Sber AI Lab Medical Imaging Group Projects

Ilya Burenko, Sr. Data Scientist

СБЕР

Outline

- Our projects;

СБЕР

Outline

- Our projects;
- Image restoration on PET-CT images;

СБЕР

Outline

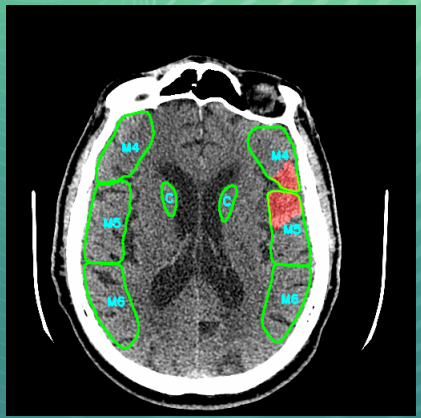
- Our projects;
- Image restoration on PET-CT images;
- Unsupervised pretraining for segmentation of CT studies.

СБЕР

The overview of projects

CT of head:

- Segmentation of lesions affected by acute stroke;

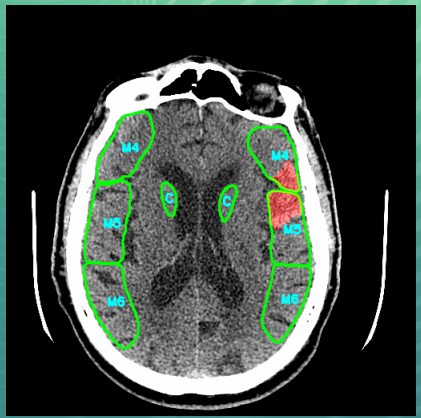


СБЕР

The overview of projects

CT of head:

- Segmentation of lesions affected by acute stroke;
- Segmentation of regions of brain (ASPECTS);



СБЕР

The overview of projects

CT of head:

- Segmentation of lesions affected by acute stroke;
- Segmentation of regions of brain (ASPECTS);
- (jointly with AIRI) Detection of very early stroke on CT (up to 12 hours);

СБЕР

The overview of projects

CT of lungs:

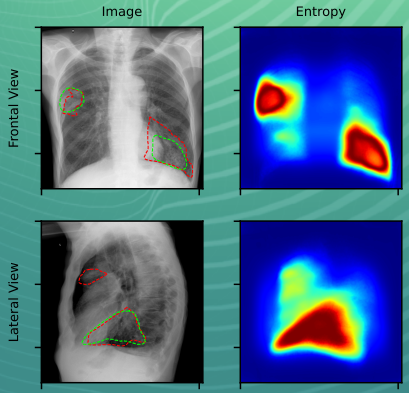
- Calculate the volume of lungs affected by covid

СБЕР

The overview of projects

X-rays of chest and breast:

- Restoration of segmentation masks using uncertainty estimation on chest X-rays;

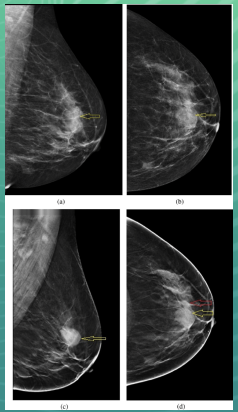


СБЕР

The overview of projects

X-rays of chest and breast:

- Restoration of segmentation masks using uncertainty estimation on chest X-rays;
- Classification and segmentation of breast cancer.



СБЕР

The overview of projects

ECG:

- Classification of QRS-complexes;

СБЕР

The overview of projects

ECG:

- Classification of QRS-complexes;
- Risk of type II diabetes;

СБЕР

Restoration of PET-CT images

PET CT is used for localize cancer cells in a body.
How a study is conducted?

- An injection of FDG (Fludeoxyglucose);

СБЕР

Restoration of PET-CT images

PET CT is used for localize cancer cells in a body.
How a study is conducted?

- An injection of FDG (Fludeoxyglucose);
- PET study;

СБЕР

Restoration of PET-CT images

PET CT is used for localize cancer cells in a body.
How a study is conducted?

