

***Inteligência Artificial
e Computacional***

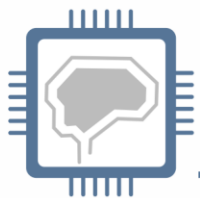
ELT578

ANÁLISE DE IMAGENS E VISÃO COMPUTACIONAL

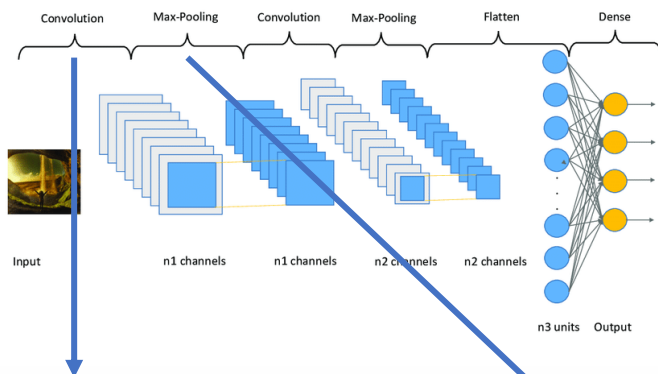
Aula 4:

Técnicas Avançadas em Visão Computacional

Conteudista: M.Sc. Talita E. Z. Santana



Classificação de Imagens



Input Volume (+pad 1) (7x7x3)

$x[:, :, 0]$	0	0	0	0	0	0	0
$x[:, :, 1]$	0	0	0	1	0	2	0
$x[:, :, 2]$	0	1	0	2	0	1	0
$x[:, :, 0]$	0	1	0	2	2	0	0
$x[:, :, 1]$	0	2	0	0	2	0	0
$x[:, :, 2]$	0	2	1	2	2	0	0
$x[:, :, 0]$	0	0	0	0	0	0	0
$x[:, :, 1]$	0	0	0	0	0	0	0
$x[:, :, 2]$	0	2	1	2	1	1	0
$x[:, :, 0]$	0	2	1	2	0	1	0
$x[:, :, 1]$	0	0	2	1	0	1	0
$x[:, :, 2]$	0	1	2	2	2	2	0
$x[:, :, 0]$	0	0	1	2	0	1	0
$x[:, :, 1]$	0	0	0	0	0	0	0
$x[:, :, 2]$	0	0	0	0	0	0	0
$x[:, :, 0]$	0	0	0	0	0	0	0
$x[:, :, 1]$	0	2	1	1	2	0	0
$x[:, :, 2]$	0	1	0	0	1	0	0
$x[:, :, 0]$	0	0	1	0	0	0	0
$x[:, :, 1]$	0	1	0	2	1	0	0
$x[:, :, 2]$	0	2	2	1	1	1	0
$x[:, :, 0]$	0	0	0	0	0	0	0

Filter W0 (3x3x3)

$w0[:, :, 0]$	-1	0	1
$w0[:, :, 1]$	0	0	1
$w0[:, :, 2]$	1	-1	1
$w0[:, :, 0]$	-1	0	1
$w0[:, :, 1]$	-1	0	1
$w0[:, :, 2]$	1	-1	1
$w0[:, :, 0]$	0	1	0
$w0[:, :, 1]$	-1	1	1
$w0[:, :, 2]$	0	-1	0
$w0[:, :, 0]$	0	-1	0
$w0[:, :, 1]$	0	-1	0
$w0[:, :, 2]$	0	-1	0

Bias b0 (1x1x1)

$b0[:, :, 0]$	1
---------------	---

Filter W1 (3x3x3)

$w1[:, :, 0]$	0	1	-1
$w1[:, :, 1]$	0	-1	0
$w1[:, :, 2]$	0	-1	1
$w1[:, :, 0]$	-1	0	0
$w1[:, :, 1]$	-1	1	-1
$w1[:, :, 2]$	1	-1	0
$w1[:, :, 0]$	1	-1	0
$w1[:, :, 1]$	-1	1	-1
$w1[:, :, 2]$	0	-1	-1
$w1[:, :, 0]$	1	0	0
$w1[:, :, 1]$	1	0	0
$w1[:, :, 2]$	1	0	0

