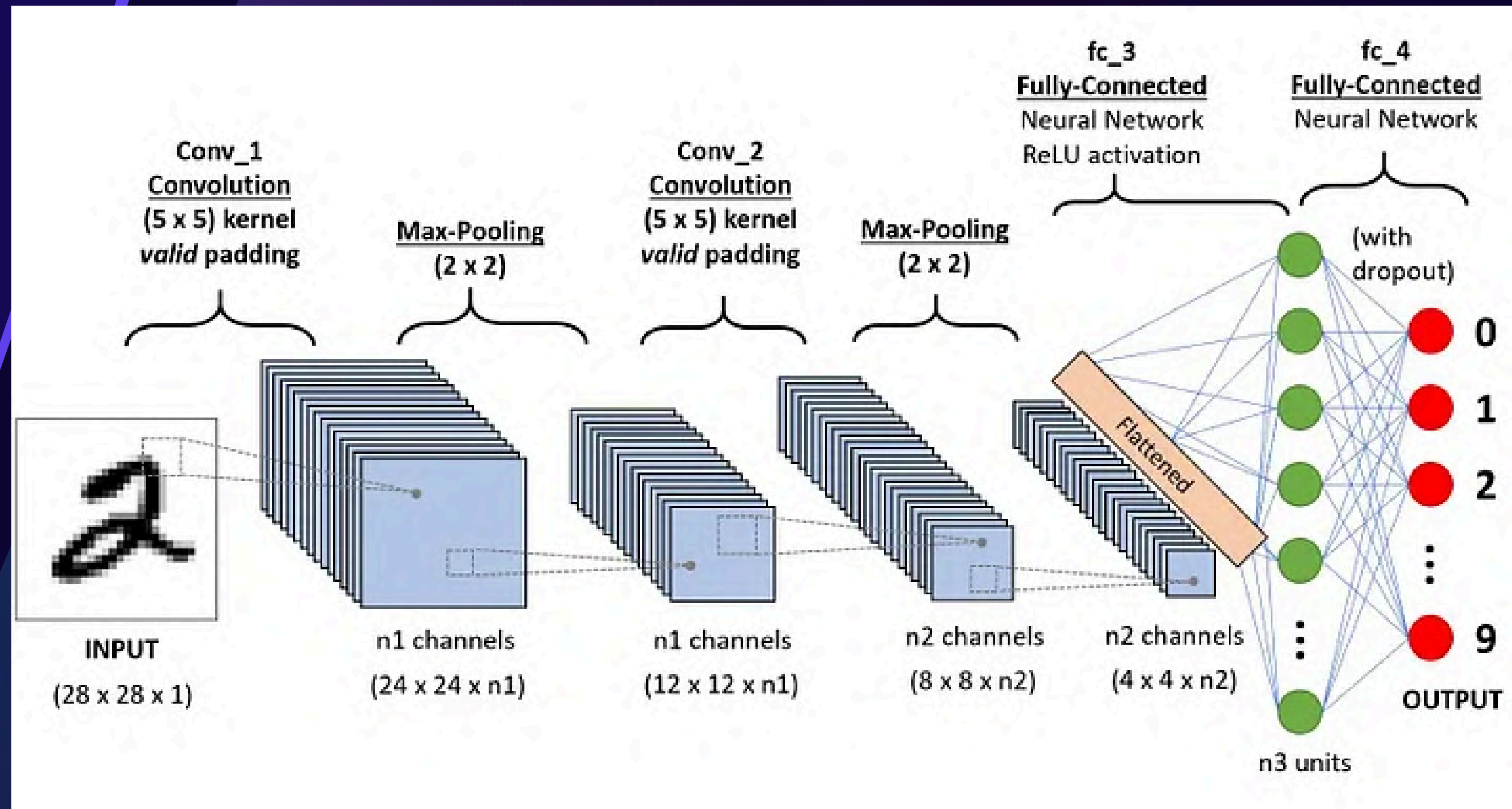




Introdução a Classificação de Imagens

COM TENSORFLOW & KERAS

REDES CONVOLUCIONAIS



CONVOLUÇÃO

0	1	2
3	4	5
6	7	8

0	1
2	3

19	25
37	43

0	1	2
3	4	5
6	7	8

0	1