- An injection of FDG (Fludeoxyglucose);
- PET study;
- CT study;

СБЕР

Restoration of PET-CT images

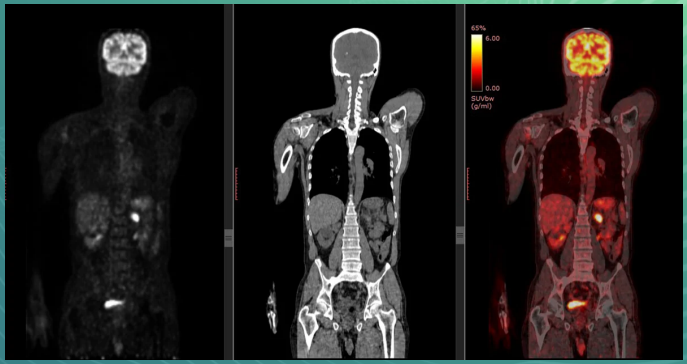


Figure: Example of PET-CT study

СБЕР

Restoration of PET-CT images

Possible tasks:

- Reduce the amount of FDG;

СБЕР

Restoration of PET-CT images

Possible tasks:

- Reduce the amount of FDG;
- Conduct the study faster.

СБЕР

Restoration of PET-CT images

The standard exposure time for PET study is 90 seconds. But we have PET data after 30 and 60 seconds of exposure as well.

Q: Is it possible to restore the standard exposure time images?

СБЕР

Restoration of PET-CT images

Possible approaches:

- Deterministic (classical image processing or deep learning approaches); – One input image -> one output image;

СБЕР

Restoration of PET-CT images

Possible approaches:

- Deterministic (classical image processing or deep learning approaches); – One input image -> one output image;
- Probabilistic; one input image -> multiple images, e.g. probability distribution.

СБЕР

Restoration of PET-CT images

Approaches:

- Gaussian filtration;

СБЕР

Restoration of PET-CT images

Approaches:

- Gaussian filtration;
- U-net like Transformer deep network;

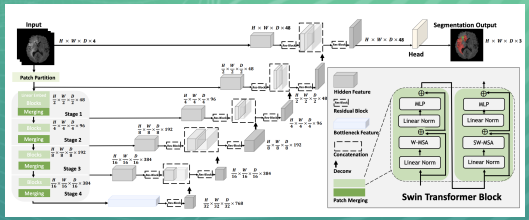


Figure: From Hatamizadeh et. al "Swin UNETR: Swin Transformers for Semantic Segmentation of Brain Tumors in MRI Images"

СБЕР

Restoration of PET-CT images

Approaches:

- Gaussian filtration;
- U-net like Transformer deep network;
- Diffusion model (Work in progress).

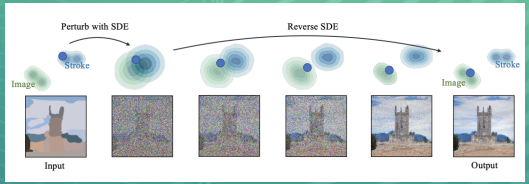
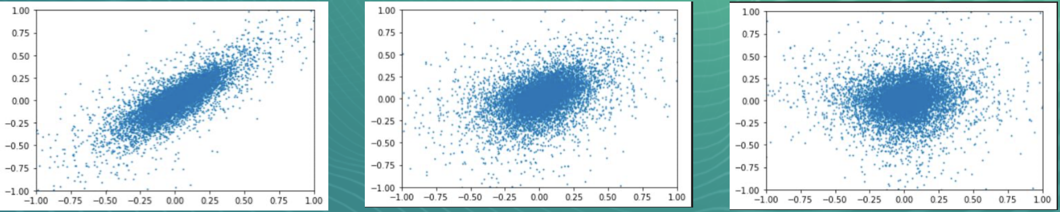


Figure: From Meng et. al "SDEdit: Guided Image Synthesis And Editing With Stochastic Differential Equations"

СБЕР

Restoration of PET-CT images

Pixel-wise difference between adjacent pixels (left), pixels at distance 2 (center), and 3 (right) of the difference Δ_o =PET$_{60}$ - PET$_{90}$:



СБЕР

Restoration of PET-CT images

Current results. Calculated as

$$100 * (1 - \frac{\Delta_r}{\Delta_o}),$$

where $\Delta_r = \text{PET}_r - \text{PET}_{90}$ and $\Delta_o = \text{PET}_{60} - \text{PET}_{90}$.

