SELF-SUPERVISED DEEP LEARNING APPROACHES FOR MEDICAL IMAGE ANALYSIS

Cristian Simionescu

October 24, 2025

Faculty of Computer Science, Iasi

SIMIONESCU

SELF-SUPERVISED LEARNING

SIMIONESCU

MOTIVATION

- From 2010 to 2020, the amount of data created, captured, copied, and consumed in the world increased from 1.2 trillion gigabytes to 59 trillion gigabytes
- By extrapolation from 2021 to 2025, more data will be created than the past 30 years
- Human labeling can't keep up with the exponential growth curve of data collection

MOTIVATION

- Percentage of labeled data will rapidly decline compared to unlabeled data
- Hardware is continuing to provide increased parallel and single threaded performance
- Deep Learning model which can scale with more data and compute are becoming state of the art in all fields (e.g Transformers)

As a result, Self-Supervised Learning will probably be the most important driver of future AI SOTA improvements

SELF-SUPERVISED LEARNING - PREDICTIVE METHODS

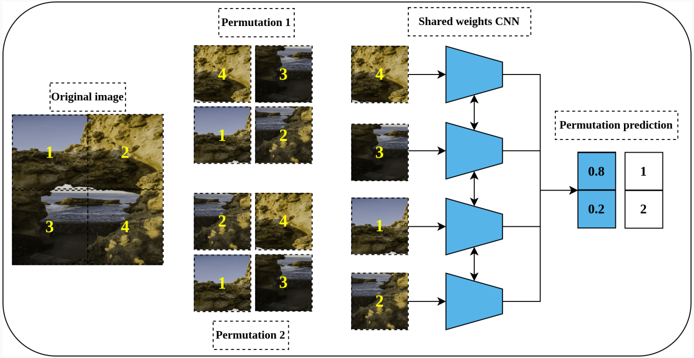


Figure: Example of Jigsaw pretext task

SELF-SUPERVISED LEARNING - GENERATIVE METHODS

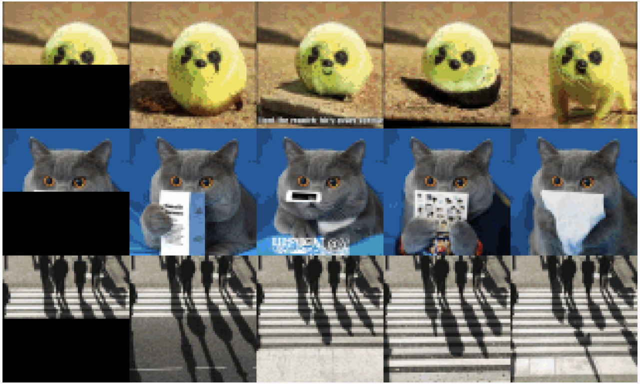


Figure: Example of Autoregressive Generation, Image GPT

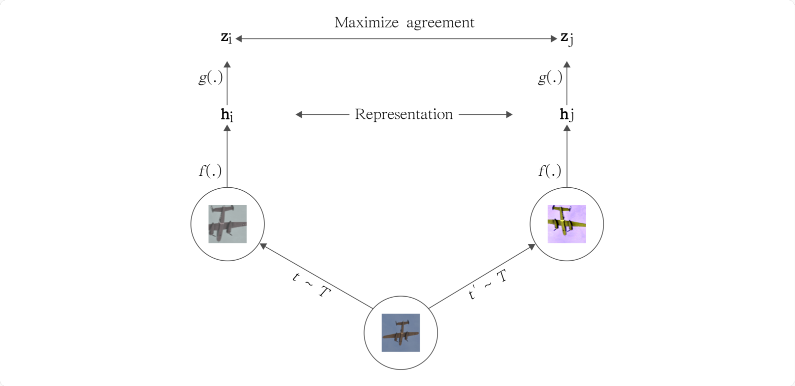
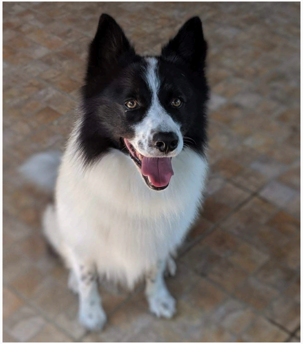


