Attention-based Image Segmentation

Author: Carlos Henrique Caloete Pena (chcp@cin.ufpe.br)
Adviser: Tsang Ing Ren (tir@cin.ufpe.br)
Area: Machine Learning

Recife 2019.
1 Resumo


Propomos um novo *framework* de fusão de segmentações, este *framework* é dividido em três partes, um conjunto de segmentadores (PoolNet), uma rede neural profunda de atenção para as segmentações (Expert), e a função de agregação (*F*), nesta função, buscamos trazer tanto funções não treináveis, como uma média aritmética ponderada, quanto funções treináveis, como uma rede neural.
2 Abstract

Ensemble Learning is widely used mainly in problems involving classification and regression tasks, such as [1] and [2], however, still needs some research to use Ensemble Learning with image segmentation. This project proposes a framework to use the concepts of Ensemble Learning with a deep neural attention network to improve results on image segmentation tasks.

We propose a new framework based on image segmentation fusion to outperform single models. This framework is divided into three parts; a set of segmentation models, the PoolNet, an attention model, the Expert, and an aggregation function, $F$. That aggregation function could be non-trainable, such as weighted arithmetic mean or trainable such as a deep neural network.
3 Introduction

Since the creation of U-Net [3] a fully convolutional network, the field of image segmentation had a significant advance, especially in microscopic images, due to a lack of images and moreover lack of hand-mand segmentation done by a specialist. The U-Net address this issue using data augmentation such as elastic transformation, and with the skip connection.

In this work, we focus on microscopic biological images [4] as shown in Figure 1, in those images, there are two principal obstacles; a presence of multiple cells stuck together, and presence of a substance that looks like a cell but they are not.

![Figure 1: Example of microscopic biological images](image)

We propose a new framework of image segmentation fusion to outperform single models. This framework is divided into three parts; a set of segmentation models such as UNetResnet, the PoolNet, an attention model, the Expert, and an aggregation function, $F$. We exploit two variants for the $F$ function, a non-trainable and trainable variant.
4 Objectives

The main propose of this project is to propose a framework for image segmentation fusion that overcomes single and static results.

The following are specific objectives of this project:

- Train at least five deep networks for image segmentation;
- Ensemble these models with a non-trainable aggregation function; such as mean, max, min rule.
- Calculate Oracle and Single Best with those models
- Train the proposed framework with at least four aggregation functions
- Compare the single model, non-trainable ensemble and our proposed framework with at least six metrics
5 Methodology

As the first objective, we select and train several deep neural networks for image segmentation, some architectures like UnetResnet and UnetVgg archive great results in competitions such as 2018 Data Science Bowl [5].

After train those models we combine them with some non-trainable aggregation function such as Majority Vote, Max Rule, Min Rule, and Median Rule, more details can be seen on [6].

We defined a framework to combine results from a group of segmentation models into a new and better segmentation. Figure 2 shows the top flowchart of the project. First, giving an image $x$, from the $SRC$ set it to pass thought every network inside PoolNet and obtained its likelihood map for each network. The $f$ function $R^{2N\times H\times W} \rightarrow R^{H\times W}$ combined those likelihood maps and weights that were obtained from a segmentation network Expert $E$, this combination is the new likelihood map proposed in this article.

In this project, we will focus our evaluation with the Panoptic Quality (PQ) metric [7] because PQ metric takes into consideration both the class and the instance label. Finally, we will compare our new framework with a single model, some non-trainable aggregation fusion, and the Oracle.
## 6 Schedule

<table>
<thead>
<tr>
<th>Activity</th>
<th>August</th>
<th>September</th>
<th>October</th>
<th>November</th>
<th>December</th>
</tr>
</thead>
<tbody>
<tr>
<td>Literature Review</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Training Deep Neural Networks</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Results Analysis</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Writing</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Presentation preparation</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>
References


7 Possible Examiner

These teachers are possible examiner for this work:

- George Darmiton da Cunha Cavalcanti (gdcc@cin.ufpe.br)
- Carlos Alexandre Barros de Mello (cabm@cin.ufpe.br)