Method	MSE	1 - SSIM
Gaussian filtration (1 layer)	11.8 %	3.2%
Gaussian filtration (3 layer)	15.2%	-58.6%
Swin Transformer	16.6%	-128.6%
Diffusion model	Work in progress	

Restoration of PET-CT images

Current results

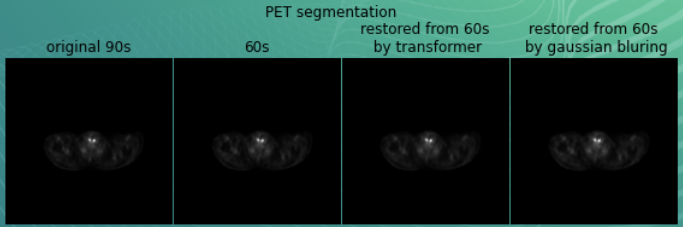
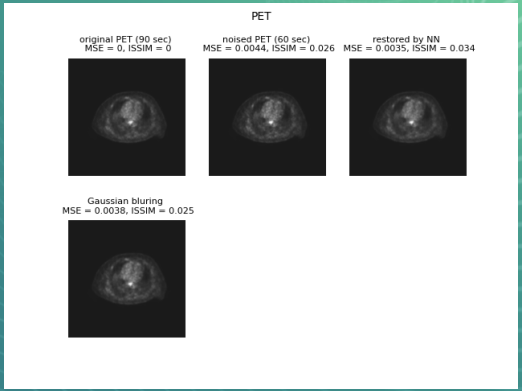


Figure: Restoration of the noised PET

Restoration of PET-CT images

Current results



СБЕР

Unsupervised pretraining on CT studies

Statement of the problem

- CT has become easily available and relatively cheap;

СБЕР

Unsupervised pretraining on CT studies

Statement of the problem

- CT has become easily available and relatively cheap;
- There are a lot of very massive CT datasets (annotated and not);

СБЕР

Unsupervised pretraining on CT studies

Statement of the problem

- CT has become easily available and relatively cheap;
- There are a lot of very massive CT datasets (annotated and not);
- A lot of medical centers want to predict a presence of some disease based on a limited amount of data;

СБЕР

Unsupervised pretraining on CT studies

Statement of the problem

- Pretrain a deep learning model that is able to generalize well for a wide range of downstream tasks.

СБЕР

Unsupervised pretraining on CT studies

We will consider segmentation task since:

- If training from scratch: requires a lot of data;

СБЕР

Unsupervised pretraining on CT studies

We will consider segmentation task since:

- If training from scratch: requires a lot of data;
- Requires networks with more parameters compared to, for instance, classification, hence more computational resources.

СБЕР

Unsupervised pretraining on CT studies

We will consider segmentation task since:

- If training from scratch: requires a lot of data;
- Requires networks with more parameters compared to, for instance, classification, hence more computational resources.
- Pretraining is a way to overcome these shortcomings!

СБЕР

Unsupervised pretraining on CT studies

Possible pretraining approaches:

- Supervised pretrain;

СБЕР

Unsupervised pretraining on CT studies

Possible pretraining approaches:

- Supervised pretrain;
- Unsupervised pretrain:

СБЕР

Unsupervised pretraining on CT studies

Possible pretraining approaches:

- Supervised pretrain;
- Unsupervised pretrain:
  1. Inpainting;

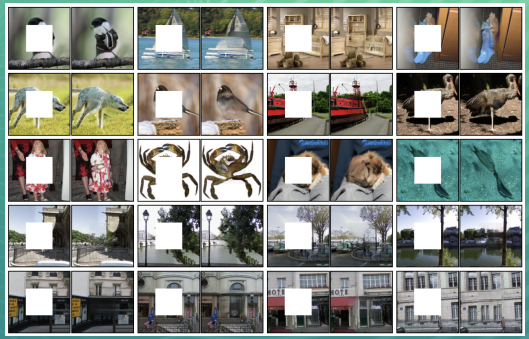


Figure: From Pathak et. al "Context Encoders: Feature Learning by Inpainting"

СБЕР

Unsupervised pretraining on CT studies

Possible pretraining approaches:

- Supervised pretrain;
- Unsupervised pretrain:
  1. Inpainting;
  2. Image rotation;

СБЕР

Unsupervised pretraining on CT studies

Possible pretraining approaches:

- Supervised pretrain;
- Unsupervised pretrain:
  1. Inpainting;
  2. Image rotation;
  3. Contrastive learning.