Figure: SimCLR Architecture

SELF-SUPERVISED LEARNING VS SUPERVISED

(a) Asist. Pixel



(b) Asist. Bruno

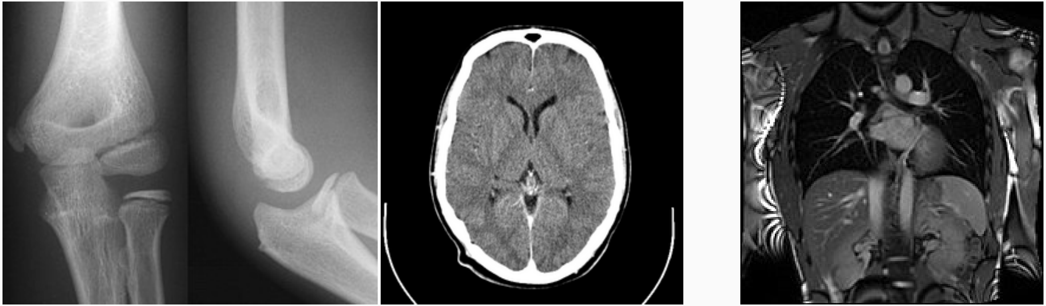
Figure: When training a classifier model in a supervised way, it only needs to "understand" enought to differentiate between the two classes (eg. only looking

SELF-SUPERVISED LEARNING FOR MEDICAL DATA

SIMIONESCU

MEDICAL IMAGE ANALYSIS

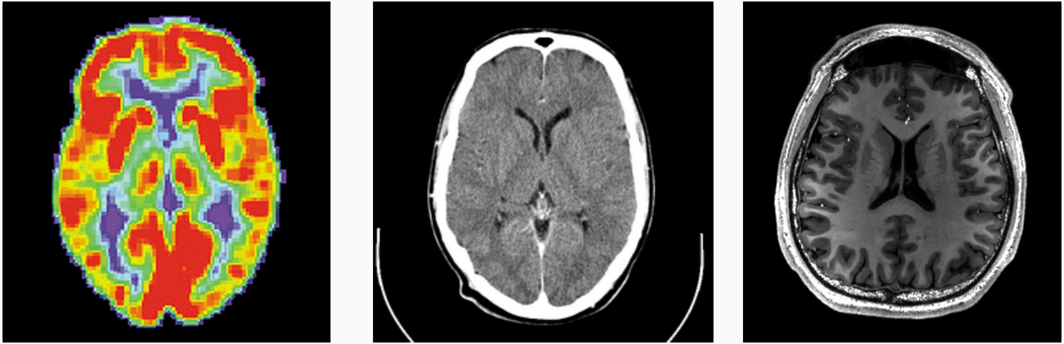
- Quality labeling can only be done by highly specialized medical professionals, slower, smaller capacity
- Non-standardization of medical practices, terminology, technology
- Individual medical datasets are small
- Research delay from natural images methods to medical imaging methods
- Smaller researcher base
- Improvements on medical tasks can have a bigger societal impact compared to natural image tasks



(a) X-ray of an arm (b) CT scan of the brain (c) MRI of the chest

Figure: Examples of heterogeneous medical imaging modalities and subjects illustrating the diversity in data (X-ray, CT, and MRI).

DATA HOMOGENEITY



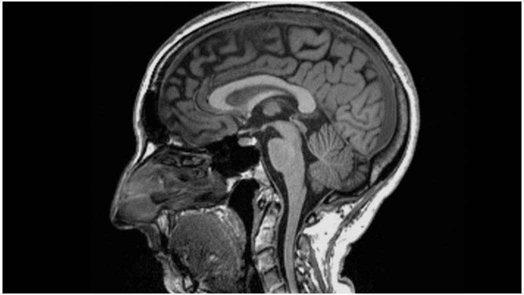
(a) PET scan of the brain (b) CT scan of the brain (c) MRI of the brain

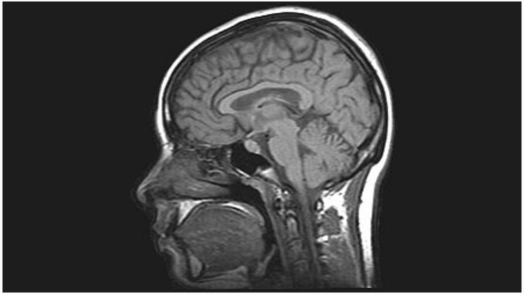
Figure: Examples of homogeneity in medical imaging modalities illustrating the similarity of the underlying subject (PET, CT, and MRI).

