Robust Image Segmentation Quality Assessment

Leixin Zhou, Wenxiang Deng, Xiaodong Wu

Department of Electrical and Computer Engineering
University of Iowa

July, 2020

Introduction

- State-of-the-art segmentation quality assessment method is deep learning (DL) [Robinson et al., 2018]
  - A regression DL network.
  - Input: original image and segmentation to be assessed.
  - Output: dice prediction.

- DL models are fragile to many factors, e.g. domain shift [Patel et al., 2015], adversarial noise [Goodfellow et al., 2015], low image quality
  - DL network may find some unrobust features.
  - The fragility can be demonstrated with adversarial attacks.

Method: "Filter" the features more related to segmentation quality

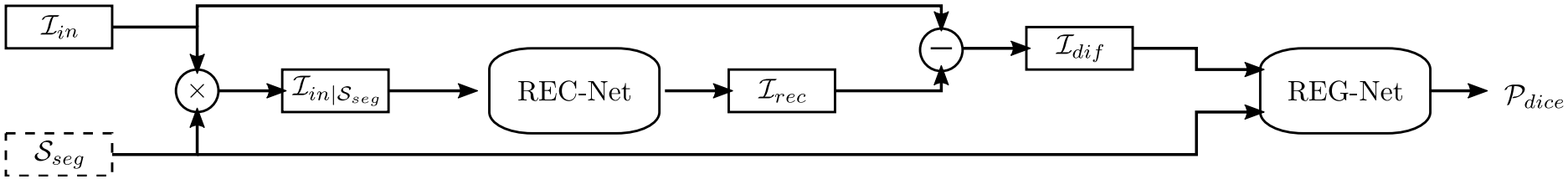


Figure 1: The work flow of proposed segmentation quality assessment method.

- State-of-the-art method: $\mathcal{I}_{in}$ includes too rich information for the regression net (REG-Net) to explore, $\mathcal{P}_{dice} = \text{REG-Net}(\mathcal{I}_{in}, \mathcal{S}_{seg})$
- Proposed method: Replace $\mathcal{I}_{in}$ with more segmentation quality related feature image $\mathcal{I}_{dif}$, defined as

$$\mathcal{I}_{dif} = \mathcal{I}_{in} - \text{REC-Net}\big(\mathcal{I}_{in} \odot (1 - \mathcal{S}_{seg}) \big)$$

  - Reconstruction network (REC-Net) is trained with original image and its ground truth segmentation only.
  - Reconstruction and then the difference image is dependent on the segmentation.

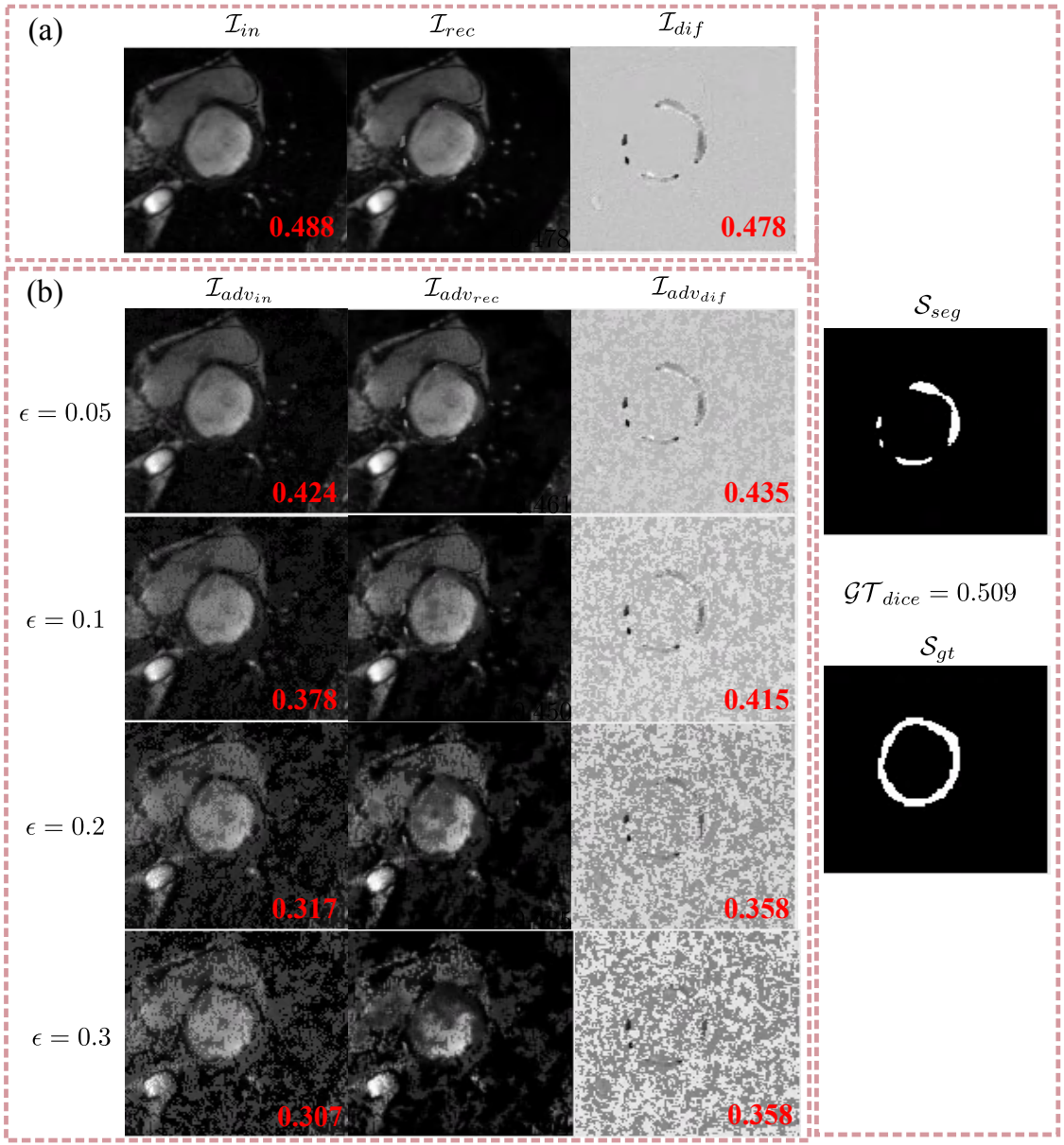
Experiments

- Data: Automated Cardiac Diagnosis Challenge (ACDC) MICCAI challenge 2017. Segmentation of left-ventricular myocardium (LVM) was considered.
- Segmentation simulation: U-nets [Ronneberger et al., 2015] with difference depths, filter number, and training epochs. The finale segmentation pool obeys uniform distribution with repect to dice.
- Adversarial attack method: fast gradient sign [Kurakin et al., 2016].

Method	$\epsilon = 0$	$\epsilon = 0.05$	$\epsilon = 0.1$	$\epsilon = 0.2$	$\epsilon = 0.3$
Robinson *et al.*	0.04±0.05	0.08±0.06	0.11±0.07	0.14±0.08	0.16±0.09
proposed	0.04±0.05	0.07±0.06	0.09±0.06	0.09±0.07	0.12±0.09

Table 1: Mean absolute errors of dice prediction under different levels of adversarial attack.

Visualization and Future Work



- The proposed method is more robust than state-of-the-art.
- To be tested with more applications.
- To be tested with more adversarial attack methods.