2	3

19	25
37	43

0	1	2
3	4	5
6	7	8

0	1
2	3

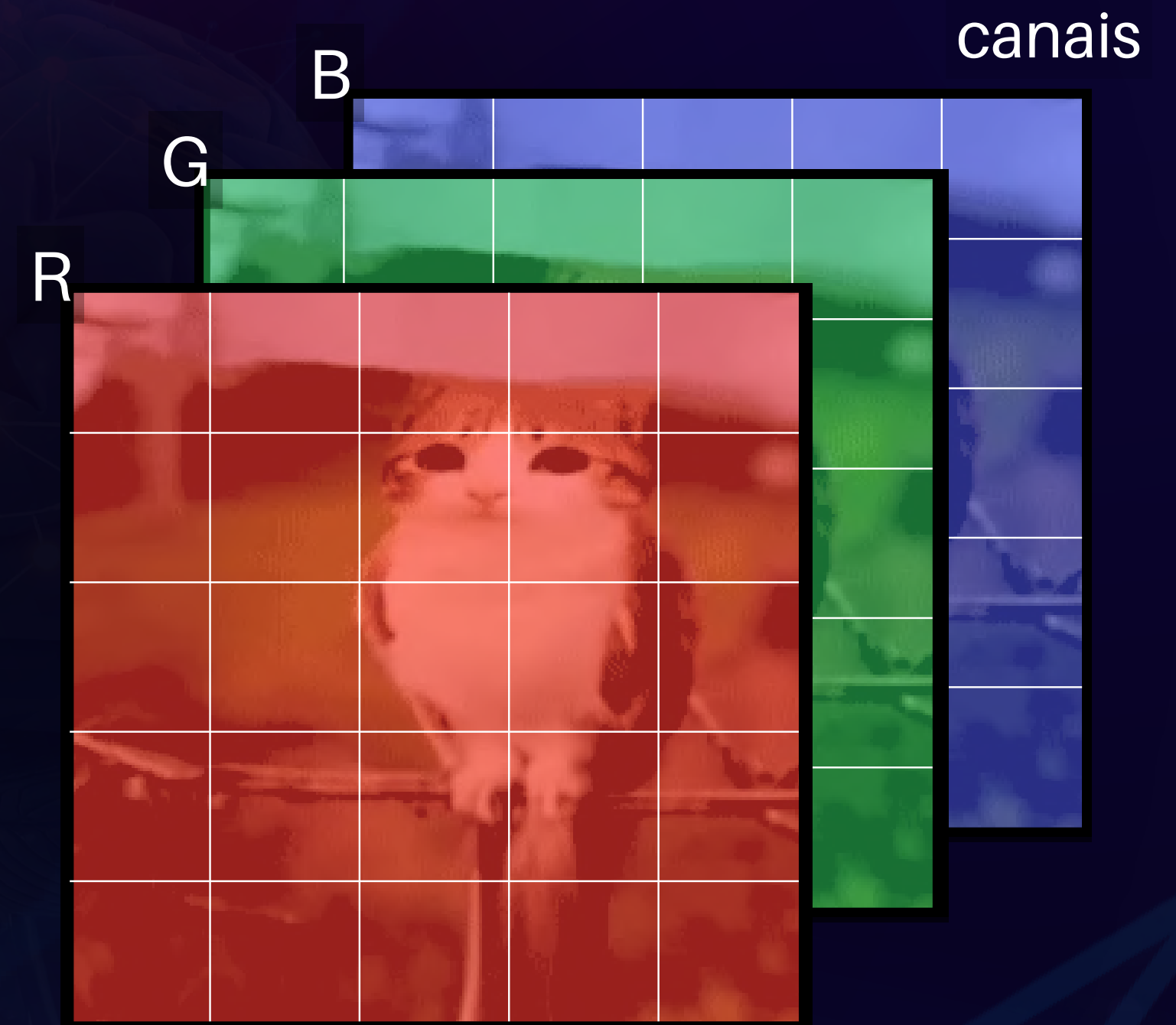
19	25
37	43

0	1	2
3	4	5
6	7	8

0	1
2	3

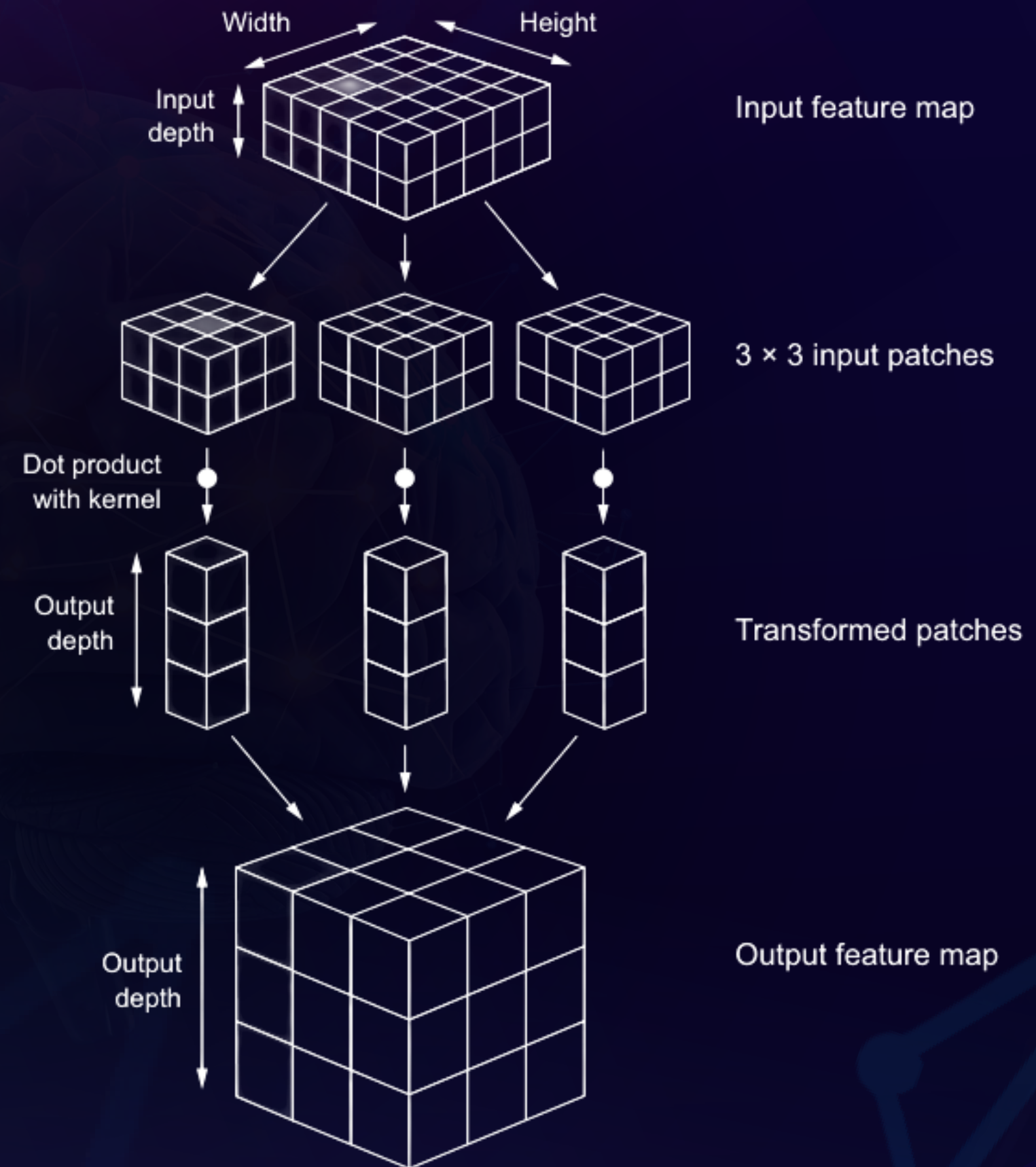
19	25
37	43

REPRESENTAÇÃO DE IMAGENS

