Technische Universiteit Eindhoven University of Technology

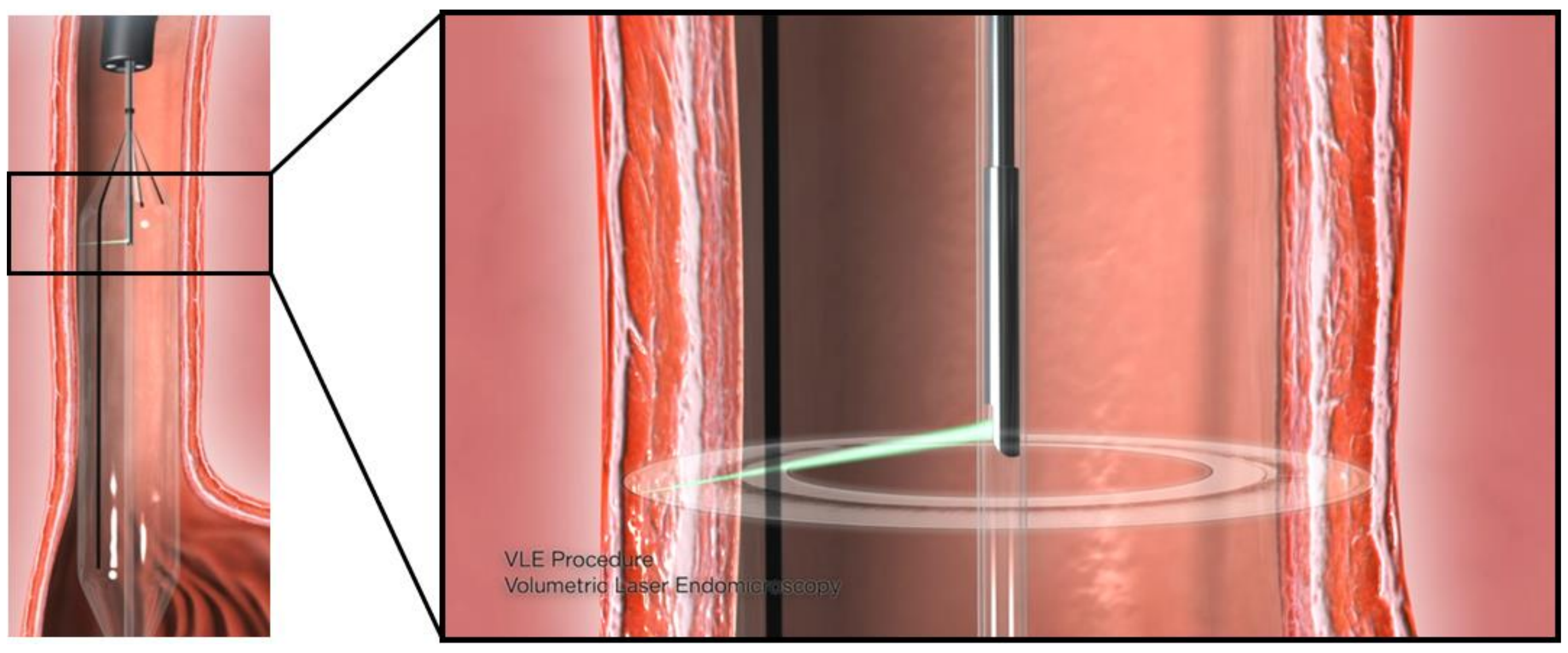
TISSUE SEGMENTATION IN VOLUMETRIC LASER ENDOMICROSCOPY DATA USING FUSIONNET AND A DOMAIN-SPECIFIC LOSS FUNCTION

van der Putten J.[1], van der Sommen F.[1], Struyvenberg M.[2], de Groof J.[2], Curvers W.L.[3], Schoon E.J.[3], Bergman J.J.G.H.M.[2], de With P.H.N.[1]

[1]Eindhoven University of Technology, the Netherlands
[2]Academic Medical Center, Amsterdam, the Netherlands
[3]Catharina Hospital Eindhoven, the Netherlands

Abstract

Volumetric Laser Endomicroscopy (VLE) is a promising balloon based imaging technique for detecting early neoplasia in Barrett's Esophagus. Especially Computer Aided Detection (CAD) techniques show great promise compared to medical doctors, who cannot reliably find disease patterns in the VLE signal. However, an essential pre-processing step for the CAD system is tissue segmentation. At present, tissue segmentation is selected manually and is therefore not scalable for full VLE scans of 1,200 × 4,096 × 2,048 pixels. Furthermore, the current CAD methods cannot use the VLE scans to their full potential as only a small section is selected while an automated system can delineate the entire image. This paper explores the possibility of automatically segmenting relevant tissue for VLE scans using a convolutional neural network.



1. Background

- Volumetric Laser Endoscopy (VLE) imaging systems have enabled the **identification of early cancerous lesions** up to 3 mm deep within the mucosa in Barrett's Esophagus (BE).
- VLE scans are complex to interpret and consists of a large amount data (typically 1200 slices of 4096 × 1024 pixels).
- **Computer vision techniques** provide a helpful tool for quantitative analysis of VLE scans for Barrett's cancer detection but require manual region of interest selection.

