Healthcare Technology Innovation Centre

Collaborate · Innovate · Impact

- A joint initiative of -

Indian Institute of Technology Madras

&

Department of Biotechnology
Ministry of Science and Technology

A context based deep learning approach for unbalanced medical image segmentation

Balamurali Murugesan, Kaushik Sarveswaran, Vijaya Raghavan S, Keerthi Ram, Sharath M Shankaranarayana, Mohanasankar Sivaprakasam

Healthcare Technology Innovation Centre
Indian Institute of Technology Madras, India

April 7, 2020

Medical Image segmentation

- Automated medical image segmentation is a preliminary step in many medical procedures.
- Most of the medical images have fewer foreground pixels relative to larger background pixels which introduces class imbalance.
- Deep learning networks (U-Net, GAN) are successful in segmentation tasks, but shows limitations in handling class imbalance.

U-Net and Loss functions

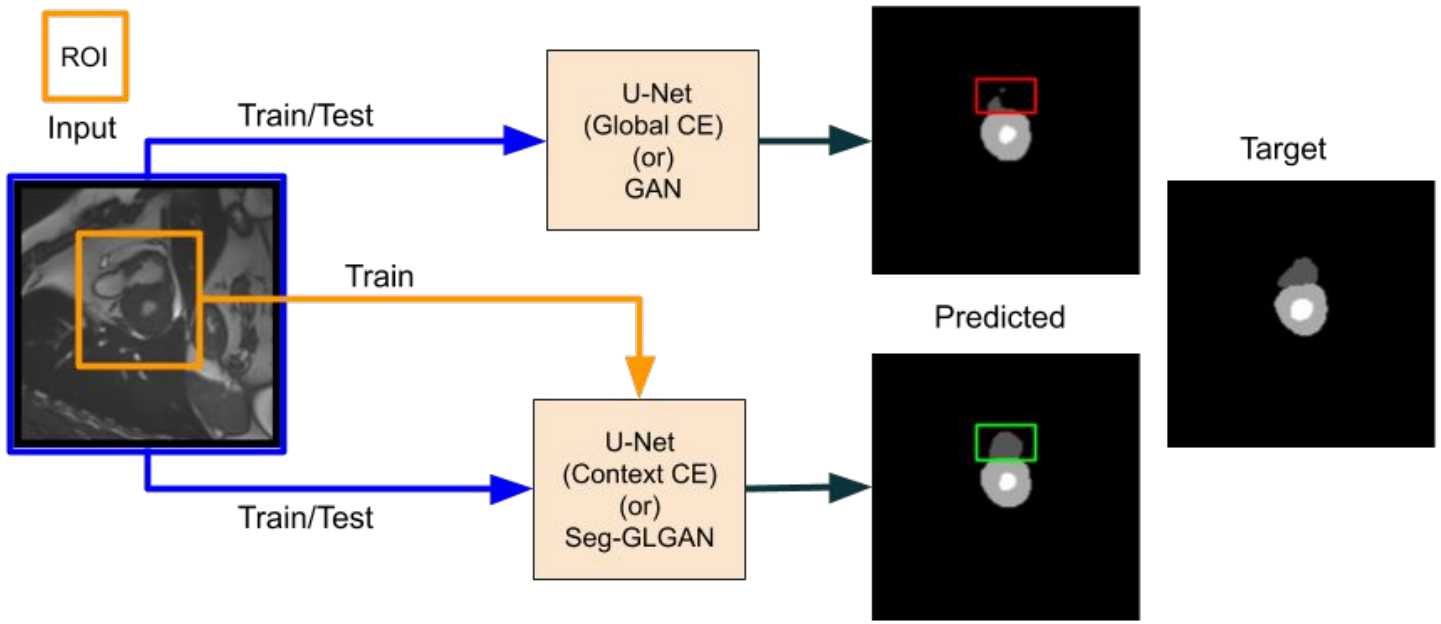
- U-Net has issues in dealing with labels having class imbalance. The issue is mainly due to the usage of cross entropy (CE) loss.
- Assign weights to classes based on inverse of its occurrences.
  - Assigning proper weights will be difficult for varying object sizes.
  - Least frequent class will be affected by noise which may result in unstable training.
- Dice coefficient and Focal loss are common practises to handle class imbalance but faces difficulty with severe class imbalance.

GAN applications and limitations

- Applications: Skin lesion segmentation, Brain tumor segmentation, Joint optic cup and disc segmentation.
- In GAN, the generator is the image to mask prediction network while discriminator is the real/fake classification network.
- The discriminator looks at the entire predicted/target label to classify it as real/fake. Because of this, the discriminator will be insensitive to minor changes.

