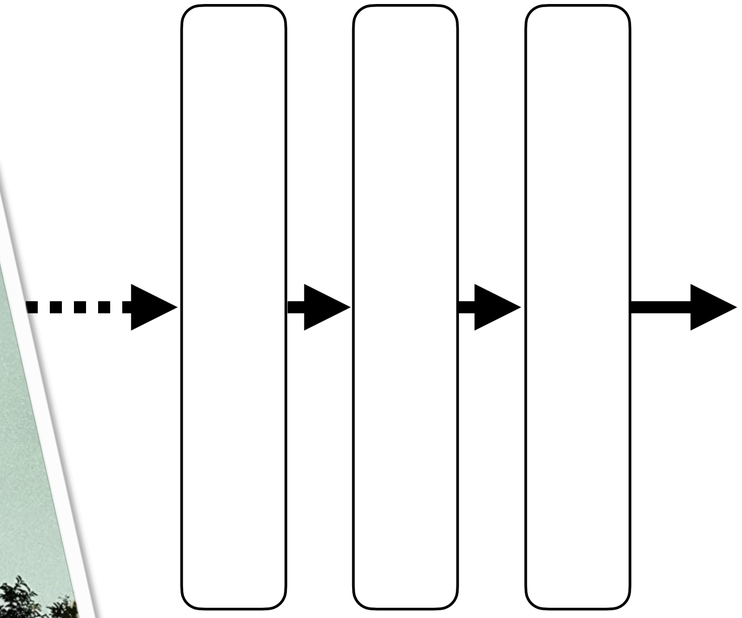


Image forensics: detecting fake images



Hany Farid

Hard to use supervised learning!



"Fake"

