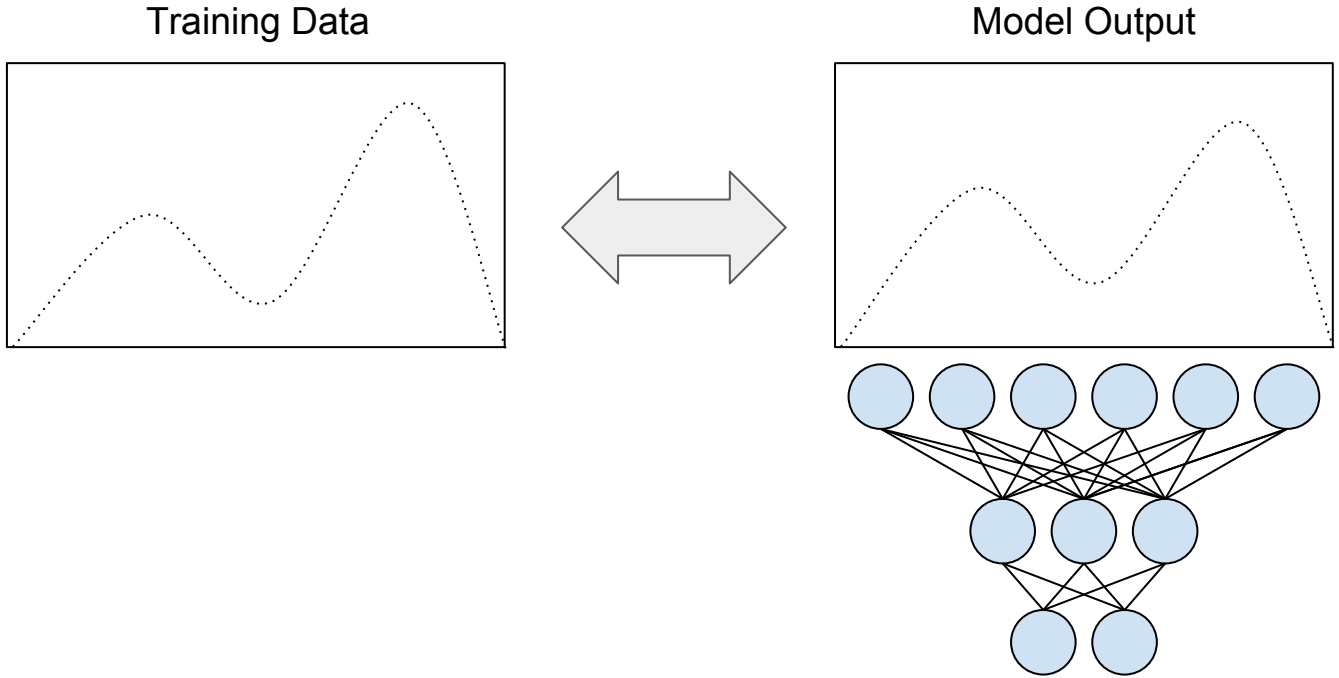
Our statement:

Image translation (via distribution matching) should not be used for direct interpretation.

