MAML is a Noisy Contrastive Learner

Prof. **Wei-Chen Chiu** (NYCU) and Dr. **Pin-Yu Chen** (MIT-IBM Watson AI Lab)

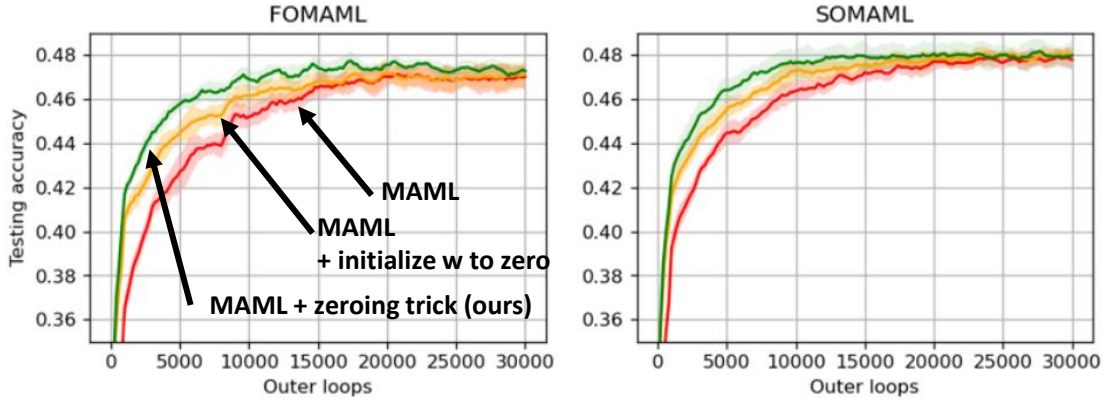
ICLR'22 submission

Presentation slides

Motivating questions and our answers

Q1	Why MAML learns generalized representation?
A1	Because MAML implicitly performs noisy, supervised **contrastive learning** (SCL).
Q2	What is the role of support and query data?
A2	In FOMAML, **support data** act as positive and negative **anchors** for SCL.
Q3	What does the bi-level optimization scheme do?
A3	In inner loops, **models memorize the support features**. In outer loops, models perform SCL.

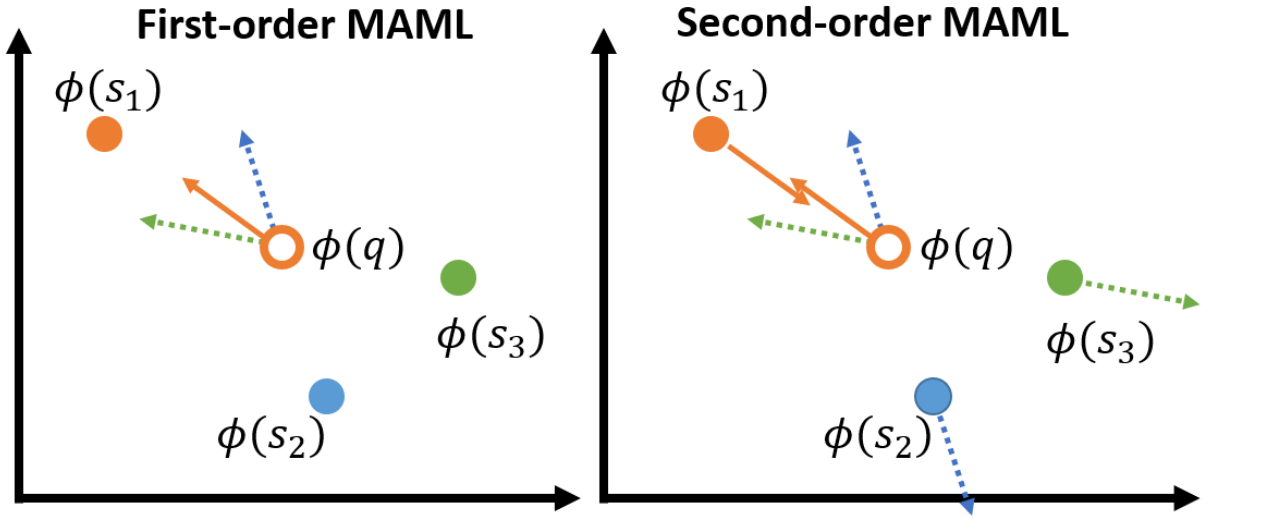
Results 5-way 1-shot on Mini-ImageNet.



Main difference between FOMAML and SOMAML

Support data: s_1, s_2, s_3
Query data: q
Encoder: ϕ

Gradient from positive sample (same color) ——→
Gradient from negative sample (diff. color) ·····→



update ϕ s.t. $\phi(q)$ is closer/further to $\phi(s)$ of the same/different class

update ϕ s.t. $\phi(q)$ and $\phi(s)$ is closer/further to each other of the same/different class

Mathematical Derivation

- We introduce the ANIL assumption, where the encoder ϕ of the models keeps frozen during the inner loop.
- We expand the outer loop loss (Softmax-NLL) and disclose the contrastiveness.

Demystifying T1-MRI to FDG18-PET Image Translation via Representational Similarity

Prof. **Yong-Sheng Chen** (NYCU), Prof. **Li-Fen Chen** (NYCU), Prof. **Wei-Chen Chiu** (NYCU)

MICCAI'21, Oral Presentation
Student Travel Award

Main ideas

Insights from clinical experience and physiology

- Brain's gray matter looks gray in T1-MRI images and white in FDG-PET images.
- Brain's white matter looks white in T1-MRI images and gray in FDG-PET images.