СБЕР

Unsupervised pretraining on CT studies

Contrastive learning:

- We would like to maximize mutual information between *similar* images [1].

$$I(x, x') = \sum_{x,x'} p(x, x') \log \frac{p(x|x')}{p(x)};$$

[1] See: van den Oord et al. Representation Learning with Contrastive Predictive Coding

Contrastive learning:

- We would like to maximize mutual information between *similar* images [1].

$$I(x, x') = \sum_{x,x'} p(x, x') \log \frac{p(x|x')}{p(x)};$$

- In fact we will minimize

$$\mathcal{L} = -\mathbb{E}_X \left[\frac{f(x, x')}{\sum_{j=1}^{N} f(x_j, x')} \right], \text{where } f(x, x') \propto \frac{p(x|x')}{p(x)}.$$

[1] See: van den Oord et al. Representation Learning with Contrastive Predictive Coding

СБЕР

Unsupervised pretraining on CT studies

But deep learning benefits from scalable approaches!

SimCLR

- Sample N images;

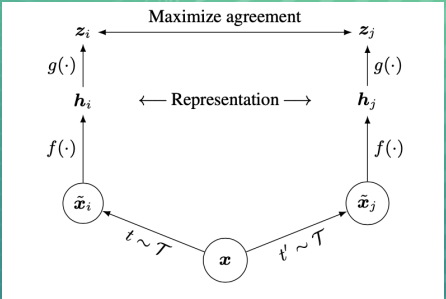


Figure: From Chen et. al "A Simple Framework for Contrastive Learning of Visual Representations"

СБЕР

Unsupervised pretraining on CT studies

But deep learning benefits from scalable approaches!

SimCLR

- Sample N images;
- Augment each image;

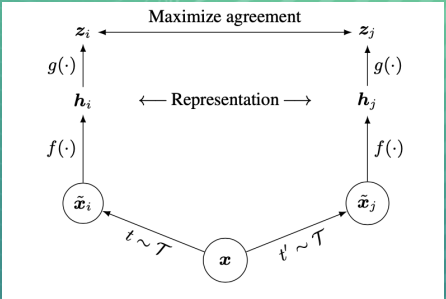


Figure: From Chen et. al "A Simple Framework for Contrastive Learning of Visual Representations"

СБЕР

Unsupervised pretraining on CT studies

But deep learning benefits from scalable approaches!

SimCLR

- Sample N images;
- Augment each image;
- Obtain representations h using neural network $f(\cdot)$;

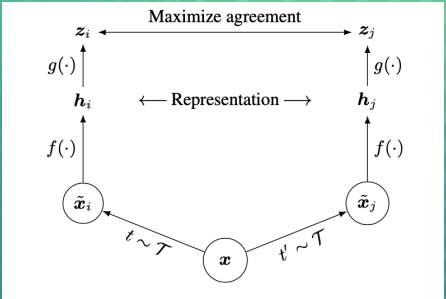


Figure: From Chen et. al "A Simple Framework for Contrastive Learning of Visual Representations"

СБЕР

Unsupervised pretraining on CT studies

But deep learning benefits from scalable approaches!

SimCLR

- Sample N images;
- Augment each image;
- Obtain representations h using neural network $f(\cdot)$;
- Obtain embeddings z using non-linear head $g(\cdot)$

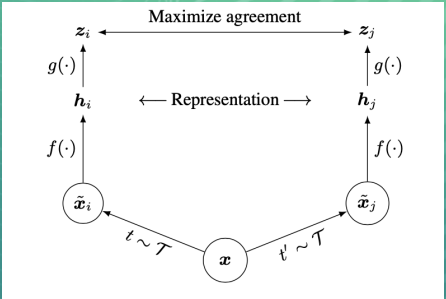


Figure: From Chen et. al "A Simple Framework for Contrastive Learning of Visual Representations"

СБЕР

Unsupervised pretraining on CT studies

But deep learning benefits from scalable approaches!