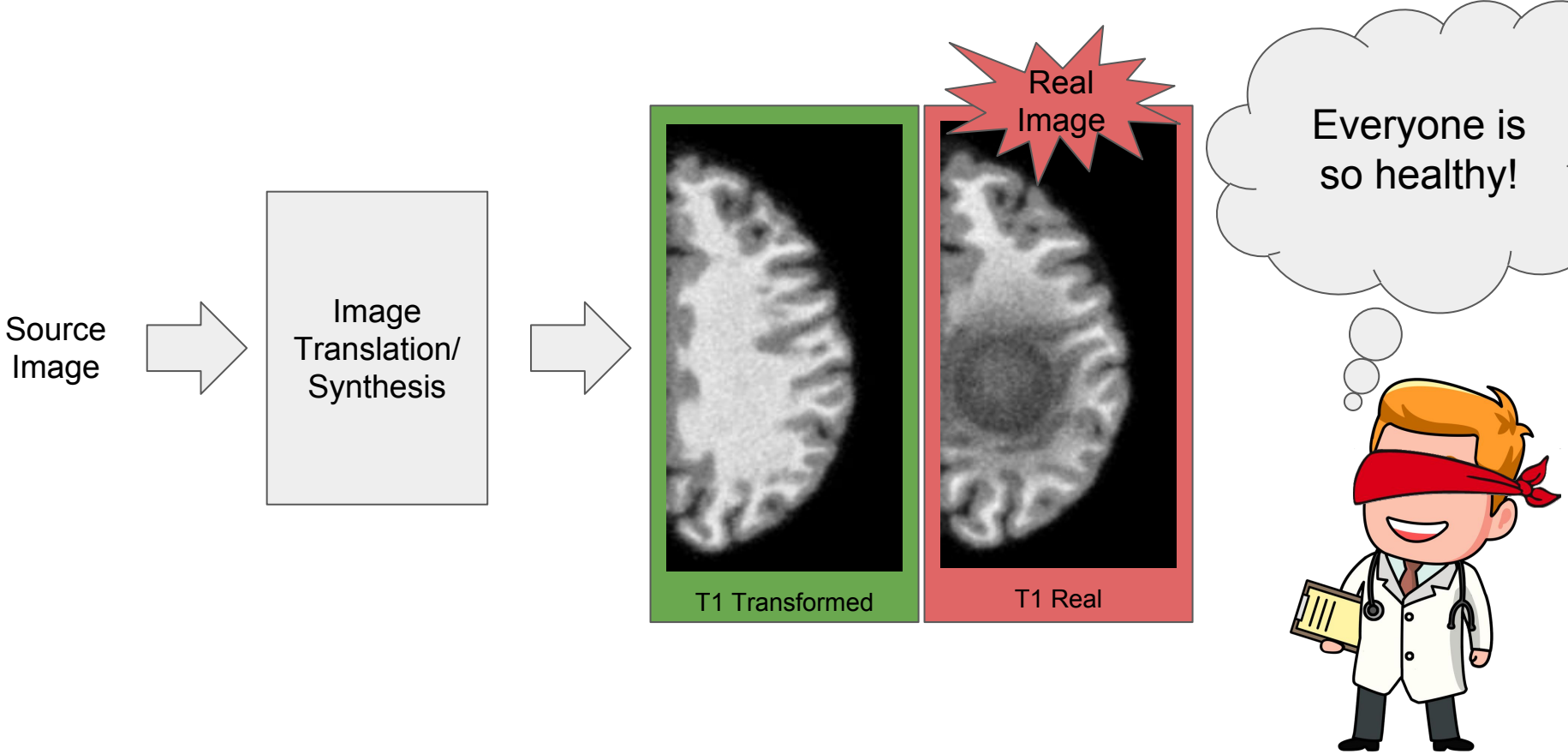


Losses like in CycleGAN just match distributions

[Karras, 2018]

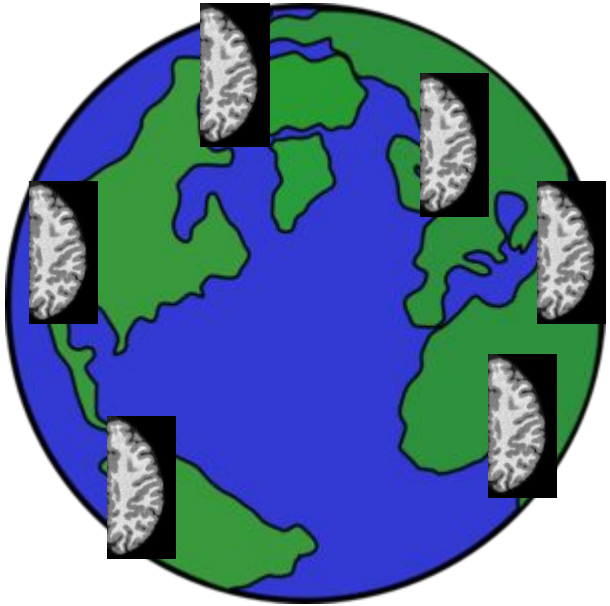
[CycleGAN, Zhu 2017]