Bias b1 (1x1x1)

$b1[:, :, 0]$	0
---------------	---

toggle movement

Output Volume (3x3x2)

$o[:, :, 0]$	2	3	3
$o[:, :, 1]$	3	7	3
$o[:, :, 2]$	8	10	-3
$o[:, :, 0]$	-8	-8	-3
$o[:, :, 1]$	-3	1	0
$o[:, :, 2]$	-3	-8	-5

Rectified feature map

1	4	2	7
2	6	8	5
3	4	0	7
1	2	3	1

max pooling with 2x2 filters and stride 2

Pooled feature map

6	8
4	7

Max(3, 4, 1, 2) = 4

roses



dandelion



tulips



sunflowers



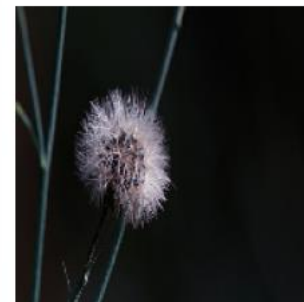
dandelion



roses



dandelion

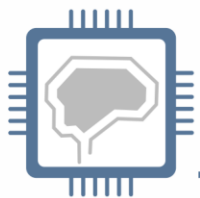


roses



tulips





Filtragem por correlação cruzada vs convulação

Correlação cruzada

$$g(i, j) = (u, v) \otimes f(i, j) = \sum_{u=-k}^{+k} \sum_{v=-k}^{+k} [(u, v) \cdot f(i + u, j + v)]$$

a	b	c
d	e	f
g	h	i

$h(u, v)$

Filtro

$u, v \in [-1, 1]$

a (-1,-1)	b (-1,0)	c (-1,1)
d (0,-1)	e (0,0)	f (0,1)
g (1,-1)	h (1,0)	i (1,1)

$u, v \in [-1, 1]$

$h(-1, -1) = a$

$h(-1, 0) = b$

$h(-1, 1) = c$

\vdots

\vdots

\vdots

$h(1, 1) = i$

Convulação

$$g(i, j) = (u, v) * f(i, j) = \sum_{u=-k}^{+k} \sum_{v=-k}^{+k} [(u, v) \cdot f(i - u, j - v)]$$

A	B	C
D	E	F
G	H	I

$f(x, y)$

Imagem

A (i-1, j-1)	B (i-1, j)	C (i-1, j+1)
D (i, j-1)	E (i, j)	F (i, j+1)
G (i+1, j-1)	H (i+1, j)	I (i+1, j+1)

$f(i, j)$

$f(i-1, j-1) = A$

$f(i-1, j) = B$

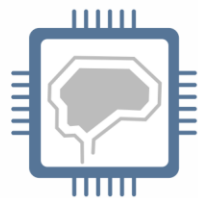
$f(i-1, j+1) = C$

\vdots

\vdots

\vdots

$f(i+1, j+1) = I$



$u, v \in [-1, 1]$

$h(-1, -1) = a$

$h(-1, 0) = b$

$h(-1, 1) = c$

.

.

.

$h(1, 1) = i$

a (-1,-1)	b (-1,0)	c (-1,1)
d (0,-1)	e (0,0)	f (0,1)
g (1,-1)	h (1,0)	i (1,1)

$h(u, v)$

A (i-1,j-1)	B (i-1,j)	C (i-1,j+1)
D (i,j-1)	E (i,j)	F (i,j+1)
G (i+1,j-1)	H (i+1,j)	I (i+1,j+1)

$f(i, j)$

$f(i-1, j-1) = A$

$f(i-1, j) = B$

$f(i-1, j+1) = C$

.

.

.