Strategy #1: use hard-to-fake physical cues

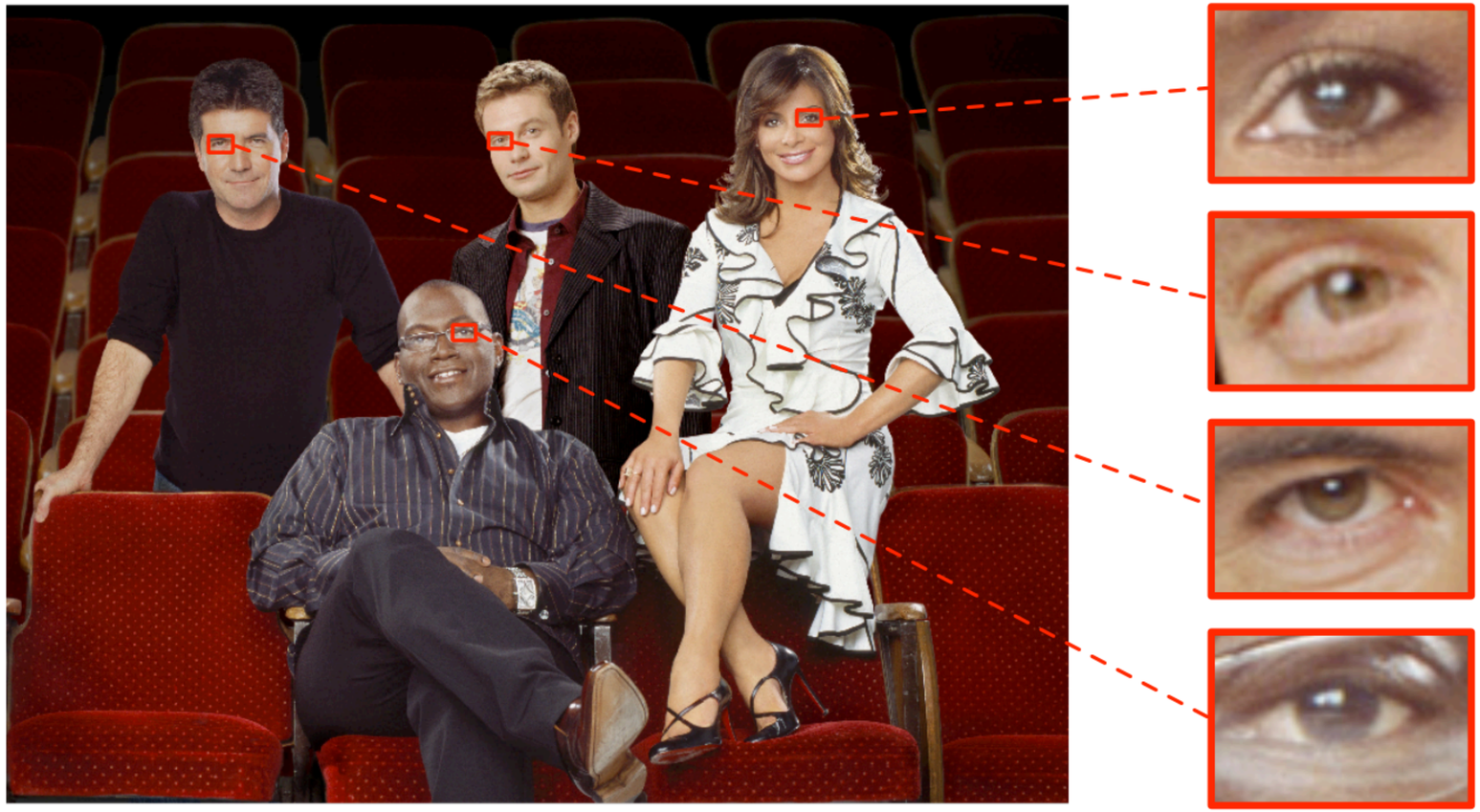
Fake photo

Real photo

[Johnson and Farid, 2005]

Source: S. Lazebnik

Strategy #1: use hard-to-fake physical cues



[Johnson and Farid, 2007]

Strategy #2: subtle signals in imaging pipeline

- Cameras compress images differently.
- During quantization, some do `round()`, others do `floor()` or `ceil()`
- Use knowledge of cameras to detect quantization type
- If a photo seems to have **both** kinds of quantization, it's probably a composite from different cameras!

[Agarwal and Farid, 2017]

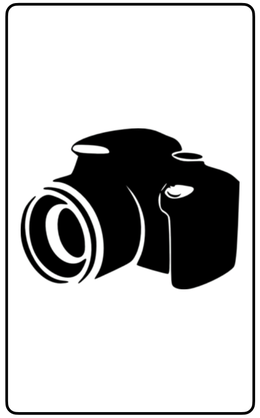
Strategy #3: learned anomaly detection

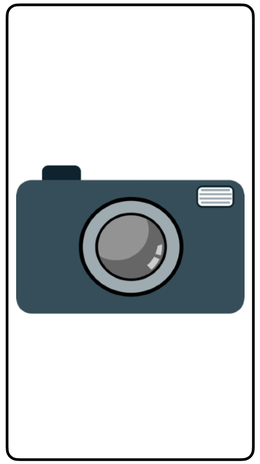
- Instead of hand-crafting cues, can we learn to detect “anomalous” images, and flag suspicious images?

[Huh, Liu et al., “Image Splice Detection via Learned Self-Consistency”, 2018]