SimCLR

- Sample N images;
- Augment each image;
- Obtain representations h using neural network $f(\cdot)$;
- Obtain embeddings z using non-linear head $g(\cdot)$
- Maximize agreement (MI) using contrastive loss;

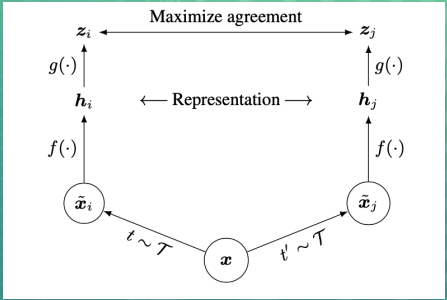


Figure: From Chen et. al "A Simple Framework for Contrastive Learning of Visual Representations"

СБЕР

Unsupervised pretraining on CT studies

Contrastive loss:

- For an augmented pair of images:

$$\mathcal{L}(x_i, x_j) = -\log \frac{\exp(z_i \cdot z_j/\tau)}{\sum_{k=1}^{2N} \mathbb{1}_{k \neq j} \left[\exp(z_i \cdot z_k/\tau)\right]};$$

Unsupervised pretraining on CT studies

Contrastive loss:

- For an augmented pair of images:

$$\mathcal{L}(x_i, x_j) = -\log \frac{\exp(z_i \cdot z_j/\tau)}{\sum_{k=1}^{2N} \mathbb{1}_{k \neq j} \left[\exp(z_i \cdot z_k/\tau)\right]};$$

- Total loss:

$$\mathcal{L} = \frac{1}{2N} \sum_{i=1}^{N} \left(\mathcal{L}(x_{2k-1}, x_{2k}) + \mathcal{L}(x_{2k}, x_{2k-1})\right).$$

Unsupervised pretraining on CT studies

Why this (and a lot of similar approaches) are so attractable?

- Contrastive learning (in the original formulation) does not require ground truth labels;

СБЕР

Unsupervised pretraining on CT studies

Why this (and a lot of similar approaches) are so attractable?

- Contrastive learning (in the original formulation) does not require ground truth labels;
- Linear evaluation protocol shows results on par with supervised methods.

СБЕР

Unsupervised pretraining on CT studies

Some observations:

- SimCLR and similar methods are based on image-level comparisons;

СБЕР

Unsupervised pretraining on CT studies

Some observations:

- SimCLR and similar methods are based on image-level comparisons;
- Which might be sub-optimal for segmentation tasks due to the lack of spatial sensitivity.

СБЕР

Unsupervised pretraining on CT studies

Requirements for pretext task:

- It should be **spatial sensitive**, i.e. discriminate spatially closed pixels for accurate predictions in boundary regions;

СБЕР

Unsupervised pretraining on CT studies

Requirements for pretext task:

- It should be **spatial sensitive**, i.e. discriminate spatially closed pixels for accurate predictions in boundary regions;
- It should be **spatial smooth**. Spatial smoothness encourage clone pixels to belong to the same class.

СБЕР

Unsupervised pretraining on CT studies

Propagate yourself[2]:

- Choose a convolutional neural network $f(\cdot)$;

[2] See Xie et al. Propagate Yourself: Exploring Pixel-Level Consistency for Unsupervised Visual Representation Learning

Propagate yourself[2]:

- Choose a convolutional neural network $f(\cdot)$;
- Process an input image $\mathbb{I} \in \mathbb{R}^{1 \times H \times W}$ using $f(\cdot)$ to obtain a representation from some convolutional layer $\tilde{\mathbb{I}} \in \mathbb{R}^{C \times H' \times W'}$;

[2] See Xie et al. Propagate Yourself: Exploring Pixel-Level Consistency for Unsupervised Visual Representation Learning

Propagate yourself[2]:

- Choose a convolutional neural network $f(\cdot)$;
- Process an input image $\mathbb{I} \in \mathbb{R}^{1 \times H \times W}$ using $f(\cdot)$ to obtain a representation from some convolutional layer $\tilde{\mathbb{I}} \in \mathbb{R}^{C \times H' \times W'}$;
- Consider each C-dimensional vector of the $\tilde{\mathbb{I}}$ as a representation of a pixel in some pixel-space.