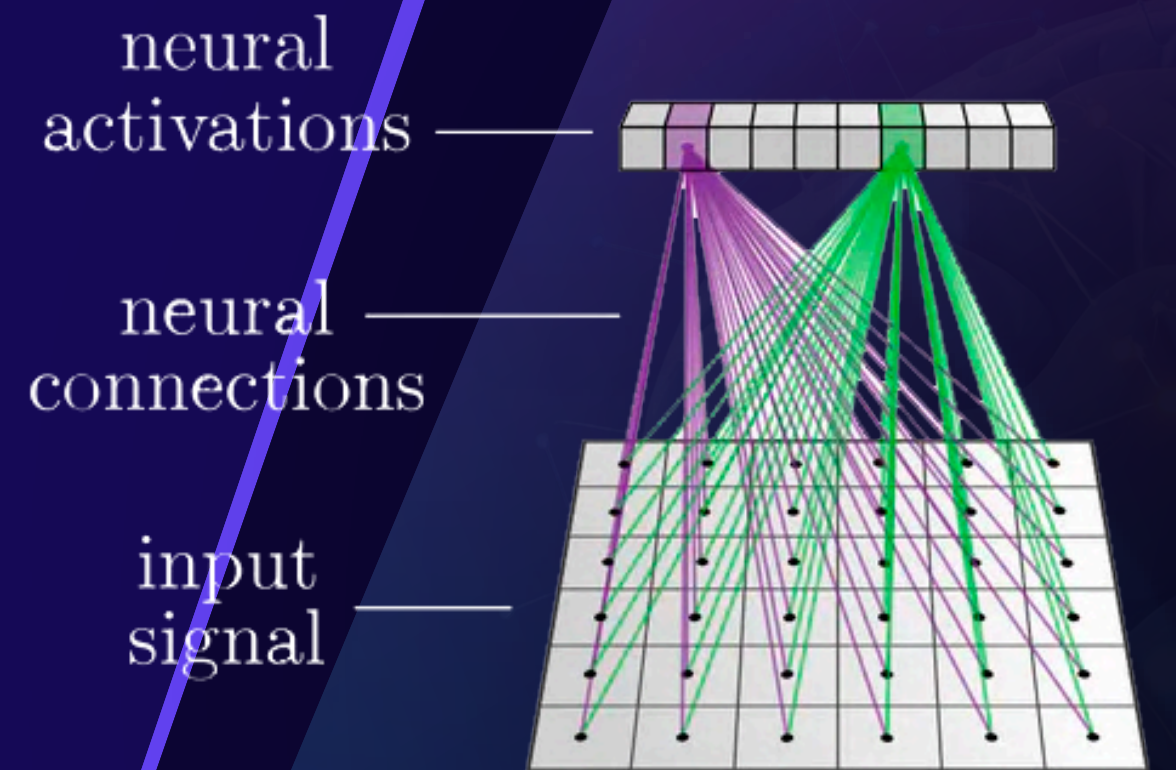


CONVOLUÇÃO

- O kernel desliza sobre a entrada 3D, extraindo as features de cada região
- Cada feature é um vetor de dimensão igual ao número de canais