$f(i+1, j+1) = I$

$$g(i, j) = \sum_{u=-1}^{+1} \sum_{v=-1}^{+1} [h(u, v) \cdot f(i+u, j+v)]$$

$u = -1$

* $v = -1 \rightarrow h(-1, -1) \cdot f(i+1, j+1) = a \cdot I$

* $v = 0 \rightarrow h(-1, 0) \cdot f(i+1, j) = b \cdot H$

* $v = 1 \rightarrow h(-1, 1) \cdot f(i+1, j-1) = c \cdot G$

$u = 0$

* $v = -1 \rightarrow h(0, -1) \cdot f(i, j+1) = d \cdot F$

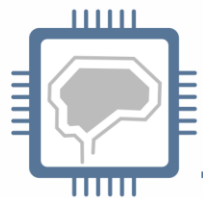
* $v = 0 \rightarrow h(0, 0) \cdot f(i, j) = e \cdot E$

* $v = 1 \rightarrow h(0, 1) \cdot f(i, j-1) = f \cdot D$

$u = 1$

...

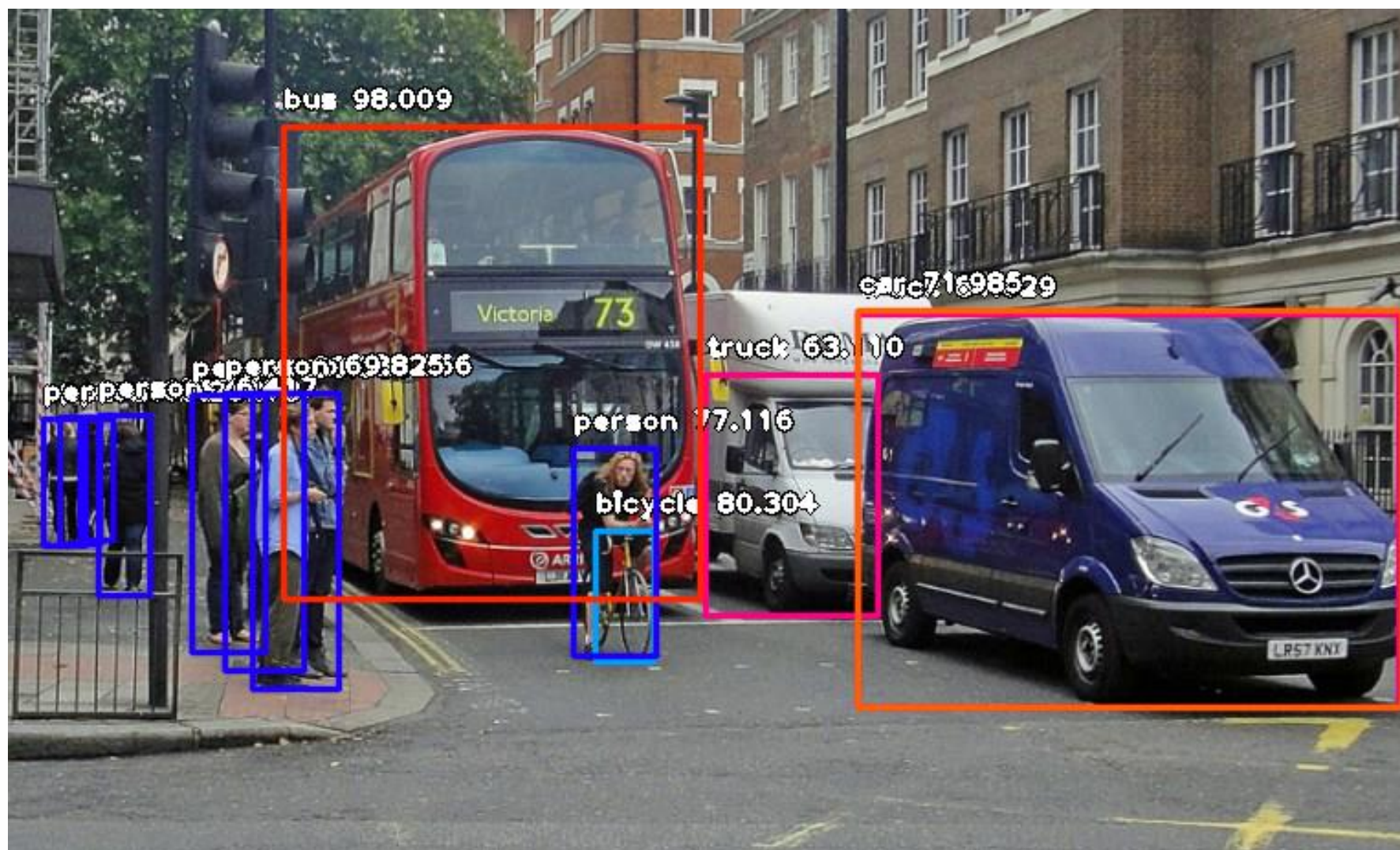
$$g(i, j) = a \cdot I + b \cdot H + c \cdot G + d \cdot F + \dots + i \cdot A$$