EXCAVATING
DAEWOO


EXCAVATING
DAEWOO

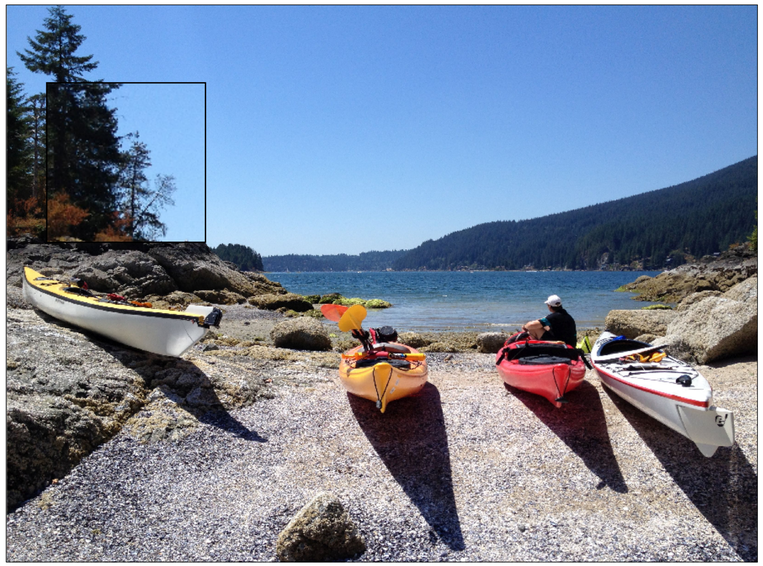




Inconsistent

Consistent

Predicting metadata consistency



```
CameraMake: Apple
CameraModel: iPhone 4s
ColorSpace: sRGB
ExifImageLength: 2448
ExifImageWidth: 3264
Flash: Flash did not fire
FocalLength: 107/2
WhiteBalance: Auto
ExposureTime: 1/2208
                …
```

```
CameraMake: NIKON CORPORATION
CameraModel: NIKON D90
ColorSpace: sRGB
ExifImageLength: 2848
ExifImageWidth: 4288
Flash: Flash did not fire
FocalLength: 18/796
WhiteBalance: Auto
ExposureTime: 1/30
                …
```

Same camera model?

Different

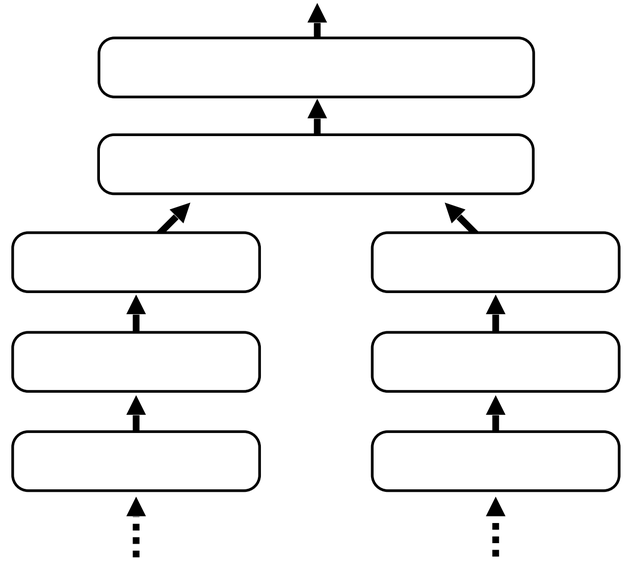
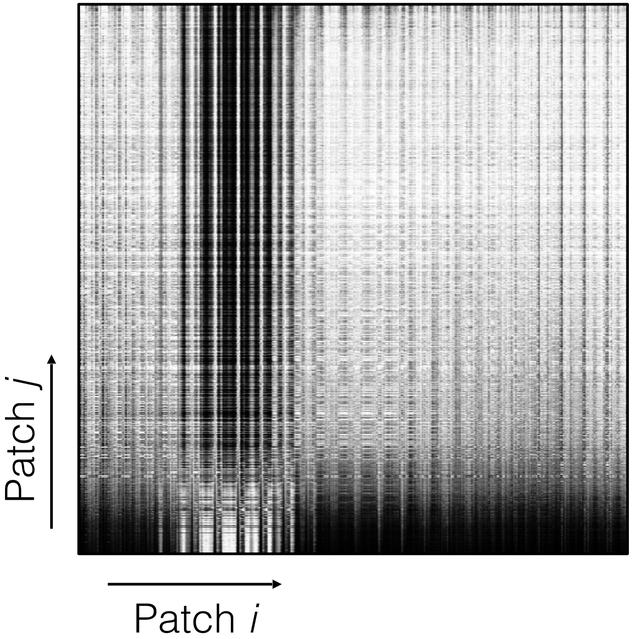