- Os vetores são reunidos em um mapa de saída
- Cada região da saída corresponde a mesma região da entrada

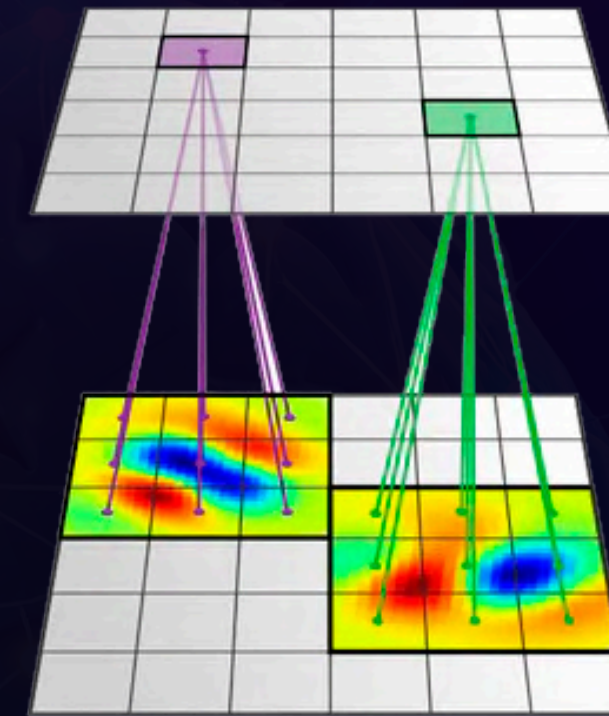


MLP X CNN

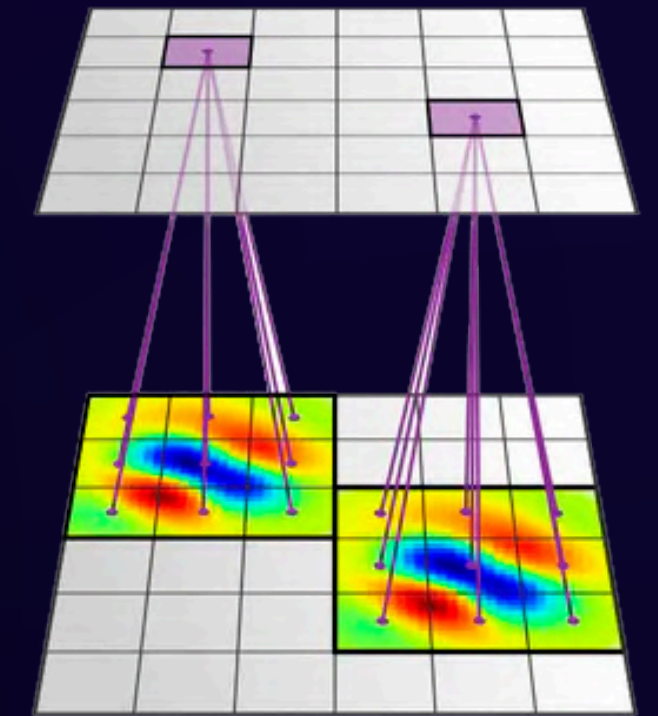


fully connected (MLP)