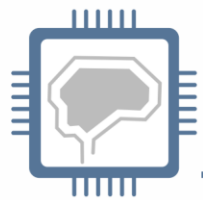


Detecção de Objetos



ceadUFV
Coordenadoria de
Educação Aberta e a Distância





Rastreamento de Objetos

Multiple Object Tracking

Distancing

Warning

Safe Social Distancing — 1.5 m

In-Store Consumer Behavior

Customer

Zone 1

Zone 2

People in the store — 12

Pedestrian Detection

Pedestrian

Face RE-ID

Recognized: Carter Arcand

Tracking by detection

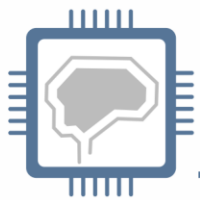
Object Detector Neural Network

Tracking by detection

Frame: 0023

Matching Detections

Frame: 0024

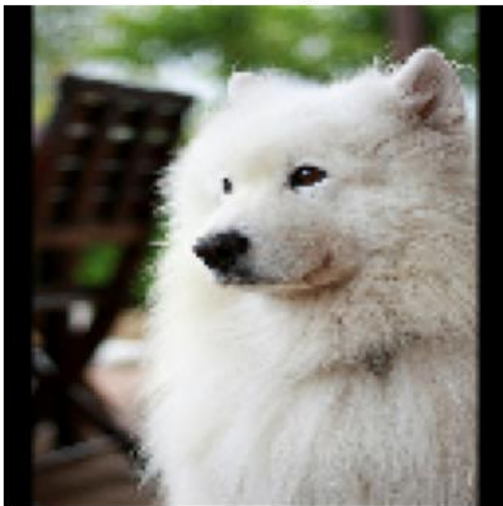


Segmentação Semântica



ceadUFV
Coordenadoria de
Educação Aberta e a Distância

Input Image



True Mask



Predicted Mask



Input Image

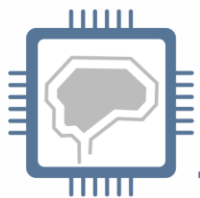


True Mask



Predicted Mask





Segmentação por Exemplo