Motivation

- Lack of local information (ROI)
  - U-Net - CE loss function operates on entire predicted/target label.
  - GAN - Discriminator operates on entire predicted/target label.



Contributions

- Context based CE loss for U-Net - Linear combination of global and local CE loss. The global CE loss takes the entire predicted and target label, while the local CE loss takes the ROI of predicted and target label.
- Seg-GLGAN - A generator and a context discriminator. The context discriminator consists of global and local feature extractor. The global feature extractor takes the entire predicted/target label, while the local feature extractor takes the ROI of predicted/target label.

Context based CE loss

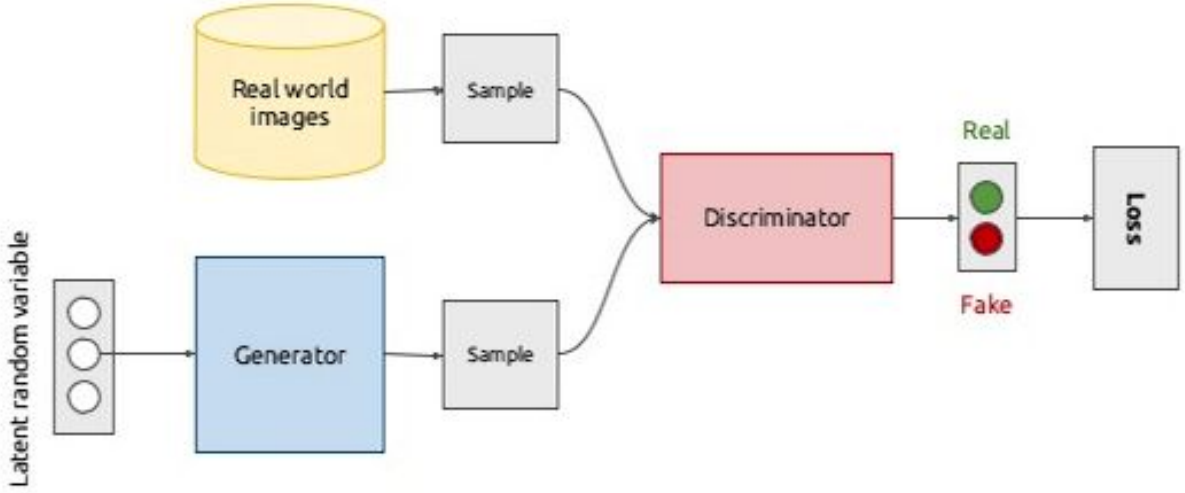
- U-Net global CE loss L_{global} (CE entire target and predicted label).
- Class imbalance: Minimization of L_{global} doesn't mean global minimum.
- U-Net context CE loss $L_{context}$. Linear combination of L_{global} and L_{local}.
- L_{local} - CE between ROI of predicted and target label.
- Minimization of L_{local} ensures better segmentation in the ROI.

$$\mathcal{L}_{context} = \mathcal{L}_{global} + \lambda \mathcal{L}_{local}$$

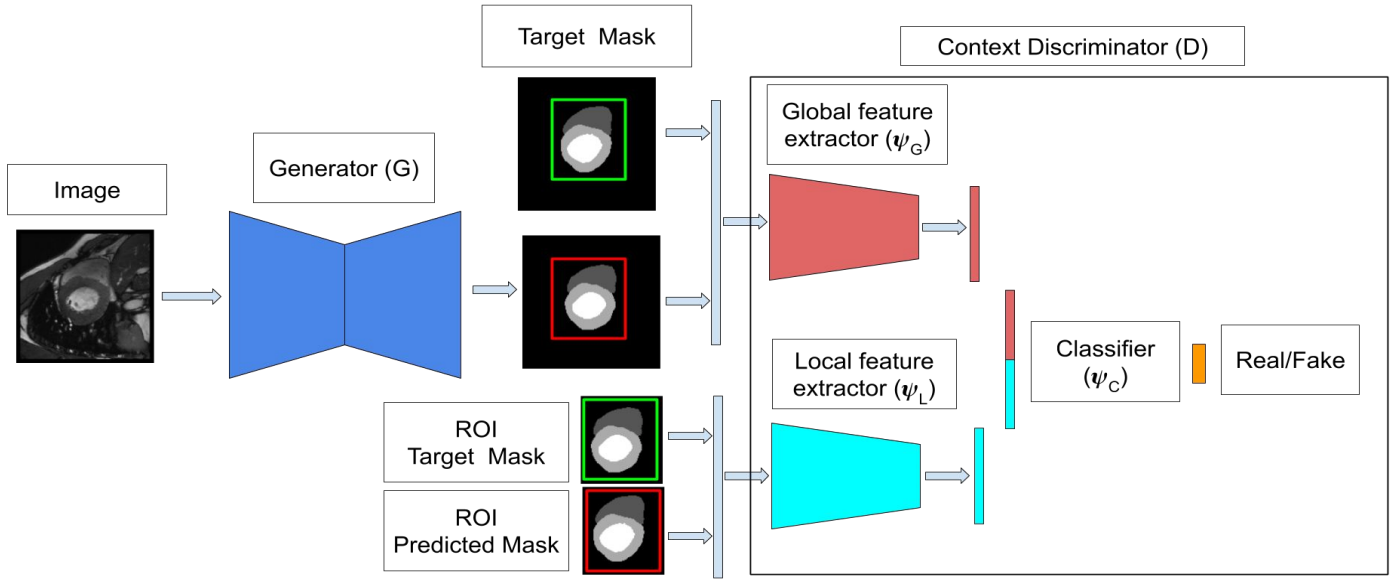
$$\mathcal{L}_{global/local} = -\sum_{\boldsymbol{i} \in \Omega_{EI/ROI}} \sum_{c=1}^{C} y_{\boldsymbol{i}c} \, ln \, \hat{y}_{\boldsymbol{i}c}$$