Photo source:
reddit.com/user/jjrosado

Photo source:
reddit.com/user/jjrosado

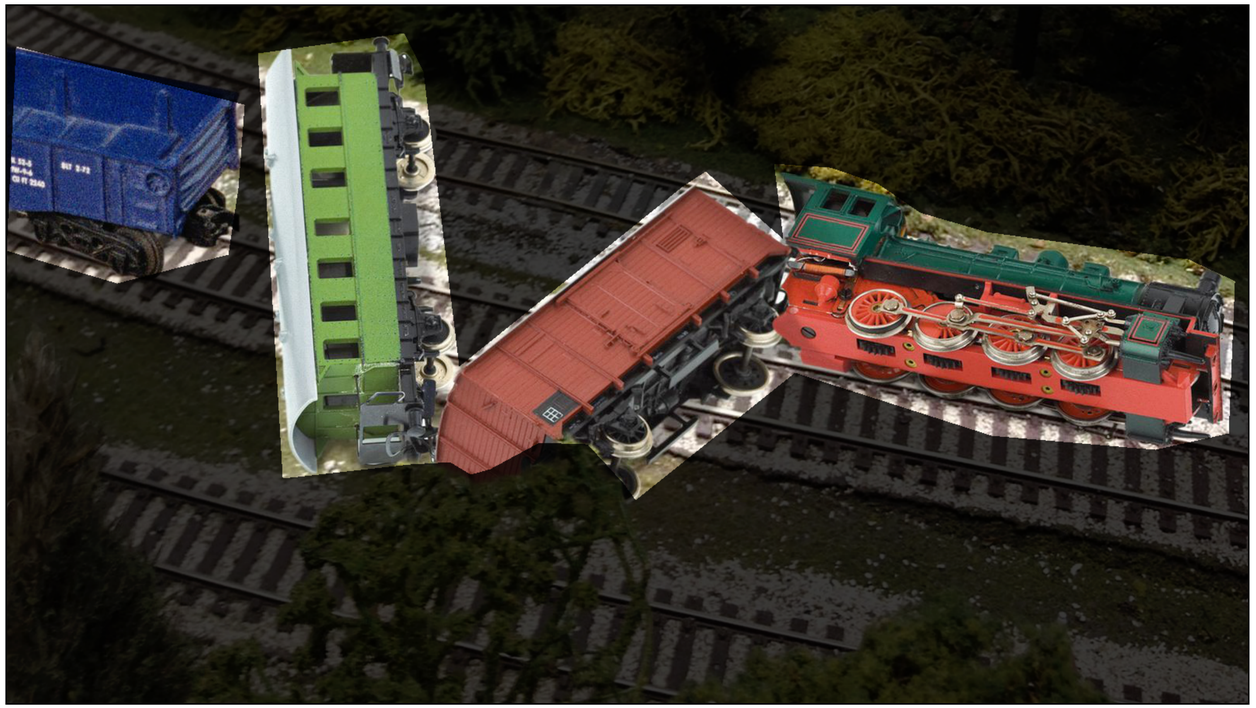