They are very good at distribution matching



But a bias in training data can lead to incorrect translation

Time t



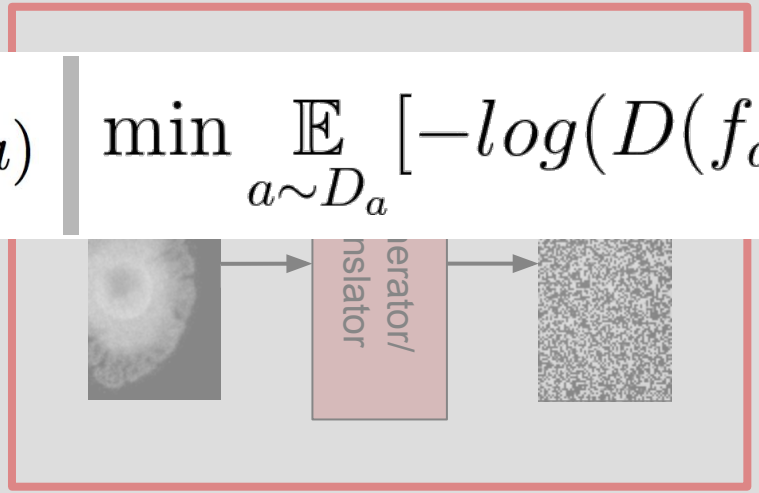
Even with all the training data in the world today.

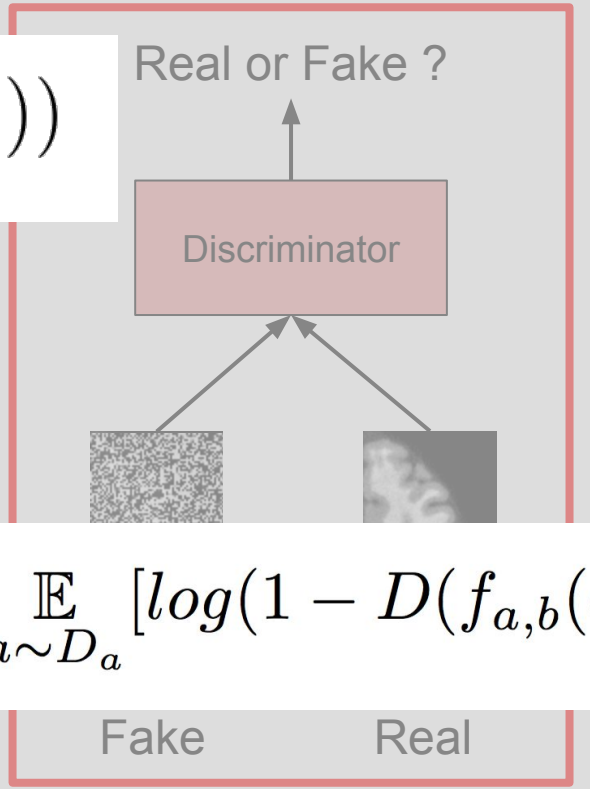
Time t+1

There will be new diseases tomorrow that are out of distribution.

What is image translation via distribution matching?

$$f_{a,b}(a) \quad \min_{a \sim D_a} \mathbb{E} \left[-log(D(f_{a,b}(a))) \right]$$





$$D \quad \max_{b \sim D_b} \mathbb{E} \left[logD(b) \right] + \mathop{\mathbb{E}}_{a \sim D_a} \left[log(1 - D(f_{a,b}(a))) \right]$$

discriminator

Model Breakdown

Optimizing

$$D \quad \max \underset{b \sim D_b}{\mathbb{E}} [logD(b)] + \underset{a \sim D_a}{\mathbb{E}} [log(1 - D(f_{a,b}(a)))]$$

Optimizing

$$f_{a,b}(a) \quad \min \underset{a \sim D_a}{\mathbb{E}} [-log(D(f_{a,b}(a)))$$

- $f_{a,b}(a)$ should produce examples in D_b
- D_a can be anything non-finite, like a Gaussian
- No guarantee mapping maintains phenotypes