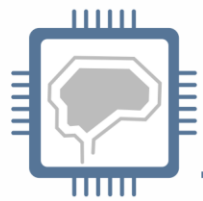
Input Image



Semantic Segmentation

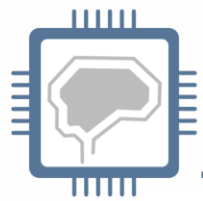


Instance Segmentation

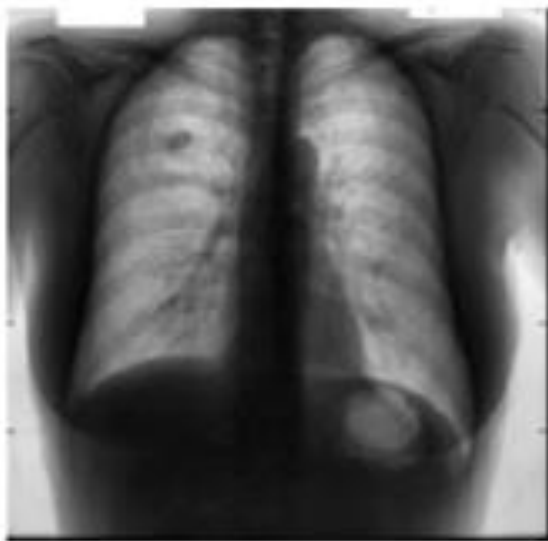


Segmentação por Exemplo

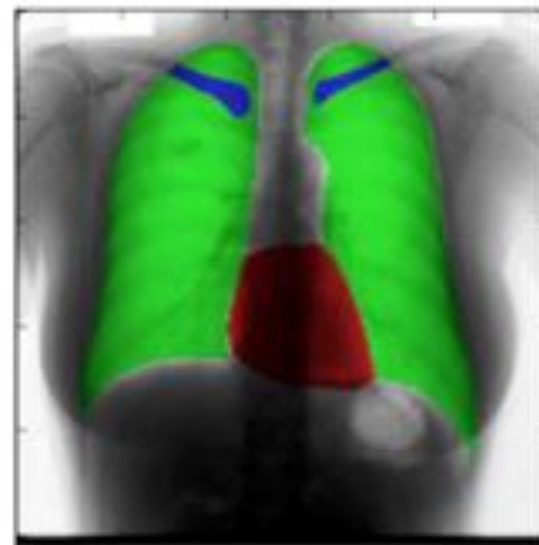




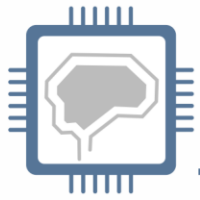
Segmentação por Exemplo



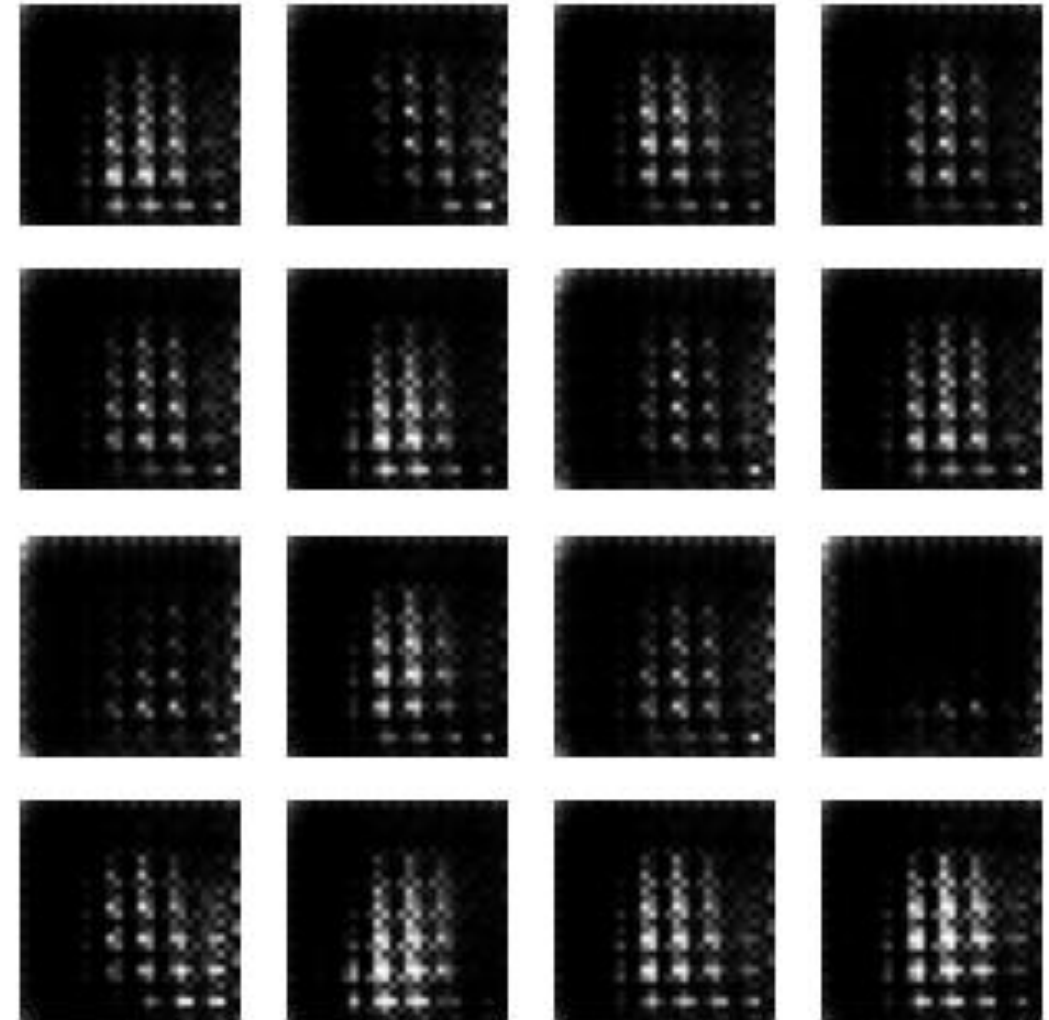
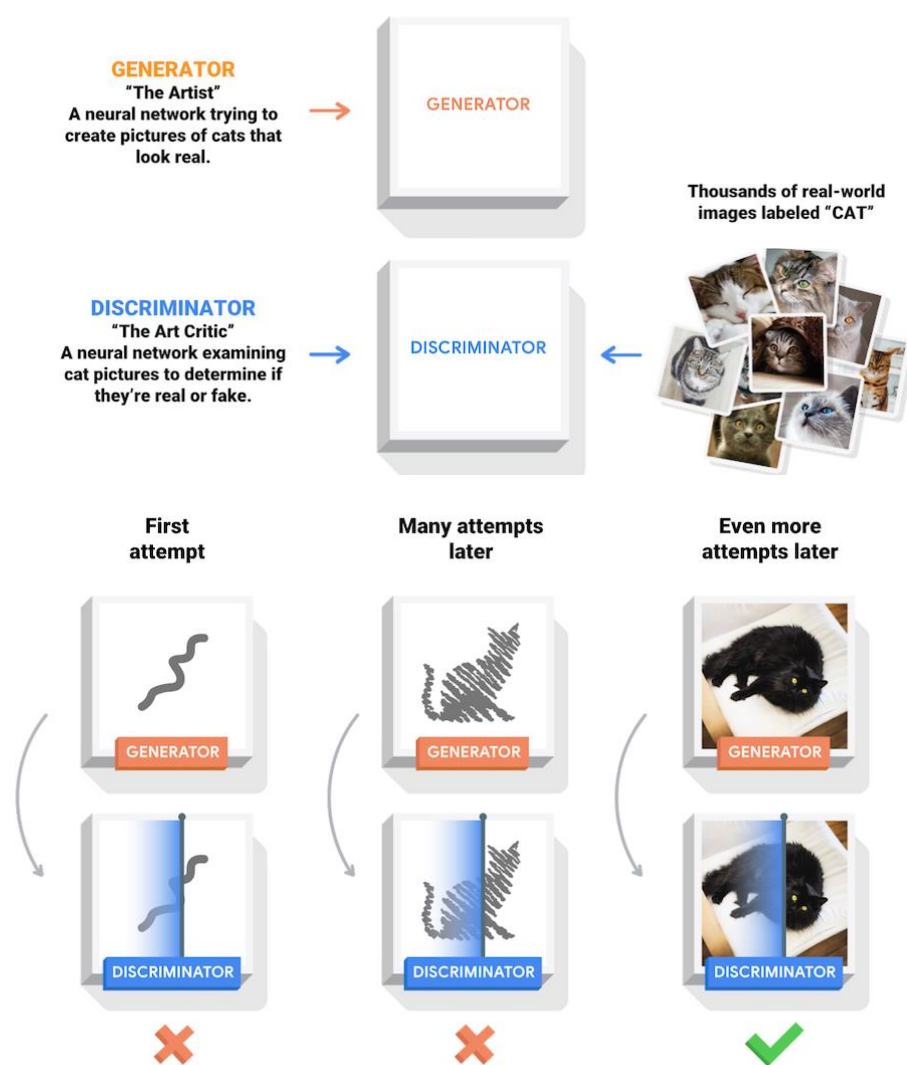
Input Image