OPPORTUNITY

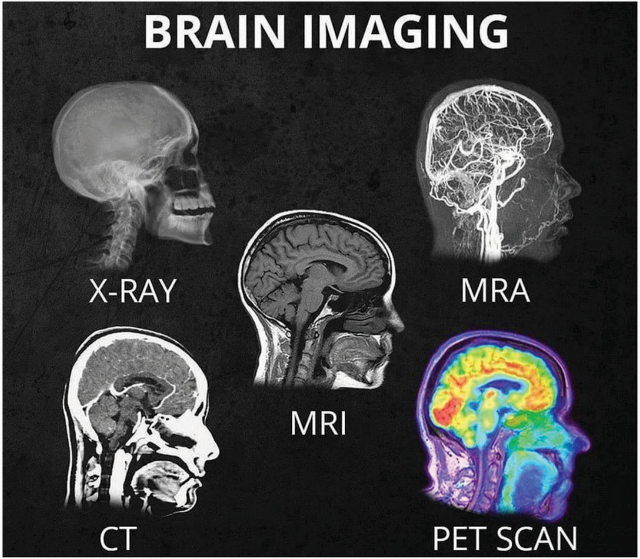
Medical images share a lot of information between them.

Scans with the same technology (CT, MRI) on the same organ, will share a lot of features, even if the pathology of the two patients is completely unrelated.

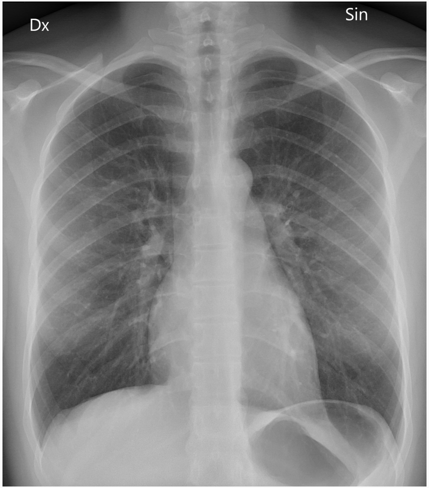




Scans with a different technology, of the same organ, will share many structural features as well as textural information.



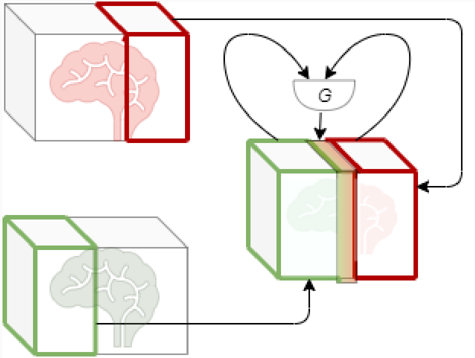
Scans of different organs with the same technology will also share features since the scans work by capturing a small set of frequencies which translate to types of tissues.



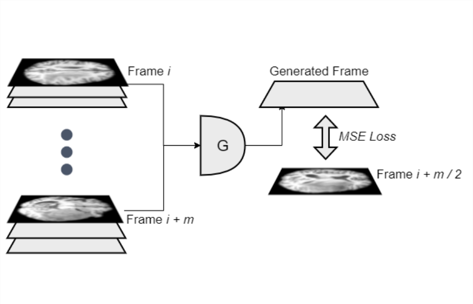
CHALLENGES IN MEDICAL IMAGING

- **Modality/Site Heterogeneity**: X-ray, CT, MRI, PET can differ significantly.
- **Intra-modality Homogeneity**: High visual uniformity within modality complicates learning.
- **Data Scarcity**: Labeled data is hard to obtain, expensive, and fragmented.
- **Generalization**: Pre-trained models (e.g., ImageNet) don't transfer well.

