Multiview Long-Short Spatial Contrastive Learning for 3D Medical Image Analysis

Gongpeng Cao[1], Yiping Wang[1], Manli Zhang[1], Jing Zhang[1], Guixia Kang[1], Xin Xu[2]

[1]*Beijing University of Posts and Telecommunications, Beijing, China*

[2]*Department of Neurosurgery, General Hospital of PLA, Beijing, China*

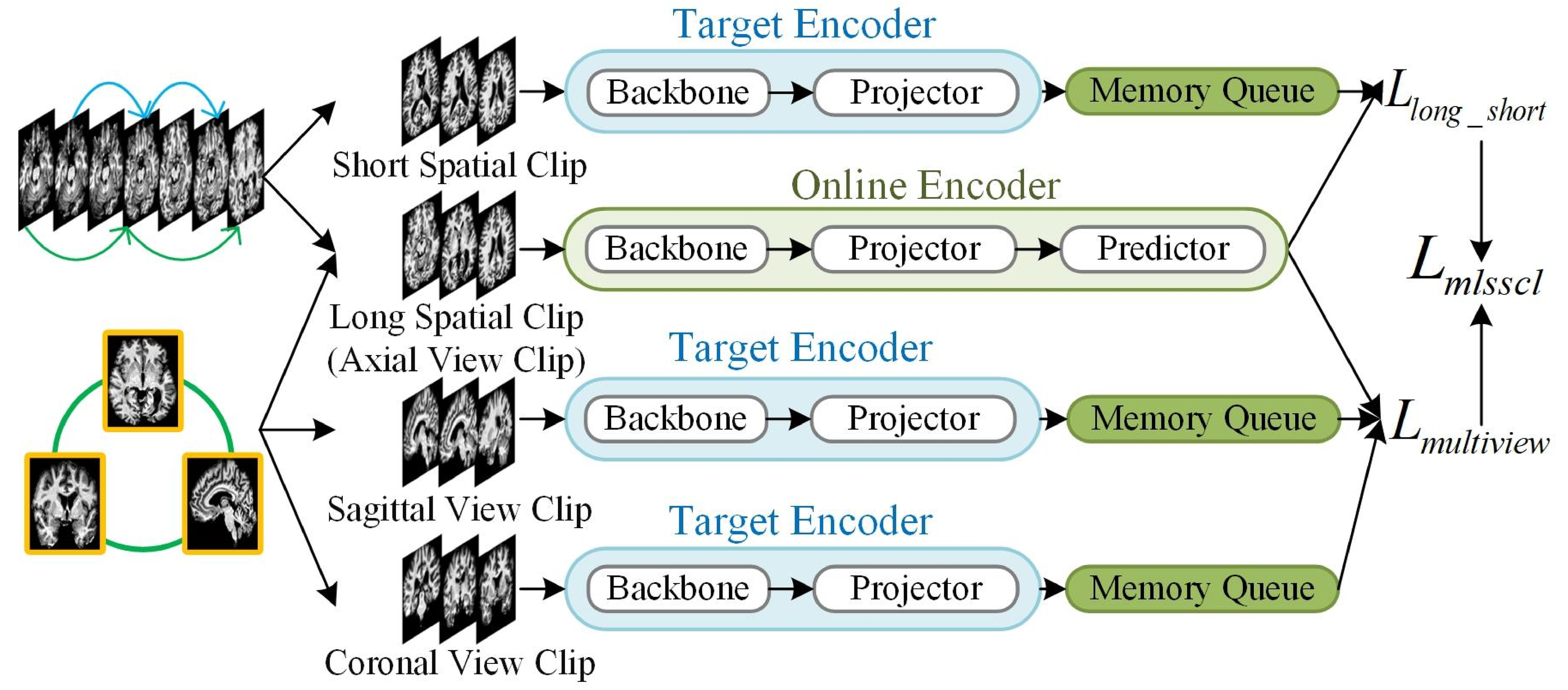
22-27 MAY 2022
ICASSP 2022
SINGAPORE
Poster Number: 1483

1. Introduction

- **Background:**
  - Supervised deep learning heavily depends on large labeled datasets whose construction is often challenging in medical image analysis.
  - Contrastive learning, an effective implementation of self-supervised learning (SSL), is a potential solution to alleviate the strong demand for human-annotations.
- **Motivations:**
  - Existing contrasting strategies ignore the intrinsic structural similarity and local representation, when applied to 3D medical images.
  - The information shared between three views (axial, coronal and sagittal views) can capture the global representation of volumetric medical image.
  - Matching the short spatial clip to long spatial clip forces the model to learn local representation.

2. Method

- **Overall Architecture:** Multiview Long-Short Spatial Contrastive Learning Framework.