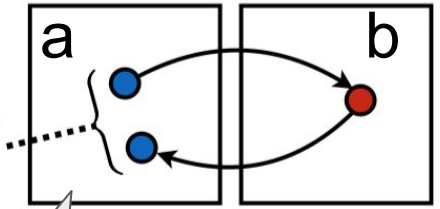
Optimizing

$$D \quad \max \underset{b \sim D_b}{\mathbb{E}} [logD(b)] + \underset{a \sim D_a}{\mathbb{E}} [log(1 - D(f_{a,b}(a)))]$$

Optimizing

$$f_{a,b}(a) \quad \min \underset{a \sim D_a}{\mathbb{E}} [-log(D(f_{a,b}(a))) + ||f_{b,a}(f_{a,b}(a)) - a||_1]$$

- Add a reconstruction loss regularizer for the func $f_{a,b}(a)$
- Loss term still matches distribution D_b
- No guarantee mapping maintains phenotypes



Cycle Loss!

Optimizing

$$D \quad \max \underset{(a,b)\sim(D_a,D_b)}{\mathbb{E}} [logD(b,a)] + \underset{a\sim D_a}{\mathbb{E}} [log(1 - D(f_{a,b}(a),a))]$$

Optimizing

$$f_{a,b}(a) \quad \min \underset{a\sim D_a}{\mathbb{E}} [-log(D(f_{a,b}(a)))$$

- D is given paired examples allowing detection of what to preserve
- D_b still plays a role in what D learns
- No guarantee mapping maintains phenotypes

Optimizing

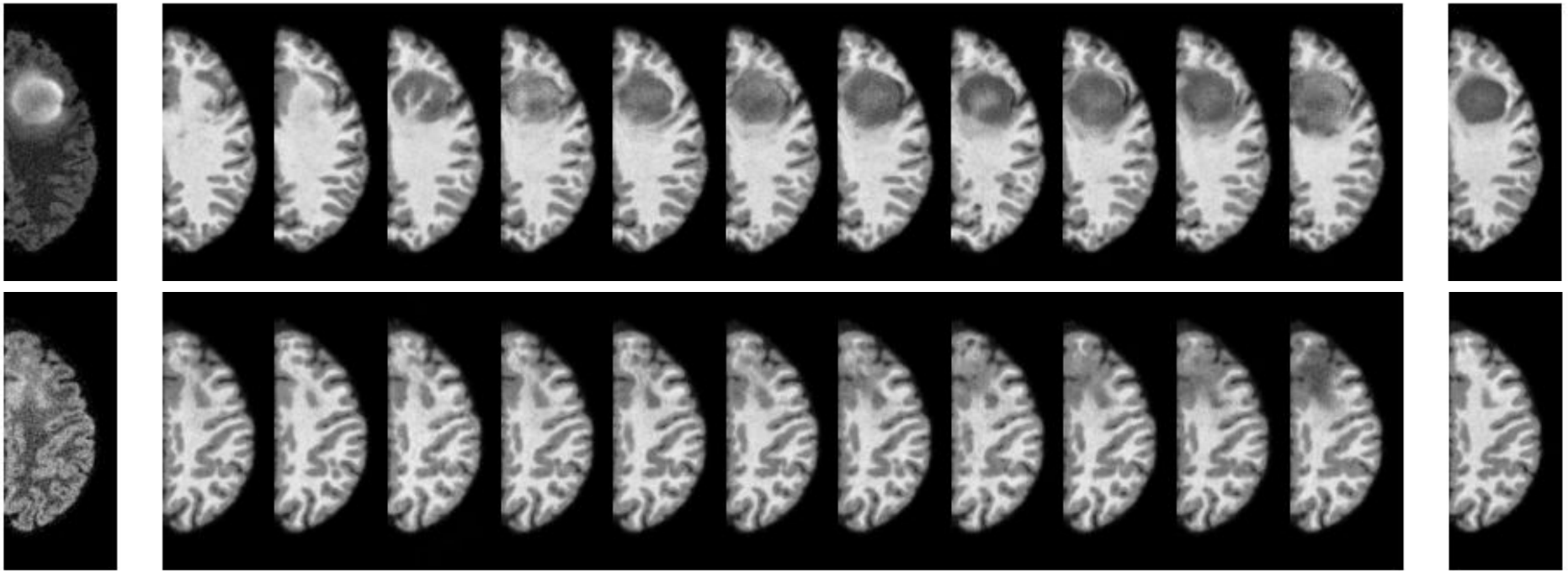
D ~~$\max \mathbb{E}_{(a,b)\sim(D_a,D_b)} \log(1 - D(f_{a,b}(a)))$~~