Affinity matrix



Input

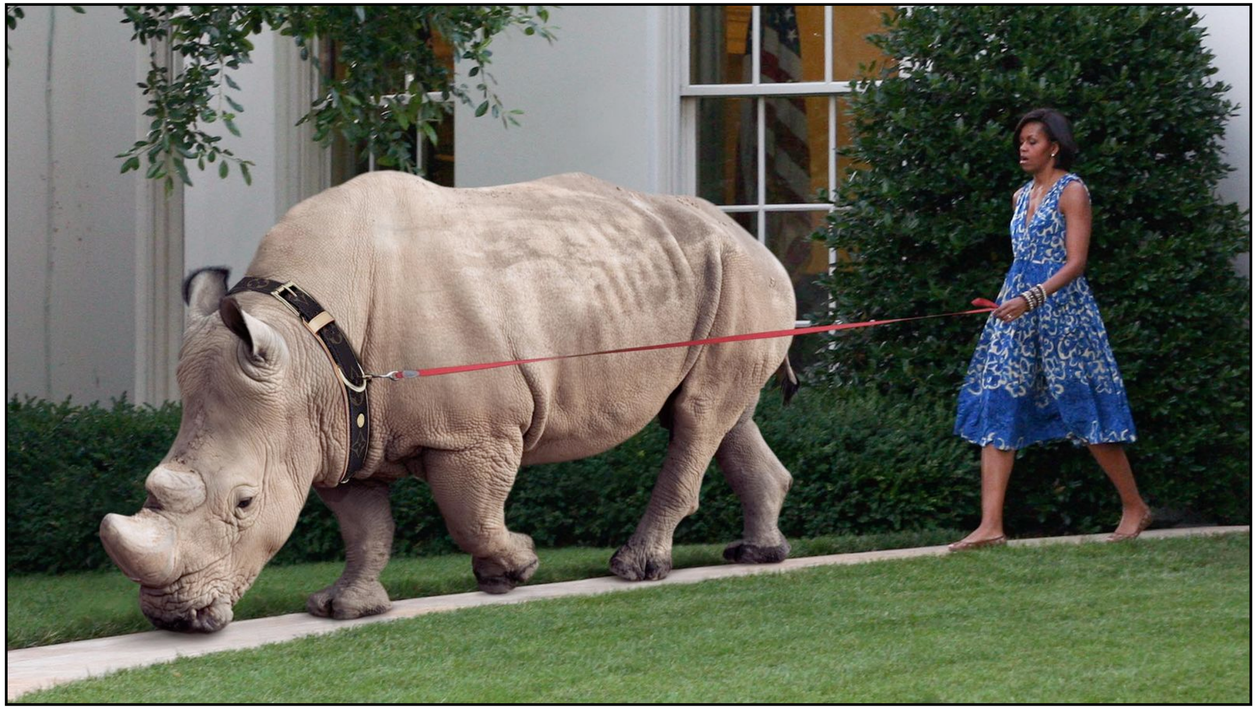
Photo source: TheOnion.com

Prediction



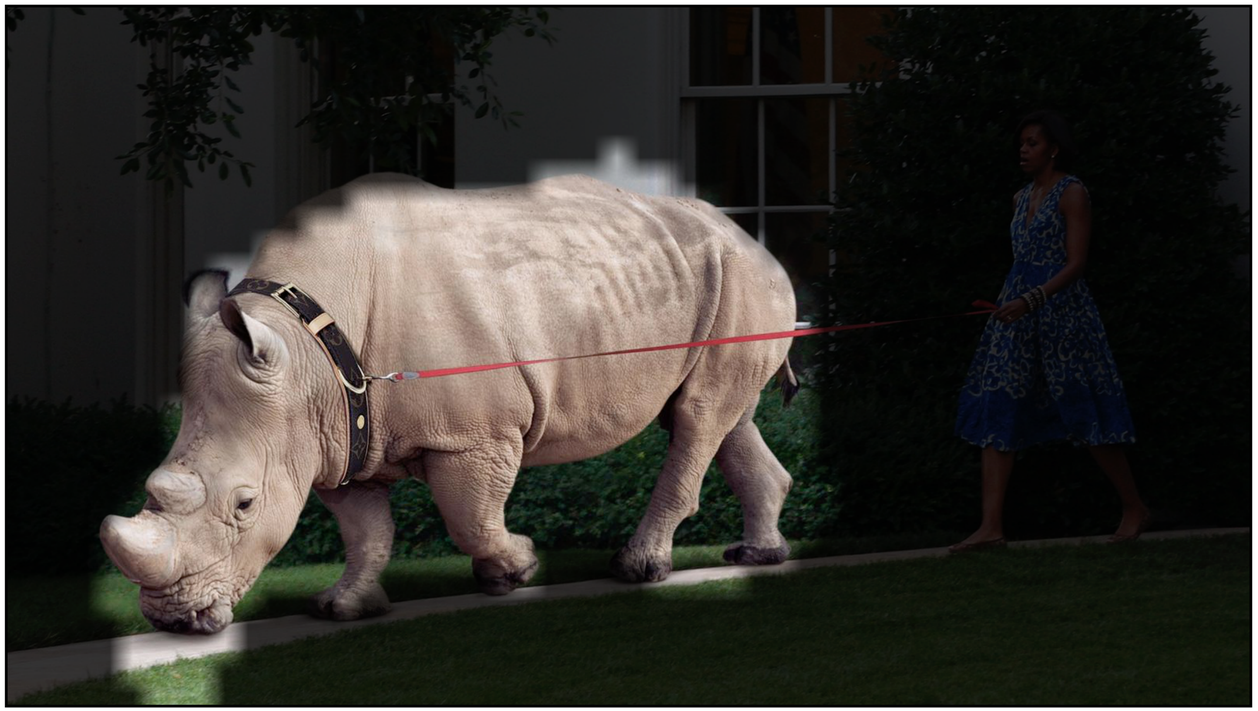