Intuition & hypothesis

- It is possible that the deep learning models learn to identify different brain tissue types and then assign each type with different FDG-PET values.

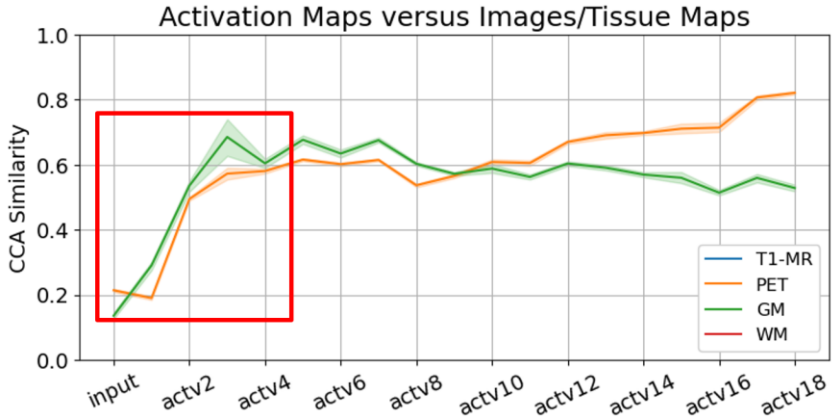
Observation in T1-MRI	Underlying anatomy and physiology meaning		Observation in PET
Gray matter	neuronal cell body	high metabolic activity	higher uptake value in PET
White matter	axon	low metabolic activity	lower uptake value in PET

Methods

Use Canonical Correlation Analysis (CCA) to understand the similarity between per-layer's activation and the brain tissue maps.

Results 1 (figure ↘) In the bottom layers, the information of brain tissues increases.

Results2 And we also show that in the middle encoding stage, the information about certain subcorticol brain regions increases. And these brain regions are known to have higher metabolic activity.



Contribution

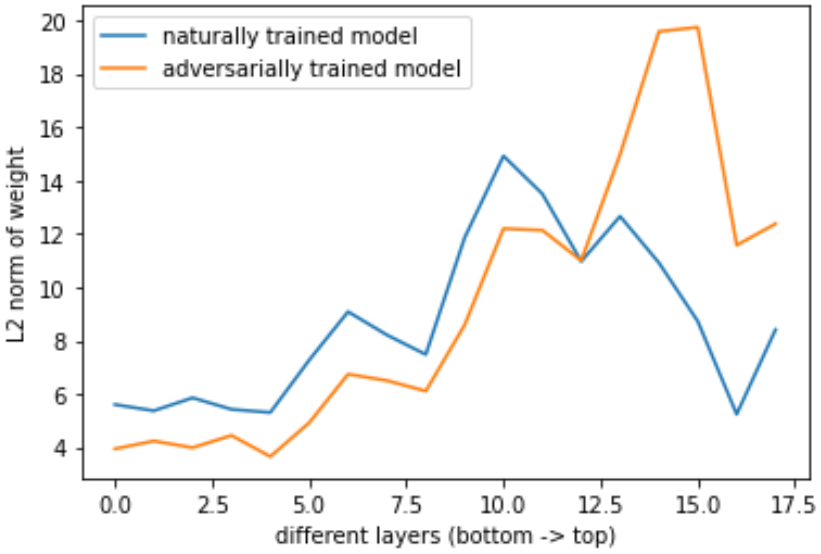
To build the groundwork for the biological plausibility of deep learning based cross-modal medical image translation models.

Ongoing project

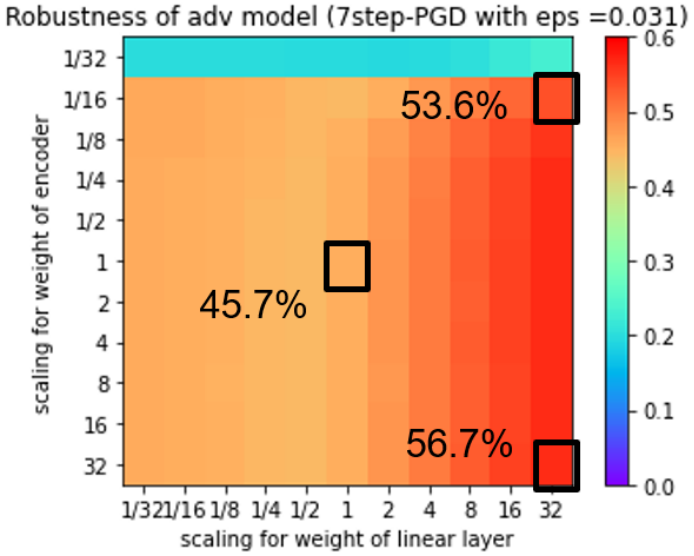
The Inductive Bias of Batch Normalization to Adversarial Training

Motivating observations

The top layers of adversarially-trained models have higher norm of weights than those of naturally trained models.



Scaling up the weight norm of the linear classifier increases robustness against PGD attack. (CIFAR10)



My preliminary thoughts

- Increasing the weight norm of the final layer scales up the logits and makes the SoftMax-output sharper. Sharper output implicitly causes gradient obfuscation, making gradient-based adversarial attack fails.
- My preliminary results indicate that the weight norm increasing phenomenon relates to batch normalization. On the contrary, label smoothing counteracts this effect and decreases the weight norm.