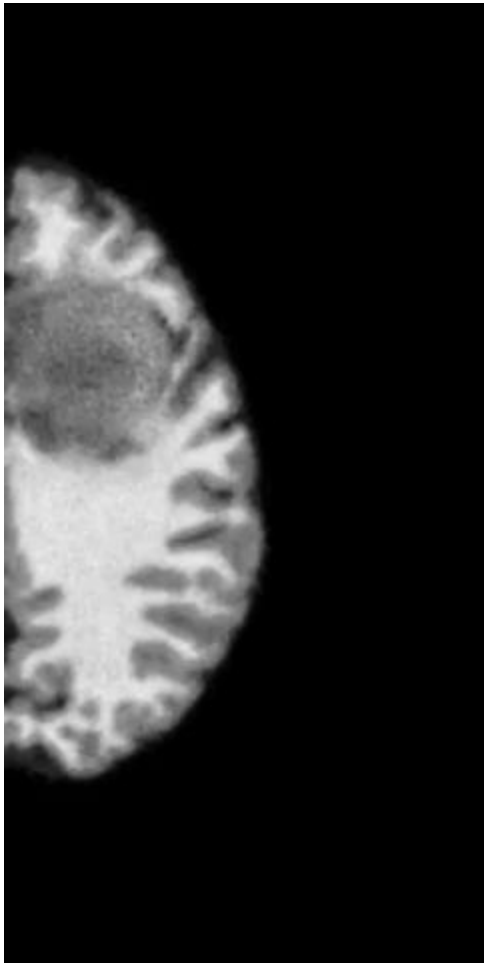
$f_{a,b}(a)$ $\min \mathbb{E}_{(a,b)\sim(D_a,D_b)} ||f_{a,b}(a) - b||_1$