2. Goals of the study

- Investigate the feasibility of a tissue segmentation algorithm in volumetric laser endomicroscopy data.

3. Methods

- Transform reference ground truth (Fig. 1b) into weighted ground truth (Fig. 1c).
- We use **FusionNet**, an upgraded fully residual version of the more widely known U-Net architecture, to segment the scans. (Fig. 2 and 3)
- **Soft dice score** was implemented to facilitate usage of the weighted ground truth.

$$DSC = \frac{2(y \cdot \hat{y})}{(y + \hat{y})}$$

- A **balloon distance** metric is introduced to assess quality of segmentation.

$$Balloon\ distance = 1 - \min\left(\frac{\sum_{n=1}^{N} APD_n}{1000}, 1\right)$$

- 4-fold cross-validation was used to evaluate performance

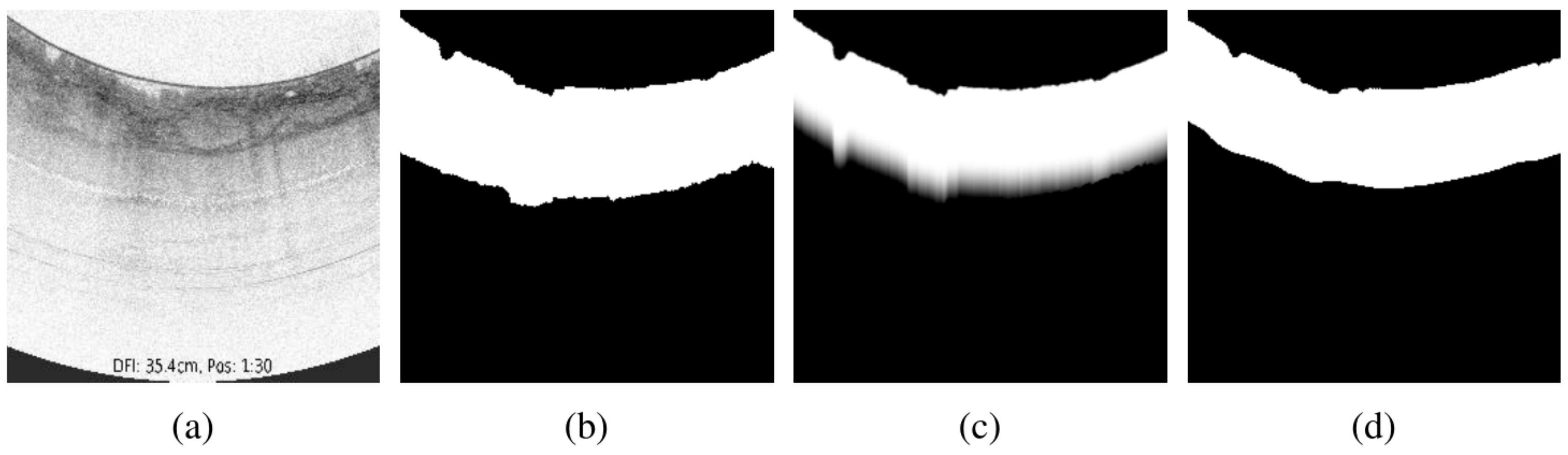


Figure 1: (a) VLE image. (b) Reference ground truth. (c) Weighted ground truth. (d) Prediction output of the CNN, using the weighted ground truth to train the FusionNet.

4. Experimental results

	Basic model		Weighted model		Assessor 2		Assessor 3	
	DSC	BD	DSC	BD	DSC	BD	DSC	BD
Assessor 1	0.95	0.87	0.97	0.88	0.96	0.83	0.96	0.84
Assessor 2	0.95	0.82	0.97	0.83	-	-	0.96	0.80
Assessor 3	0.95	0.84	0.97	0.85	-	-	-	-

Table 1: Left: DSC and balloon distance (BD) comparison of the different assessors and the weighted ground truth and basic ground truth predictions, obtained with their respective model. Right: different assessor annotations scored against each other.