- Cada neurônio recebe entradas de todos os pixels
- Total de parâmetros = n^2



locally connected

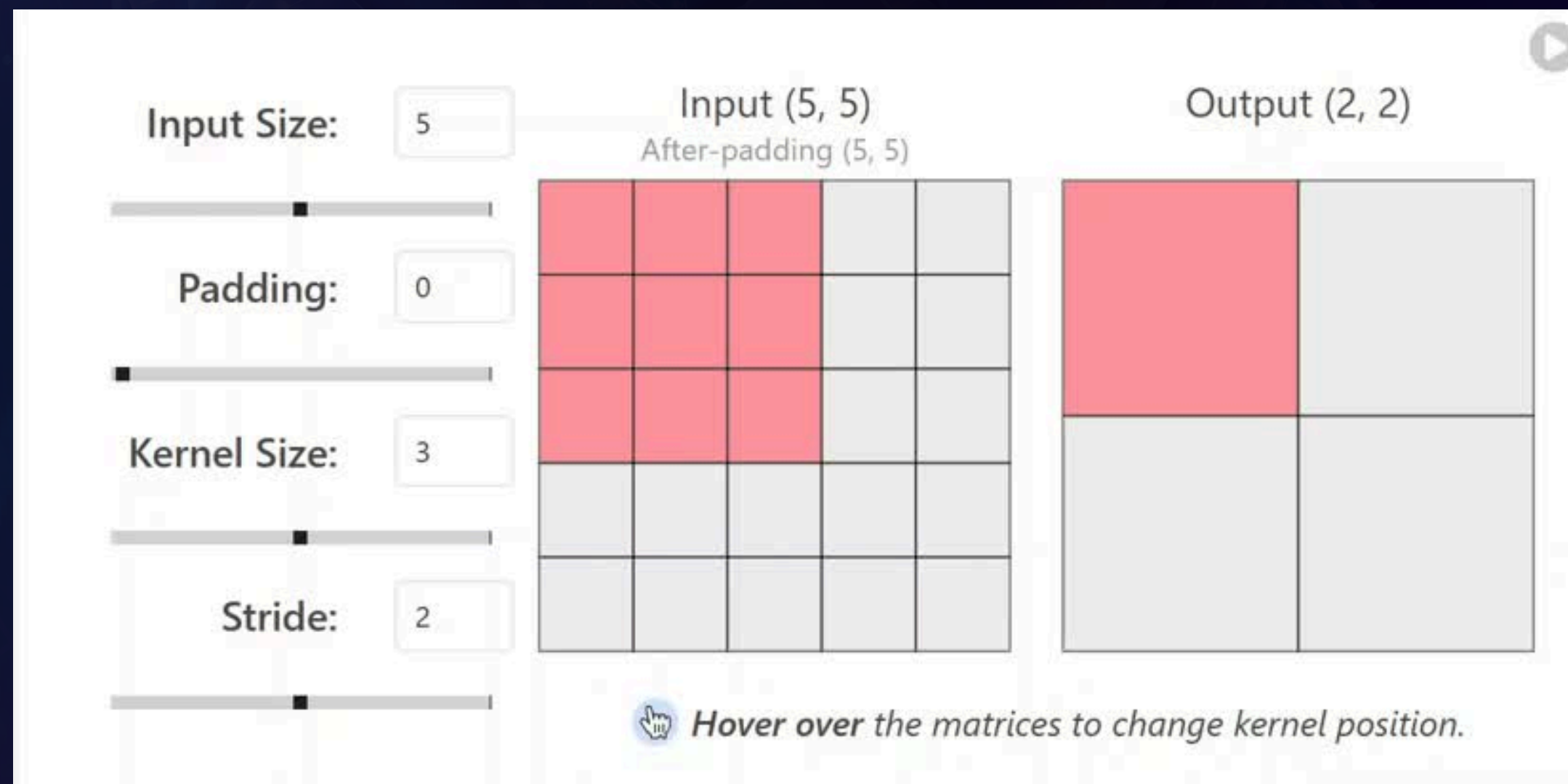


convolutional (CNN)