Ground truth

Photo source: TheOnion.com

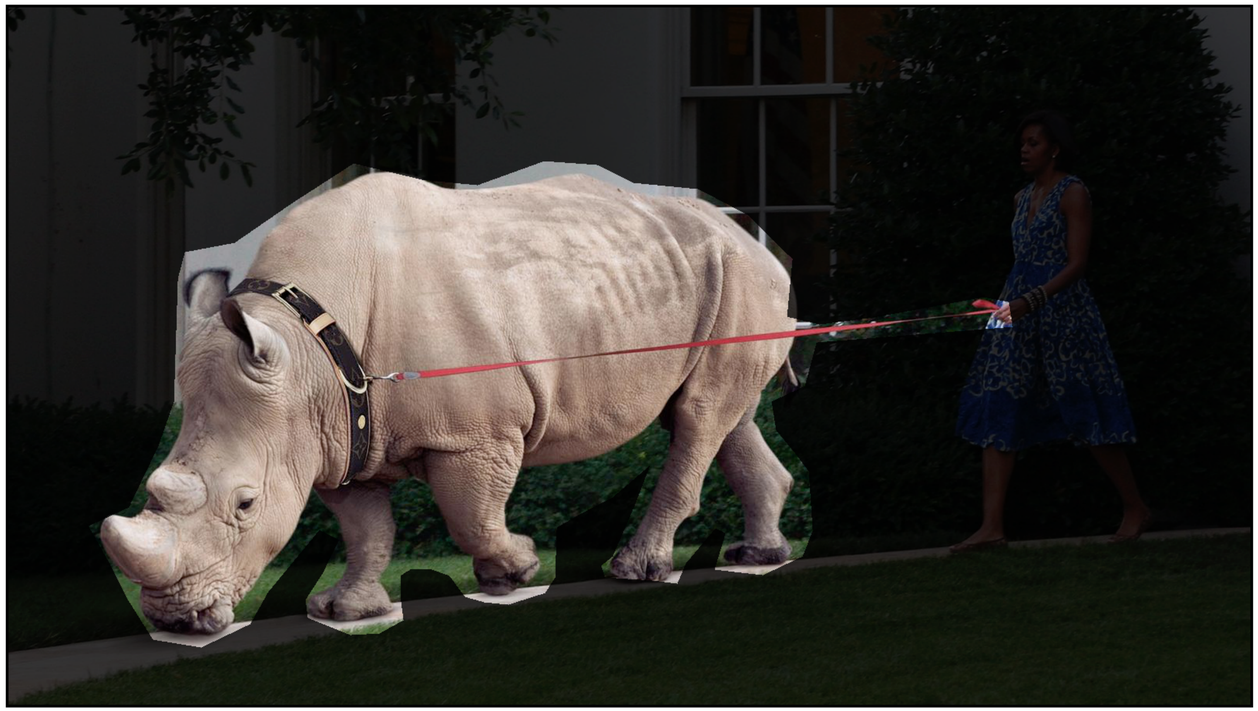


Input

Photo source: TheOnion.com