[2] See Xie et al. Propagate Yourself: Exploring Pixel-Level Consistency for Unsupervised Visual Representation Learning

СБЕР

Unsupervised pretraining on CT studies

The pipeline:

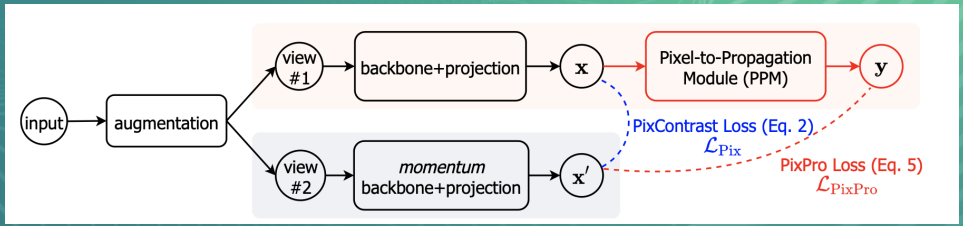


Figure: See Xie et al. Propagate Yourself: Exploring Pixel-Level Consistency for Unsupervised Visual Representation Learning

СБЕР

Unsupervised pretraining on CT studies

- Pixel-level contrastive loss:

$$\mathcal{L}_{\mathcal{P}}(i) = -log \frac{\sum\limits_{j \in \Omega_p^i} e^{\cos(x_i, x_j')/\tau}}{\sum\limits_{j \in \Omega_p^i} e^{\cos(x_i, x_j')/\tau} + \sum\limits_{k \in \Omega_n^i} e^{\cos(x_i, x_k')/\tau}};$$

СБЕР

Unsupervised pretraining on CT studies

- Pixel-level contrastive loss:

$$\mathcal{L}_{\mathcal{P}}(i) = -log \frac{\sum\limits_{j \in \Omega_p^i} e^{\cos(x_i, x_j')/\tau}}{\sum\limits_{j \in \Omega_p^i} e^{\cos(x_i, x_j')/\tau} + \sum\limits_{k \in \Omega_n^i} e^{\cos(x_i, x_k')/\tau}};$$

- Where x_i, x_j' – pixels from two augmented versions of an image x;

СБЕР

Unsupervised pretraining on CT studies

- Pixel-level contrastive loss:

$$\mathcal{L}_{\mathcal{P}}(i) = -log \frac{\sum\limits_{j \in \Omega_p^i} e^{\cos(x_i, x_j')/\tau}}{\sum\limits_{j \in \Omega_p^i} e^{\cos(x_i, x_j')/\tau} + \sum\limits_{k \in \Omega_n^i} e^{\cos(x_i, x_k')/\tau}};$$

- Where x_i, x_i' – pixels from two augmented versions of an image x;
- Ω_p^i and Ω_n^i – pixels inside and outside some vicinity of the current pixel i respectively.

СБЕР

Unsupervised pretraining on CT studies

- The total loss:

$$\mathcal{L} = \mathcal{L}_{\text{inst}} + \frac{1}{H' \times W'} \sum_{i} \mathcal{L}_{\mathcal{P}}(i)$$

Unsupervised pretraining on CT studies

We will compare the following methods:

- Training from a random initialization;

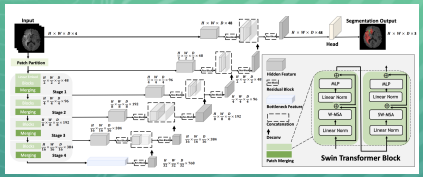


Figure: From Hatamizadeh et. al "Swin UNETR: Swin Transformers for Semantic Segmentation of Brain Tumors in MRI Images"

СБЕР

Unsupervised pretraining on CT studies

We will compare the following methods:

- Training from a random initialization;
- Fine-tuning a model pretrained on (Nvidia)

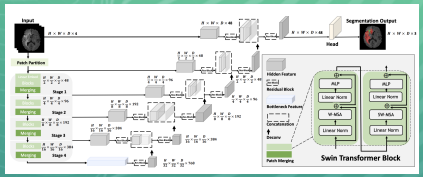


Figure: From Hatamizadeh et. al "Swin UNETR: Swin Transformers for Semantic Segmentation of Brain Tumors in MRI Images"

Unsupervised pretraining on CT studies

We will compare the following methods:

- Training from a random initialization;
- Fine-tuning a model pretrained on (Nvidia)
  - Inpainting;