BRANFUSE

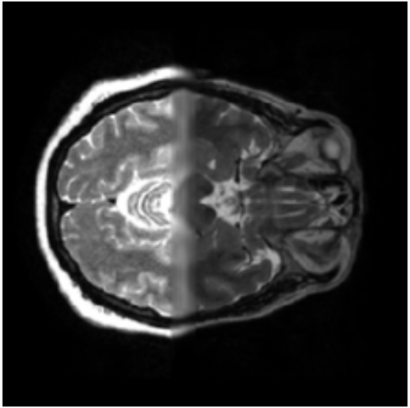


(a) Brain Mixing Idea

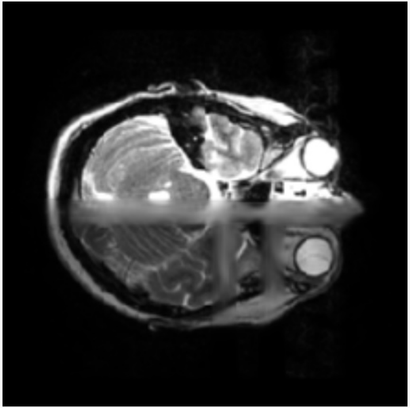


(b) Generative Model Training Idea

Figure: Overview.



(a) Generated brain by fusing two patients.



(b) Generated brain by fusing eight patients.

Method	Top1 Acc	Top3 Acc	Top5 Acc
NT (Normal Training)	$32.1 \pm 0.27\%$	$53.8 \pm 0.18\%$	$64.21 \pm 0.32\%$
NT + BT (Barlow/Twins)	$40.1 \pm 0.77\%$	$62.8 \pm 0.11\%$	$73.01 \pm 0.1\%$
NT + BT + BrainFuse	$50.11 \pm 0.3\%$	$71.16 \pm 0.12\%$	$82.9 \pm 0.22\%$
NT + BT + RawBrainFuse	$45.1 \pm 0.94\%$	$67.5 \pm 0.41\%$	$80.01 \pm 0.1\%$
NT + BT + BrainFuse + Consistency Regularisation	**51.4** $\pm 0.84\%$	**72.1** $\pm 0.61\%$	**83.8** $\pm 0.52\%$
NT + BT + RawBrainFuse + Consistency Regularisation	$46.6 \pm 0.9\%$	$68.8 \pm 0.78\%$	$81.1 \pm 0.31\%$

Table: Results for the different training procedures. Each procedure has been run 3 times

MEDFORMER OVERVIEW

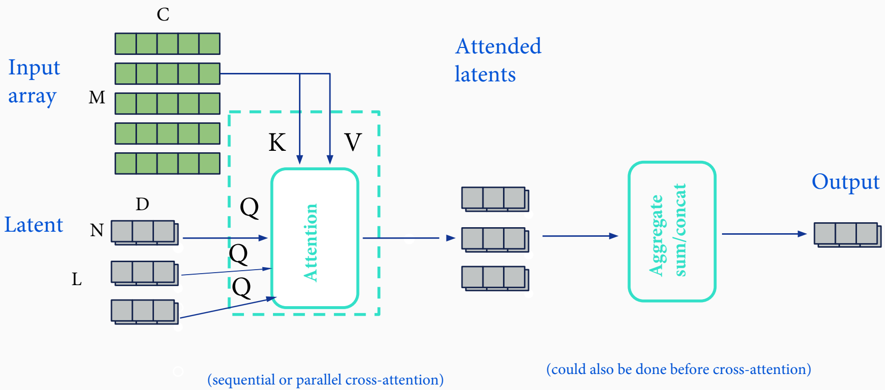
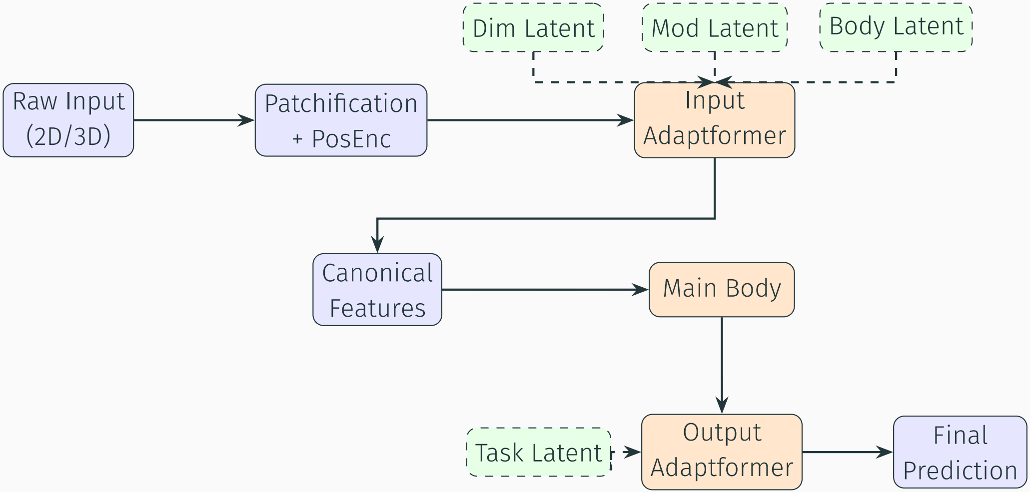


