

## How do robots see?

An introduction to semantic segmentation in images

Alano Martins Pinto Computer vision research leader

email: alano\_pinto@atlantico.com.br fone: +5585999509376 Paulo Bruno Serafim Software Development Engineer

email: paulo.bruno@hp.com phone: +5585999972923



Atlântico

# I. Segmentation





### **Segmentation**

Segmentation is an important part of image analysis. It refers to the process of partitioning an image into multiple segments.

More precisely, image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain visual characteristics.

The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyse.



Topology

(Watershed)

### **Methods**



Clustering

(C-Means)







Threshold - Histogram (Otsu)

From: https://scikit-image.org/docs/dev/auto\_examples/segmentation/plot\_watershed.html https://slideplayer.com/slide/5031931/ http://www.labbookpages.co.uk/software/imgProc/otsuThreshold.html



# **Autoencoders**

Neural Networks



## Convolution Kernel $\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$

## **Neural Network**

#### Brief introduction





#### CNN

During the forward pass, each filter is convolved across the width and height of the input volume, computing the dot product between the entries of the filter and the input and producing a 2-dimensional activation map of that filter.



#### **Max-Pooling**

Max pooling is a sample-based discretization process. The objective is to down-sample an input representation (image, hidden-layer output matrix, etc.), reducing its dimensionality and allowing for assumptions to be made about features contained in the sub-regions binned.



#### Softmax

Softmax function, also known as softargmax or normalized exponential function, is a function that takes as input a vector of K real numbers, and normalizes it into a probability distribution consisting of K probabilities proportional to the exponentials of the input numbers.

From: https://towardsdatascience.com/generating-images-with-autoencoders-77fd3a8dd368 https://computersciencewiki.org/index.php/Max-pooling\_/\_Pooling https://youtu.be/lvNdl7yg4Pg



#### **Autoencoders**

An autoencoder is a neural network that is trained to attempt to copy its input to its output. Internally, it has a hidden layer h that describes a code used to represent the input.

The network may be viewed as consisting of two parts: an encoder function: **h** = **f(x)** and a decoder that produces a reconstruction: **r** = **g(h)** 



# CODE





### **Semantic Segmentation**

Semantic segmentation is a natural step in the progression from coarse to fine inference: The origin could be located at classification, which consists of making a prediction for a whole input.

The next step is localization / detection, which provide not only the classes but also additional information regarding the spatial location of those classes.

Finally, semantic segmentation achieves fine-grained inference by making dense predictions inferring labels for every pixel, so that each pixel is labeled with the class of its enclosing object ore region.

-----

Ð,

~ \*\*

Class based

## **Segmentation types**



Original



**Instance based** 



**Context based** 

From: http://www.cs.toronto.edu/~tingwuwang/semantic\_segmentation.pdf



### SegNet

SegNet, is designed to be an efficient architecture for pixel-wise semantic segmentation. It is primarily motivated by road scene understanding applications which require the ability to model appearance (road, building), shape (cars, pedestrians) and understand the spatial-relationship (context) between different classes such as road and sidewalk



# CODE







## Thank you!

#### SEDE FORTALEZA

Salinas Shopping Av. Washington Soares, 909, Lojas 42, 43, 44 e 45 - Edson Queiroz CEP: 60811-341 Fortaleza - CE

Fone: +55 (85) 3216-7800 Fax: +55 (85) 3216-7864

#### UNIDADE SÃO PAULO

Alphaville Comercial Alameda Araguaia, 933 Alphaville CEP: 04752-005 - São Paulo - SP

Fone/Fax: +55 (11) 3506-3700