GAN background

1. GAN consists of generator (G) and discriminator (D) networks.
2. The goal of the G is to map a latent variable to the distribution of the given true data that we are interested in imitating.
3. The D aims to distinguish the true data from the synthesized data.



Proposed architecture - Seg-GLGAN



Generator : U-Net

Context Discriminator :

1) Global feature extractor
   a) 3 Conv layers + 2 FC
   b) Activation function: ReLU.
   c) Output - 64d vector.
2) Local feature extractor
   a) 3 Conv layers + 2 FC
   b) Activation function: ReLU.
   c) Output - 64d vector.
3) Classifier
   a) 128d vector
   b) FC layer (128 x 1)

Loss function

$$L_{Seg-GLGAN} = L_{MCE}(G) + \lambda L_{GAN}(G, D)$$

$$L_{MCE}(G) = -\mathbb{E}_{x,y}\left[\sum_{c=1}^{C} y_c \, ln \, G_c(x)\right]$$

$$L_{GAN}(G, D) = \mathbb{E}_y[\ln D(y)] + \mathbb{E}_x[\ln(1 - D(G(x))]$$

$$D(y) = \Psi_C(\Psi_G(y)||\Psi_L(\Phi(y)))$$

$$D(G(x)) = \Psi_C(\Psi_G(G(x))||\Psi_L(\Phi(G(x))))$$

Dataset and Evaluation metrics

- Prostate MR Image segmentation 2012 (PROMISE12):
  - 50 T2-weighted MR volumes of prostate and its binary segmentation masks. 2D slices from each volume resulted in 1127 and 530 slices for training and validation respectively.
- Automated Cardiac Diagnosis Challenge (ACDC):
  - 150 patient records of cardiac MR volumes with its multi-class segmentation masks. 2D slices from patient records amounted to 1309 and 532 slices for training and validation respectively.
- Dice Similarity Coefficient (DICE) and Hausdorff Distance (HD) are used to evaluate the segmentation performance.

ROI Selection

- Static ROI (StROI): During training, the dimensions of ROI for PROMISE12 and ACDC are empirically set to 50 x 50 and 60 x 60.
- Dynamic ROI (DyROI): The dimensions of the ROI is adapted dynamically based on the size of the object.
- In the case of Seg-GLGAN, the fully connected network in the context discriminator is replaced with Global Average Pooling to handle ROI with different dimensions.