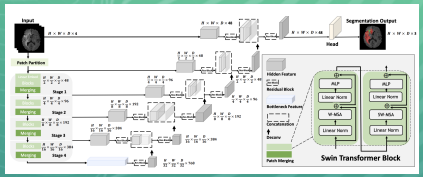


Figure: From Hatamizadeh et. al "Swin UNETR: Swin Transformers for Semantic Segmentation of Brain Tumors in MRI Images"

СБЕР

Unsupervised pretraining on CT studies

We will compare the following methods:

- Training from a random initialization;
- Fine-tuning a model pretrained on (Nvidia)
  - Inpainting;
  - Rotation prediction;

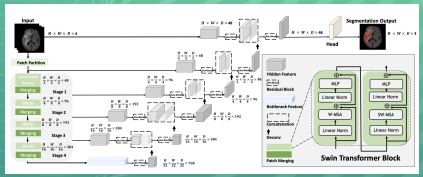


Figure: From Hatamizadeh et. al "Swin UNETR: Swin Transformers for Semantic Segmentation of Brain Tumors in MRI Images"

Unsupervised pretraining on CT studies

We will compare the following methods:

- Training from a random initialization;
- Fine-tuning a model pretrained on (Nvidia)
  - Inpainting;
  - Rotation prediction;
  - Instance-level CL.

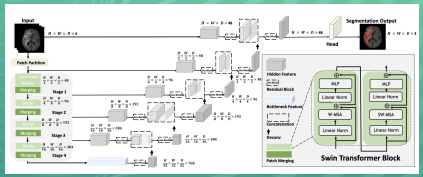


Figure: From Hatamizadeh et. al "Swin UNETR: Swin Transformers for Semantic Segmentation of Brain Tumors in MRI Images"

Unsupervised pretraining on CT studies

We will compare the following methods:

- Training from a random initialization;
- Fine-tuning a model pretrained on (Nvidia)
  - Inpainting;
  - Rotation prediction;
  - Instance-level CL.
- Fine-tuning a model pretrained on (Sber)

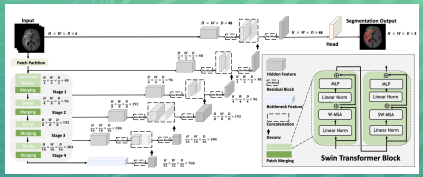


Figure: From Hatamizadeh et. al "Swin UNETR: Swin Transformers for Semantic Segmentation of Brain Tumors in MRI Images"

Unsupervised pretraining on CT studies

We will compare the following methods:

- Training from a random initialization;
- Fine-tuning a model pretrained on (Nvidia)
  - Inpainting;
  - Rotation prediction;
  - Instance-level CL.
- Fine-tuning a model pretrained on (Sber)
  - Instance-level CL;

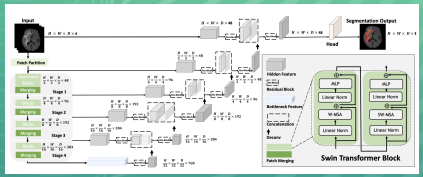


Figure: From Hatamizadeh et. al "Swin UNETR: Swin Transformers for Semantic Segmentation of Brain Tumors in MRI Images"

СБЕР

Unsupervised pretraining on CT studies

We will compare the following methods:

- Training from a random initialization;
- Fine-tuning a model pretrained on (Nvidia)
  - Inpainting;
  - Rotation prediction;
  - Instance-level CL.
- Fine-tuning a model pretrained on (Sber)
  - Instance-level CL;
  - Pixel-level CL.

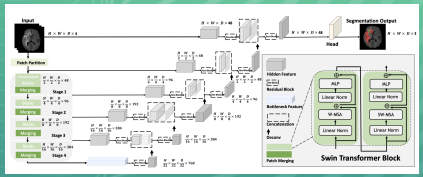


Figure: From Hatamizadeh et. al "Swin UNETR: Swin Transformers for Semantic Segmentation of Brain Tumors in MRI Images"

СБЕР

Unsupervised pretraining on CT studies

- Pretraining data:

СБЕР

Unsupervised pretraining on CT studies

- Pretraining data:
  - Nvidia: LUNA16, TCIA Covid19, LiDC (Total: 2124 CTs);

СБЕР

Unsupervised pretraining on CT studies

- Pretraining data:
  - Nvidia: LUNA16, TCIA Covid19, LiDC (Total: 2124 CTs);
  - Sber: LUNA16 (Total 888 CTs).