Figure: Adaptformer Architecture

MEDFORMER OVERVIEW



MEDMNIST DATASET

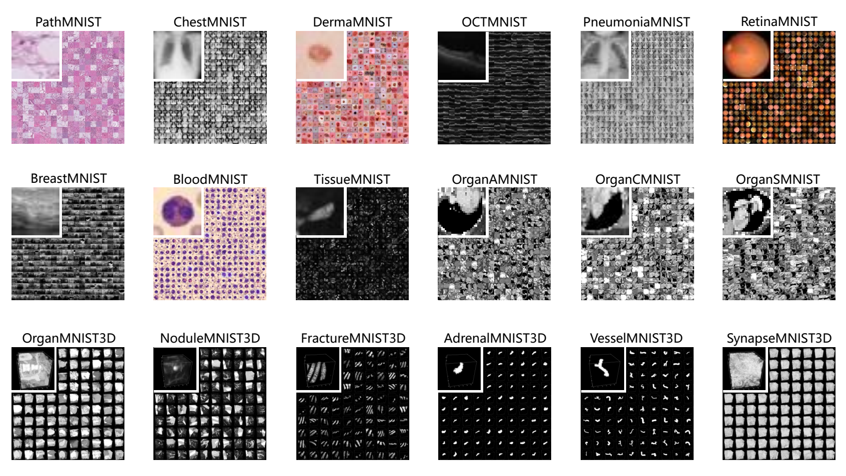


Figure: MedMNIST Dataset

RESULTS

Dataset	MT – ST		SSL – ST		SSL – MT		SSL Steps
	AUC	Acc	AUC	Acc	AUC	Acc	
PathMNIST	-0.004	-1.0	-0.001	-5.3	+0.003	-4.3	200
DermaMNIST	-0.569	+23.0	+0.094	+2.9	+0.663	-20.1	150
ChestMNIST	+0.192	-0.1	+0.178	0.0	-0.014	+0.1	50
RetinaMNIST	+0.025	-5.5	+0.038	0.0	+0.013	+5.5	50
PneumoniaMNIST	-0.030	-3.0	+0.007	+1.9	+0.037	+4.9	10
OCTMNIST	-0.495	+6.6	-0.001	-0.4	+0.494	-7.0	500
BreastMNIST	+0.083	-2.6	+0.109	-6.4	+0.026	-3.8	30
BloodMNIST	+0.000	+0.0	+0.025	+14.2	+0.025	+14.2	30
TissueMNIST	-0.002	-11.2	+0.003	+1.3	+0.005	+12.5	500
OrganAMNIST	+0.000	-2.4	+0.000	+0.3	+0.000	+2.7	50

RESULTS

Dataset	MT – ST		SSL – ST		SSL – MT		SSL Steps
	AUC	Acc	AUC	Acc	AUC	Acc	
OrganMNIST3D	-0.013	-20.5	-0.005	-6.2	+0.008	+14.3	40
NoduleMNIST3D	-0.015	+1.8	-0.006	+1.8	+0.009	0.0	20
AdrenalMNIST3D	+0.024	+9.2	+0.044	+16.3	+0.020	+7.1	60
FractureMNIST3D	+0.123	+0.0	+0.131	+3.9	+0.008	+3.9	120
VesselMNIST3D	+0.019	-1.0	+0.032	+1.6	+0.013	+2.6	50
SynapseMNIST3D	+0.016	+0.0	+0.056	-1.1	+0.040	-1.1	125

/cristian.simionescu.92

@CrisSimi2525

simi2525

cristian.simionescu.me

/cristian-simionescu/

Cristian Simionescu