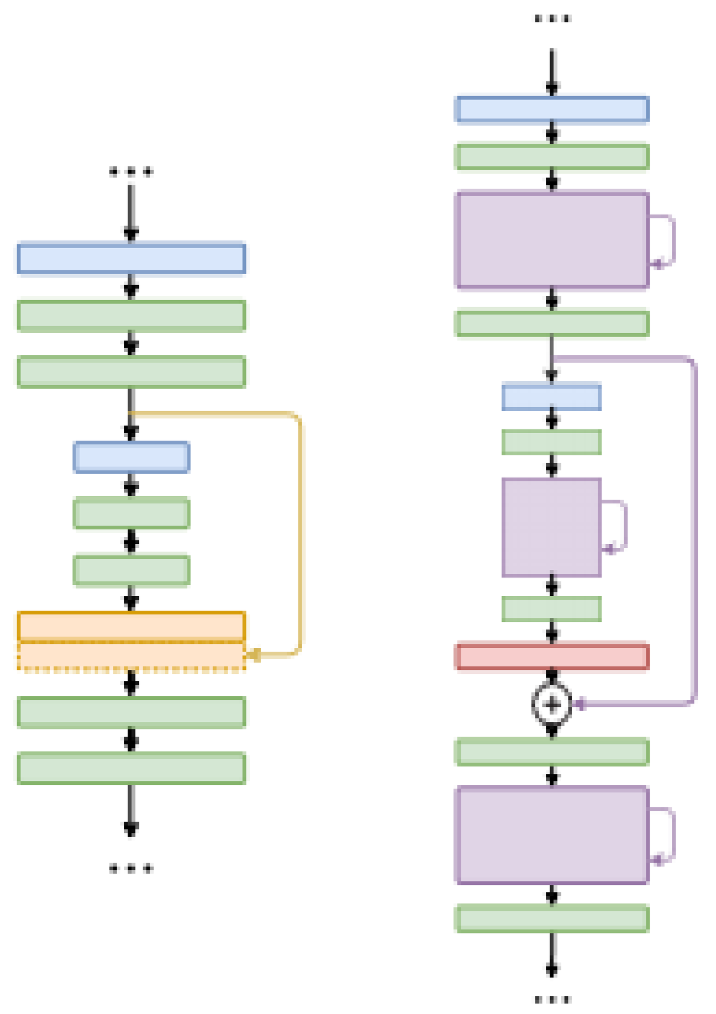


Figure 2: difference between U-net and FusionNet.

5. Conclusions

- First segmentation algorithm for automatically finding the region of interest in VLE scans.
- Weighted ground truth model improves on reference model.
- Resulting segmentation within inter-observer variability of 3 expert assessors. Making our predictions interchangeable with expert annotations.

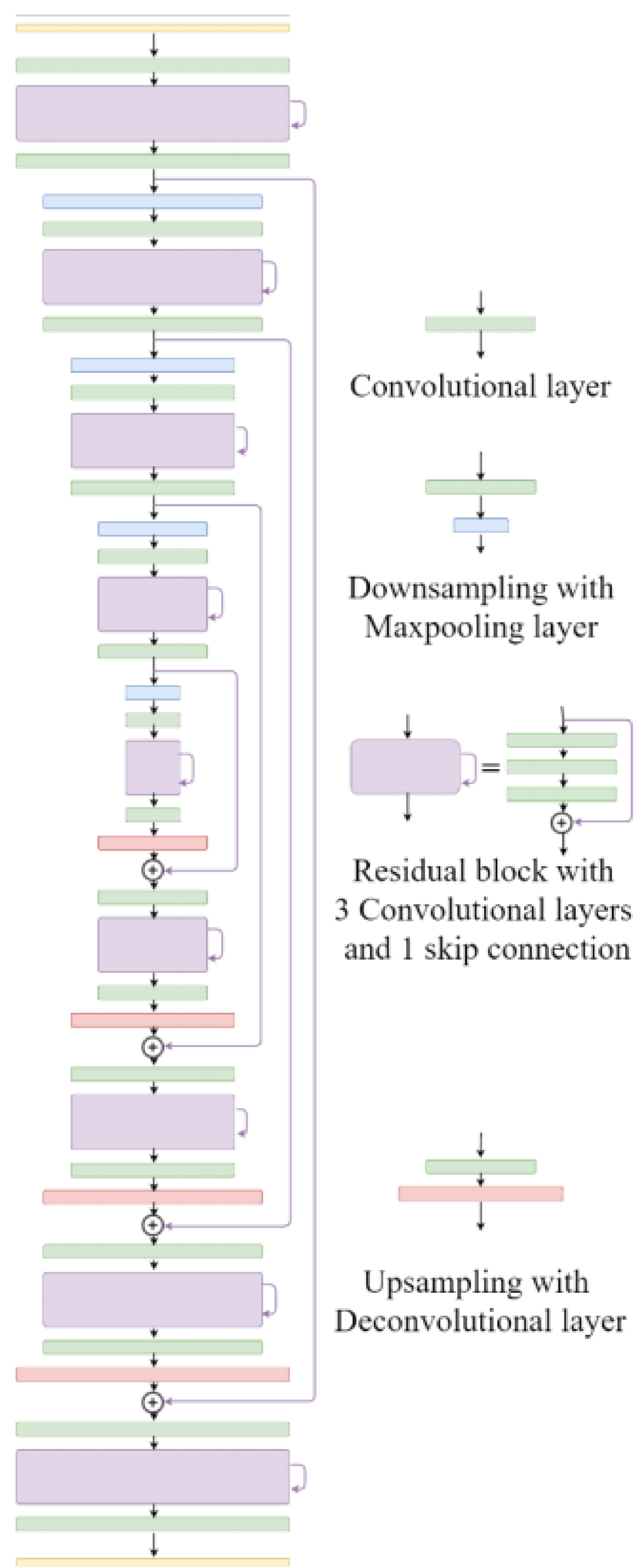


Figure 3: FusionNet architecture overview.