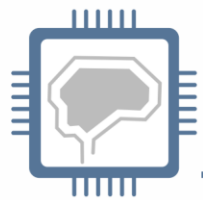


Segmented Image



Deep Convolutional Generative Adversarial Network

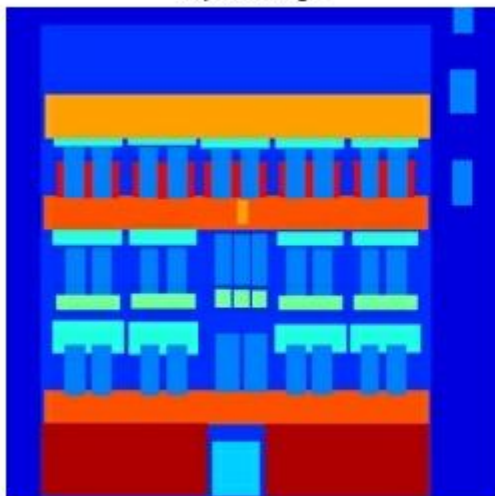




Redes Adversárias Generativas: pix2pix



Input Image



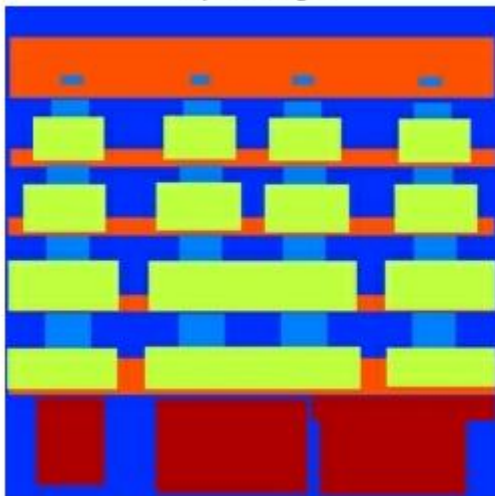
Ground Truth



Predicted Image



Input Image

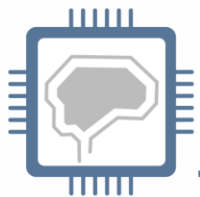


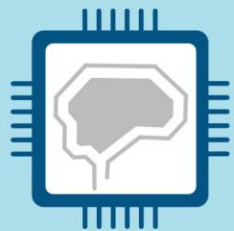
Ground Truth



Predicted Image







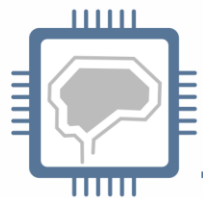
Inteligência Artificial e Computacional

ELT578

ANÁLISE DE IMAGENS E VISÃO COMPUTACIONAL



Conteudista:
Talita E. Z. Santana
talita.santana@ufv.br



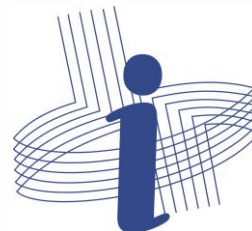
Realização

UFV

Universidade Federal de Viçosa

ENGENHARIA
ELÉTRICA

Universidade Federal de Viçosa



NIAS

Núcleo Interdisciplinar de Análise de Sinais