- **Long-Short Spatial Contrasting Strategy**
  - We maximize the representation similarity of long spatial clip v_L and short spatial clip v_S to learn local representation:

$$L_{long-short} = L_N(v_S, v_L)$$

- **Multiview Contrasting Strategy**
  - To learn global representation, we need to maximize the mutual information between three views (v_a, v_c, v_s) of volumetric image:

$$\max\{I(v_a; v_c) + I(v_a; v_s) + I(v_c; v_s)\}$$

  - However, the mutual information is difficult to compute for high-dimensional data, we use InfoNCE loss L_N to estimate the lower bound of mutual information. For two views v_1, v_2, :

$$I(v_1; v_2) \geq log(K) - L_N(v_1, v_2)$$

    where K is the number of negative samples.
  - Therefore, we transform the problem of maximizing mutual information between three views into a multiview contrastive learning problem:

$$L_{multiview} = L_N(v_a, v_c) + L_N(v_a, v_s) + L_N(v_c, v_s)$$

3. Experiments

- **Evaluation on MS Lesion Segmentation**

Table 2. The segmentation results of different approaches on the ISBI 2015 longitudinal MS lesion segmentation test set.

Method	DSC†	PPV†	LTPR†	LFPR†
Training-from-Scratch	0.6176	0.8229	0.4451	0.3485
SSL				
Age-Aware [13]	0.6320	0.8103	0.4586	0.3034
BYOL [12]	0.6337	0.7991	0.4675	0.3442
MoCo [10]	0.6369	0.7972	0.4641	0.3092
Model Genesis [8]	0.6434	0.8200	0.4647	0.3082
MS SOTA				
Aslania et al. [3]	0.6114	**0.8992**	0.4103	0.1393
Andermatt et al. [18]	0.6298	0.8446	0.4870	0.2013
Valverde et al. [4]	0.6304	0.7866	0.3669	0.1529
Hu et al. [5]	0.6345	0.8682	0.4787	**0.1299**
MLSSCL	**0.6482**	0.8007	**0.4933**	0.2796

- **Evaluation on AD Classification**

Table 1. Results(mean±std) for AD classification (AD vs. HC) on the ADNI-AD classification test set.

Method	ACC	SEN	SPE	AUC
Training-from-Scratch	0.793 ± 0.011	0.874 ± 0.055	0.711 ± 0.058	0.896 ± 0.006
BYOL [12]	0.809 ± 0.004	0.866 ± 0.032	0.752 ± 0.037	0.886 ± 0.016
MoCo [10]	0.825 ± 0.020	0.886 ± 0.043	0.764 ± 0.060	0.895 ± 0.001
Model Genesis [8]	0.827 ± 0.004	0.911 ± 0.061	0.744 ± 0.061	0.904 ± 0.009
Age-Aware [13]	0.831 ± 0.007	0.882 ± 0.007	0.780 ± 0.012	0.899 ± 0.011
MLSSCL	**0.858 ± 0.013**	**0.911 ± 0.019**	**0.805 ± 0.044**	**0.907 ± 0.012**

- **Ablation to Contrasting Strategies**

Table 3. Ablation to contrasting strategies on AD classification task (mean ± std).

Contrasting Strategy	ACC	AUC
Long-Short	0.823 ± 0.012	0.892 ± 0.014
Multiview	0.833 ± 0.027	0.906 ± 0.024
Multiview & Long-Short	**0.858 ± 0.013**	**0.907 ± 0.012**

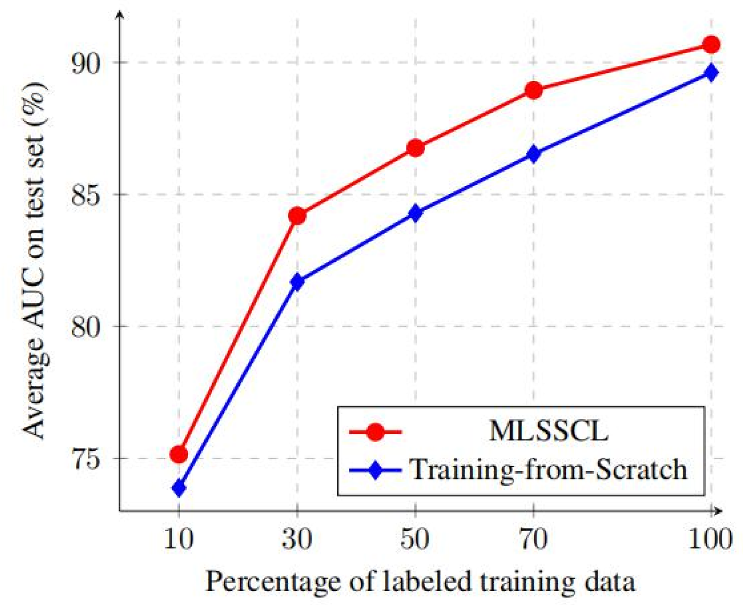


Fig. 2. The AD classification performance of networks trained with different amounts of labeled data.

4. Conclusion

- ✓ We proposed multiview long-short spatial contrastive learning framework for self-supervised 3D visual representation learning, involving multiview contrasting strategy and long-short spatial contrasting strategy.
- ✓ Extensive experiments demonstrate that our framework significantly outperforms learning from scratch and other SSL methods.