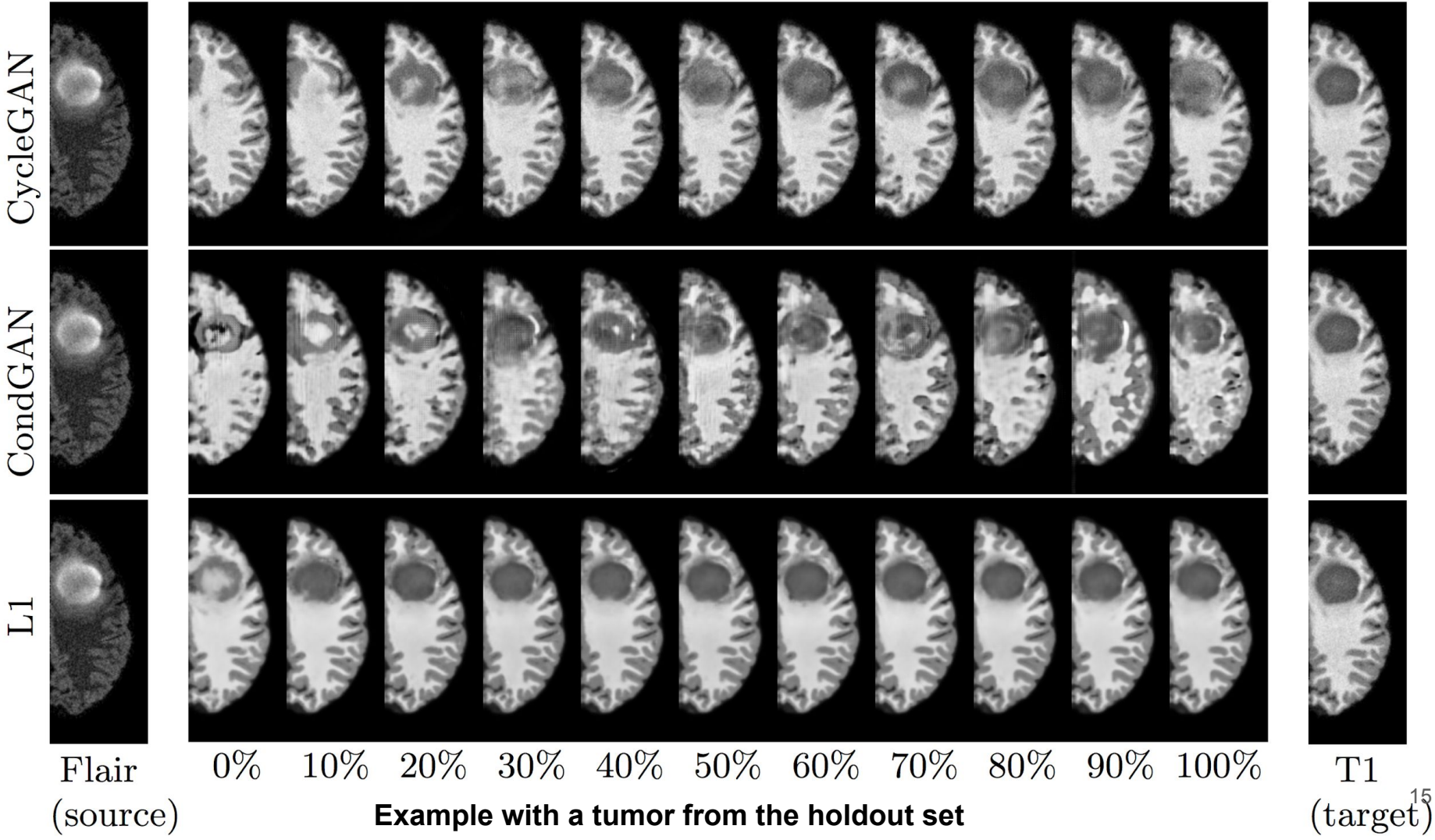
- $f_{a,b}(a)$ should produce examples in D_b
- Pixel-wise loss
- No guarantee mapping maintains phenotypes