СБЕР

Unsupervised pretraining on CT studies

Evaluation protocol:

- Divide the training data randomly into two halves, one – for training, another – for validation;

СБЕР

Unsupervised pretraining on CT studies

Evaluation protocol:

- Divide the training data randomly into two halves, one – for training, another – for validation;
- Use for training 20%, 50% or 100% of training data;

СБЕР

Unsupervised pretraining on CT studies

Evaluation protocol:

- Divide the training data randomly into two halves, one – for training, another – for validation;
- Use for training 20%, 50% or 100% of training data;
- Choose the best on validation;

СБЕР

Unsupervised pretraining on CT studies

Evaluation protocol:

- Divide the training data randomly into two halves, one – for training, another – for validation;
- Use for training 20%, 50% or 100% of training data;
- Choose the best on validation;
- Measure the model on test data;

СБЕР

Unsupervised pretraining on CT studies

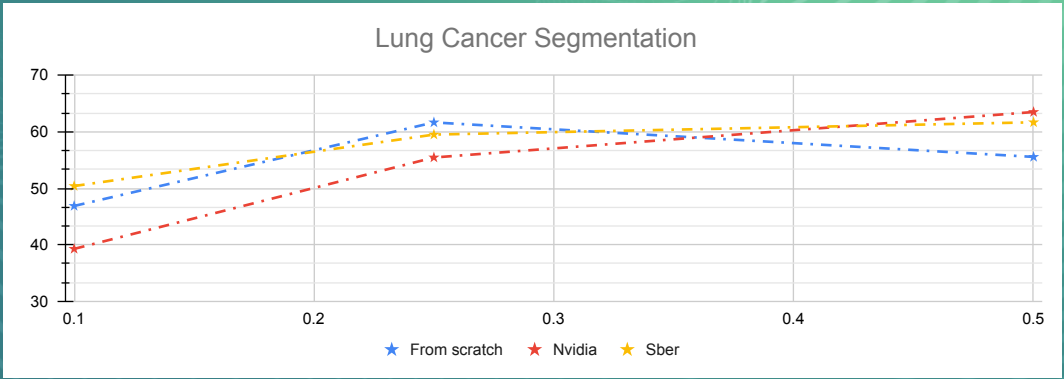
Evaluation protocol:

- Divide the training data randomly into two halves, one – for training, another – for validation;
- Use for training 20%, 50% or 100% of training data;
- Choose the best on validation;
- Measure the model on test data;
- Average between six runs.

СБЕР

Unsupervised pretraining on CT studies

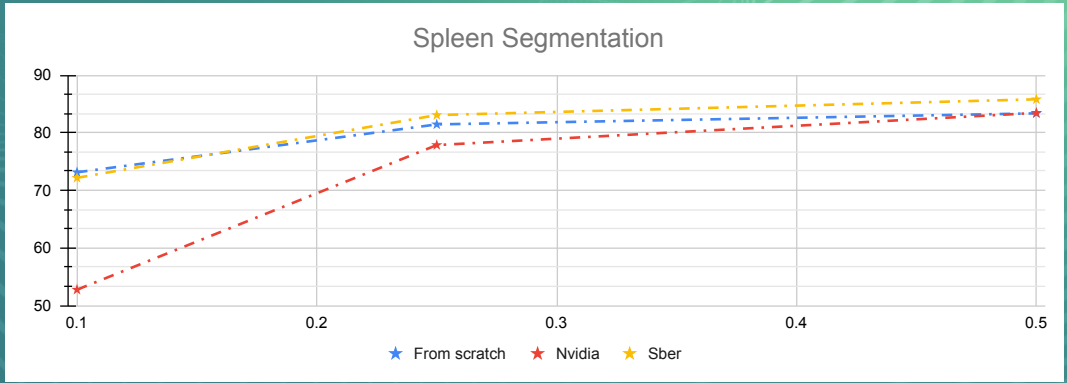
Results: Task06_Lung from Medical Decathlon



СБЕР

Unsupervised pretraining on CT studies

Results Task09_Spleen from Medical Decathlon:



Thank you for your softmax$\left(\frac{QK^T}{\sqrt{d_k}}\right)V$!