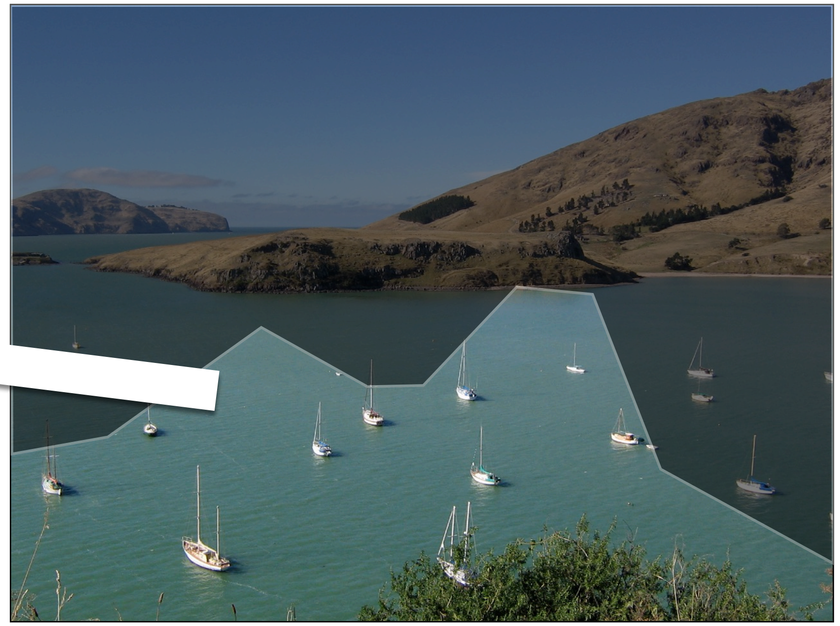
Prediction



Ground truth

Photo source: TheOnion.com

Input



(Hays & Efros 2009)