Experiments and results

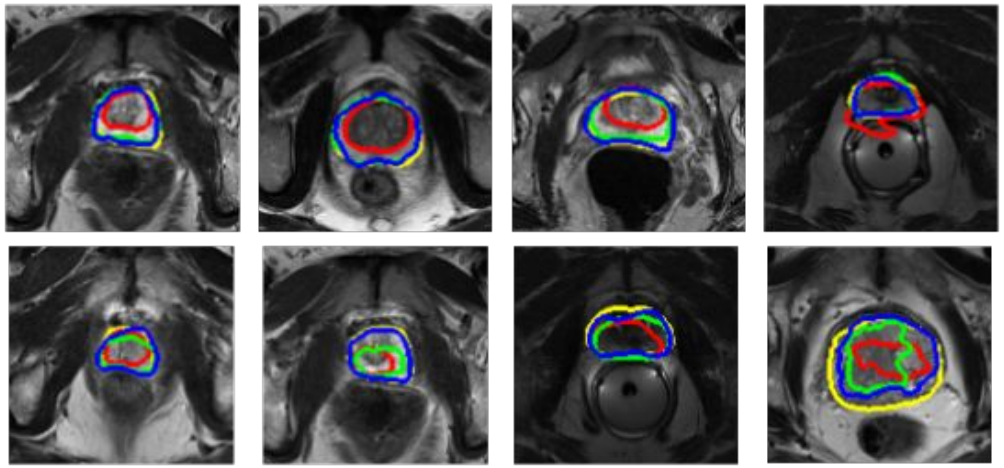
Quantitative comparison

		PROMISE 12		ACDC	
		Dice	HD	Dice	HD
U-Net	$\mathcal{L}_{CE}$	0.7964	9.11	0.822	5.69
	$\mathcal{L}_{WCE}$	0.8374	7.82	0.8385	5.31
	$\mathcal{L}_{DICE}$	0.8166	8.25	-	-
	$\mathcal{L}_{FOCAL}$	0.8269	8.00	0.8245	5.48
	$\mathcal{L}_{CONTEXT}(StROI)^*$	**0.861**	**6.91**	**0.85**	**4.55**
	$\mathcal{L}_{CONTEXT}(DyROI)^*$	0.861	7.30	0.8402	5.30
GAN	GAN	0.7877	8.76	0.833	4.70
	Seg-GLGAN(StROI)*	0.8019	8.40	0.835	4.69
	Seg-GLGAN(DyROI)*	**0.86**	**7.01**	**0.856**	**4.38**

Experiments and results

Qualitative comparison

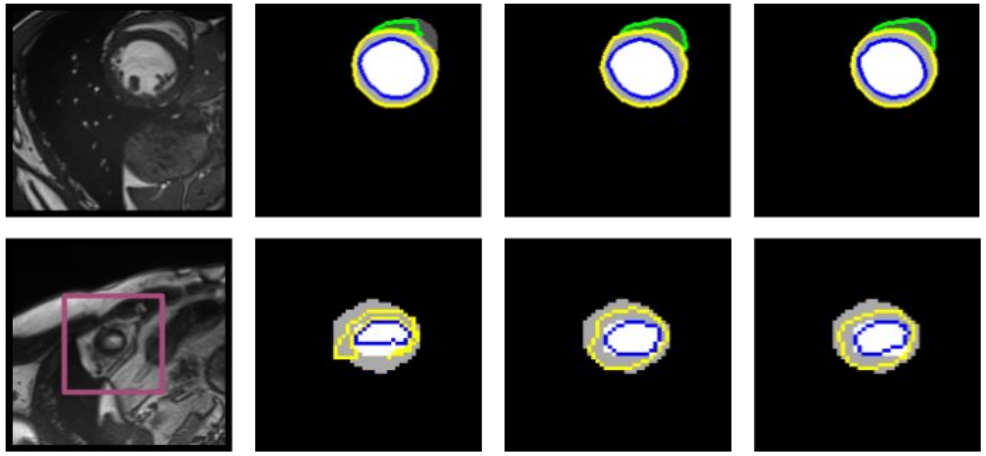
PROMISE12



Top Row: Red - U-Net with CE, Green - U-Net with Context CE(StROI), Blue - U-Net with Context CE(DyROI) Yellow - Target.

Bottom Row: Red - GAN, Green - Seg-GLGAN(StROI), Blue - Seg-GLGAN(DyROI), Yellow - Target.

ACDC



Green(RV), Yellow(Myo) and Blue(LV).

Top Row - Input, U-Net with CE, U-Net with Context CE(StROI), U-Net with Context CE(DyROI).

Bottom Row -Input, GAN, Seg-GLGAN(StROI), Seg-GLGAN(DyROI). (The ROI is zoomed for better visualization)

Conclusion and Future work

- We have proposed a context based cross entropy loss for U-Net and a GAN based network, Seg-GLGAN to alleviate class imbalance problem in segmentation by considering both global and local context.

- The proposed methods can be successfully applied to regular object segmentation tasks because of the ROI. However, the idea of ROI limits the method's extension to sparse segmentation tasks (vessel extraction for retinal images). The future research will be to address this limitation.

Thank you

Paper

https://arxiv.org/abs/2001.02387

Code

https://github.com/Bala93/Context-aware-segmentation

Contact

balamurali@htic.iitm.ac.in