Visual Evaluation

CycleGAN results

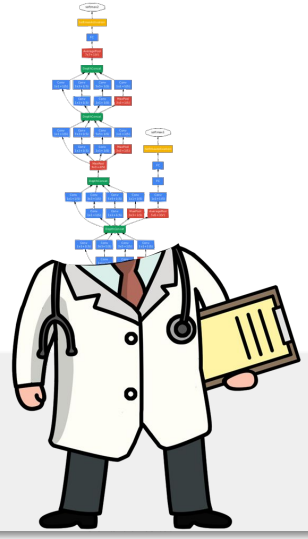


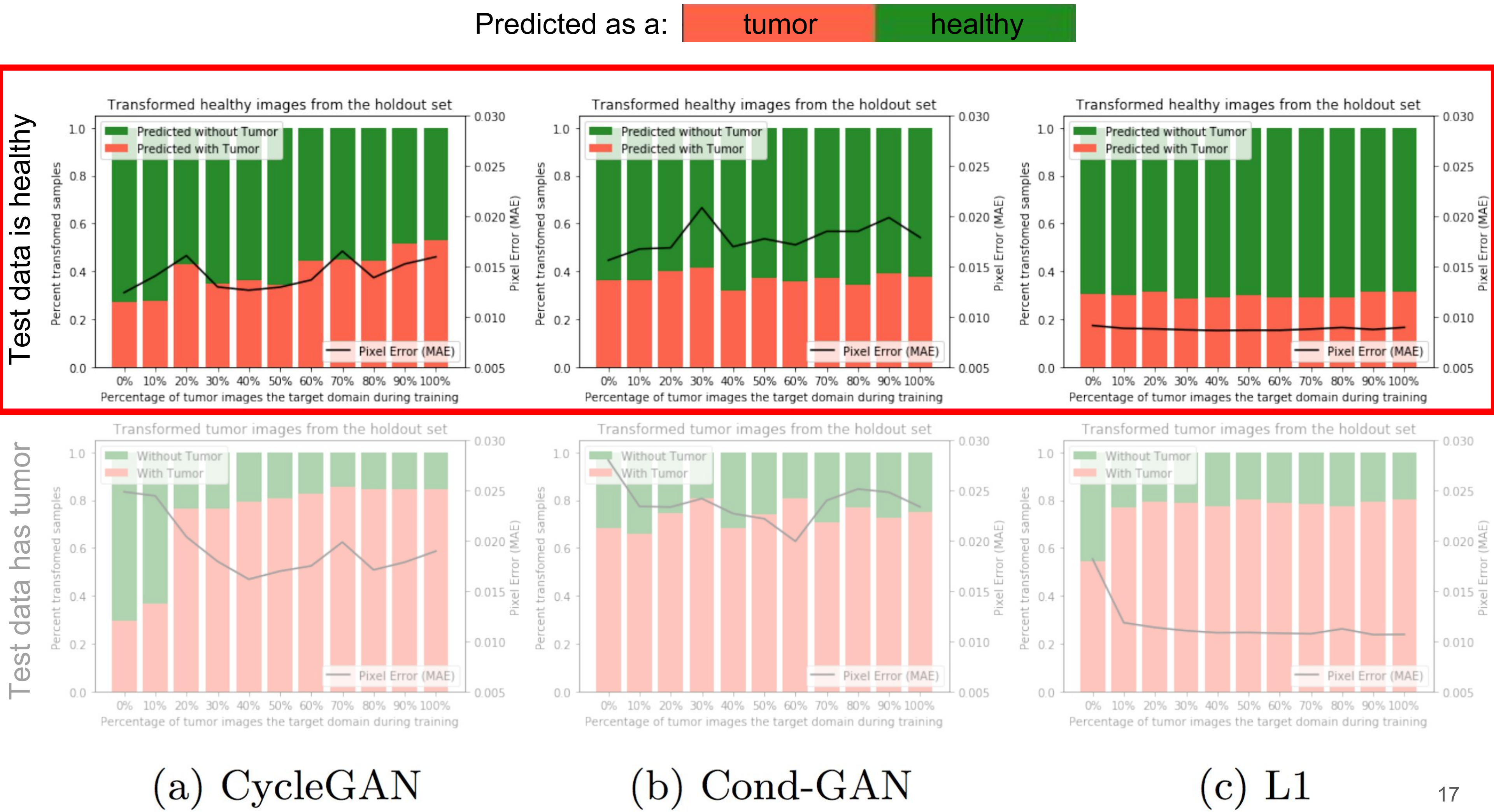




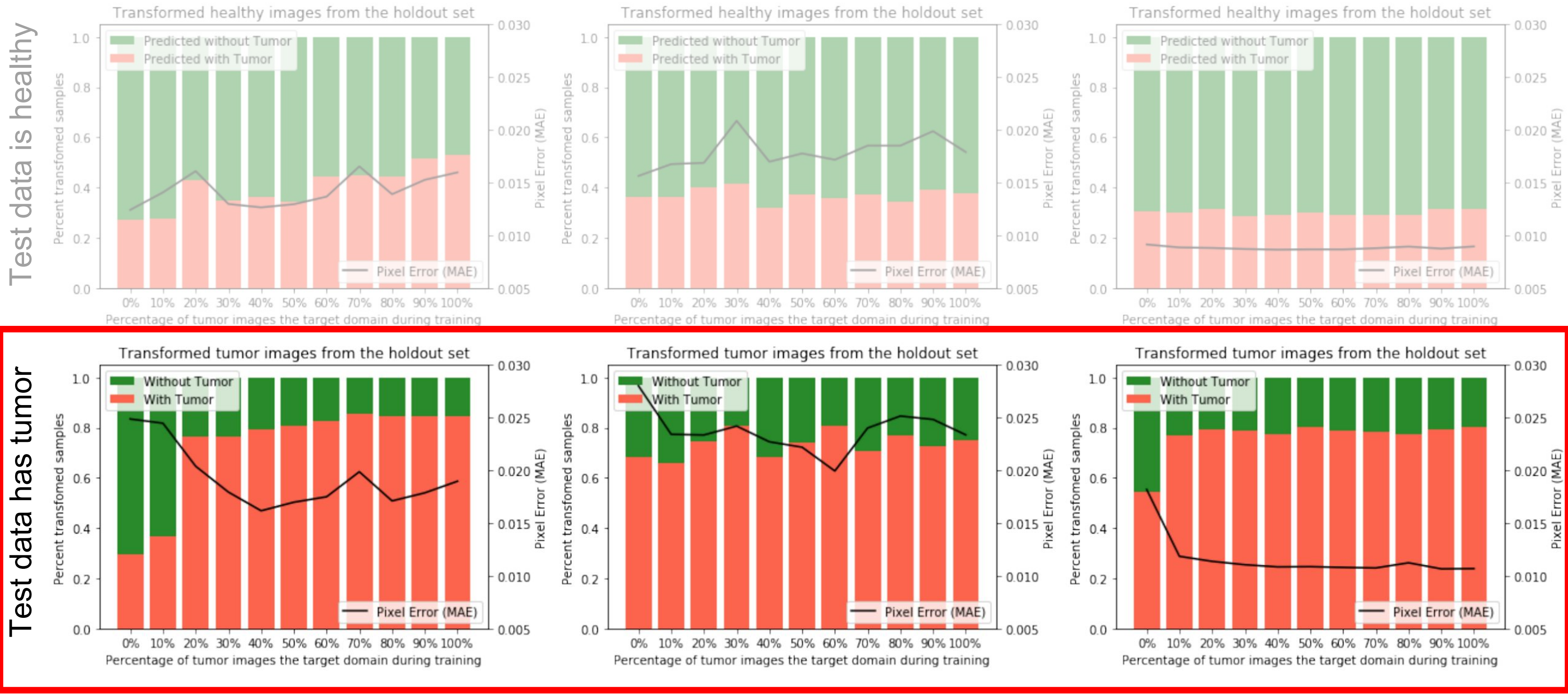
Example with a tumor from the holdout set

Quantitative Evaluation




Predicted as a:
tumor
healthy
Test data is healthy
Test data has tumor
Transformed healthy images from the holdout set
Predicted without Tumor
Predicted with Tumor
Pixel Error (MAE)
Percent transfomed samples
Pixel Error (MAE)
Percentage of tumor images the target domain during training
0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%
Transformed tumor images from the holdout set
Without Tumor
With Tumor
(a) CycleGAN
(b) Cond-GAN
(c) L1

Predicted as a: tumor healthy



(a) CycleGAN (b) Cond-GAN (c) L1

Our statement:

Image translation (via distribution matching) should not be used for direct interpretation.

Where do go from here?

1. How to guarantee image translation? (I doubt it)

2. Where should distribution matching be used in medical imaging?
    a. Data augmentation (for classification, segmentation, registration)
    b. Better features (for unsupervised learning)
    c. To correct model predictions [Zhang MICCAI 2017]

Limitations

- We test only a subset of loss terms which compose most methods
- The synthetic BRATS 2013 data had tumors added to healthy brains (in real data the entire brain is sick)

Mila Medical

Joseph Paul Cohen, PhD

Pr. Yoshua Bengio, PhD

Francis Dutil

Martin Weiss

Shawn Tan

Tristan Sylvain

Margaux Luck, PhD

Sina Honari

Assya Trofimov

Vincent Frappier, PhD

Thanks!


Mila
Medical


Université
de Montréal
DIRO


IVADO
INSTITUT DE VALORISATION DES DONNÉES


NSF


NSERC
CRSNG


computecanada


Calcul Québec


NVIDIA

Distribution Matching Losses Can Hallucinate Features in Medical Image Translation

Joseph Paul Cohen

Margaux Luck

Sina Honari

See us at poster M-60

https://arxiv.org/abs/1805.08841