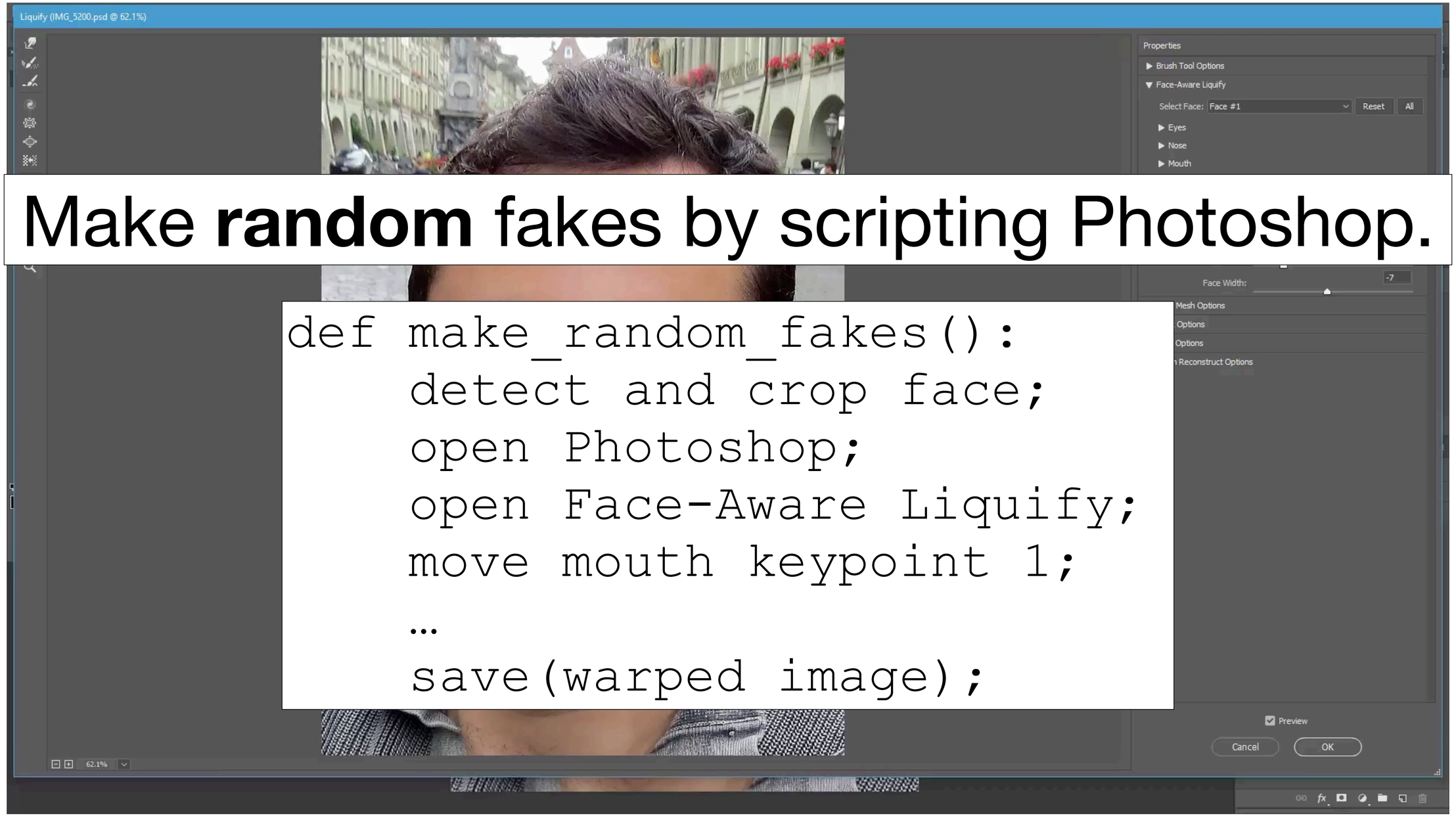
Prediction

Ground truth

Strategy #4: supervised learning

- Malicious image editors often use the same tools, e.g. Photoshop
- Can we “overfit” to these tools and detect them well?

[Wang et al., “Image Splice Detection via Learned Self-Consistency”, 2018]



Make **random** fakes by scripting Photoshop.

```
def make_random_fakes():
    detect and crop face;
    open Photoshop;
    open Face-Aware Liquify;
    move mouth keypoint 1;
    …
    save(warped image);
```

Photoshop Face-Aware Liquify tutorial. Source: https://youtu.be/5Qqv_C6iVvQ?t=86

Manipulated Photo

Warp Prediction

Suggested “Undo”

Original Photo

Suggested “Undo”

Manipulated Photo

New challenges on the horizon

“Deepfakes”

New challenges on the horizon

- Any internet troll can make a fake video!
- They don’t look *that* convincing yet, but of course quickly improving!
- Supervised learning methods can detect “known”deepfake algorithms
- But what about methods we’ve never seen before?
- Getting harder to tell what’s real vs. fake

What’s real and what’s fake?

[“The suspicious video that helped spark an attempted coup in Gabon” Washington Post. 2020]

https://www.youtube.com/watch?v=F5vzKs4z1dc

What’s real and what’s fake?

[“The suspicious video that helped spark an attempted coup in Gabon” Washington Post. 2020]

What’s real and what’s fake?

[“The suspicious video that helped spark an attempted coup in Gabon” Washington Post. 2020]

Next class: ethics in computer vision (part 2)