- Cada unidade da camada recebe entradas de uma região local
- Compartilhamento de pesos por toda a imagem
- Total de parâmetros = dl
 - $d \rightarrow$ núm. de filtros
 - $l \rightarrow$ núm. de pesos

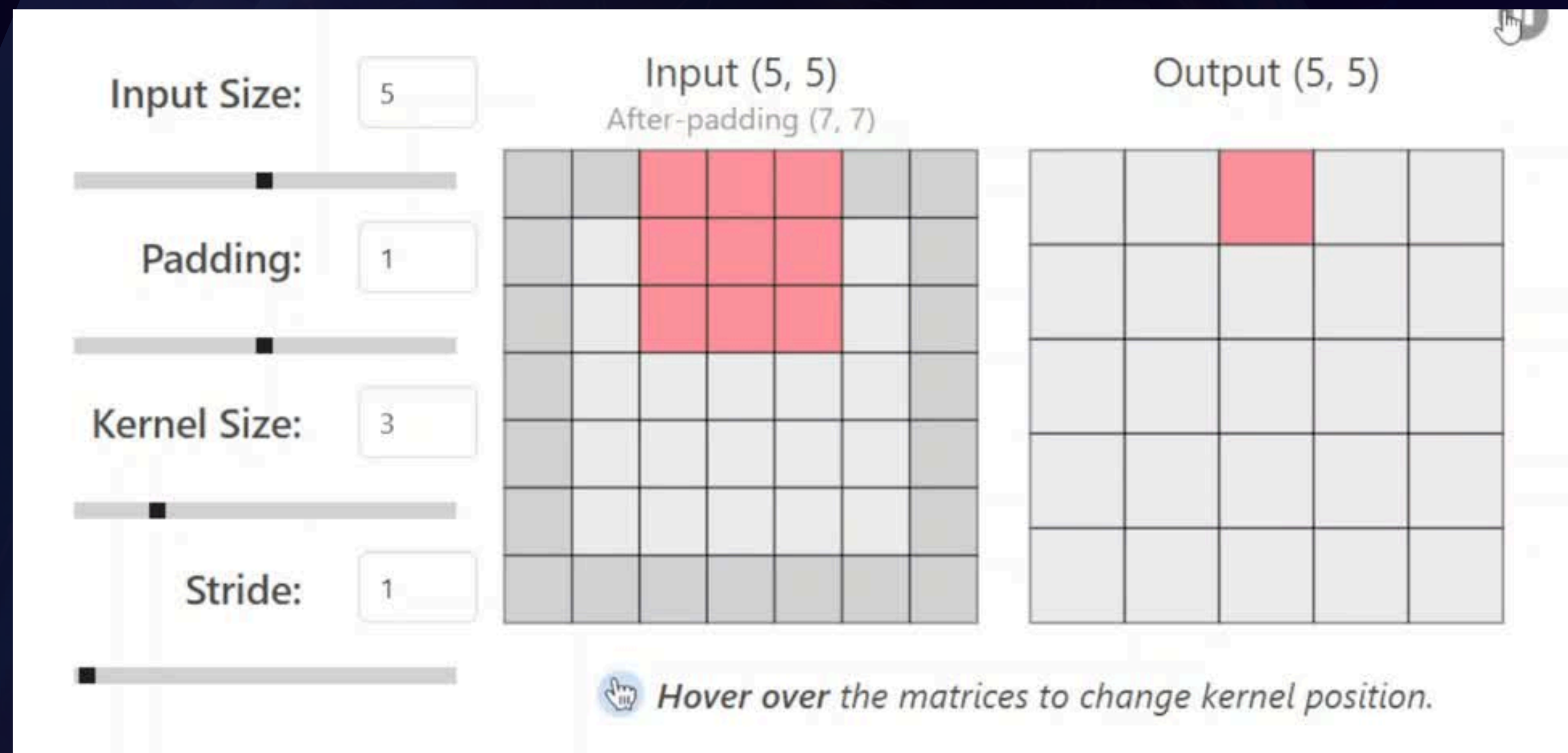
STRIDE

- Stride se refere ao número de pixels em que o kernel é deslocado a cada passo da convolução
- Maior stride:
 - Perda de detalhes devido a menor sobreposição de regiões
 - Saída com menor dimensão

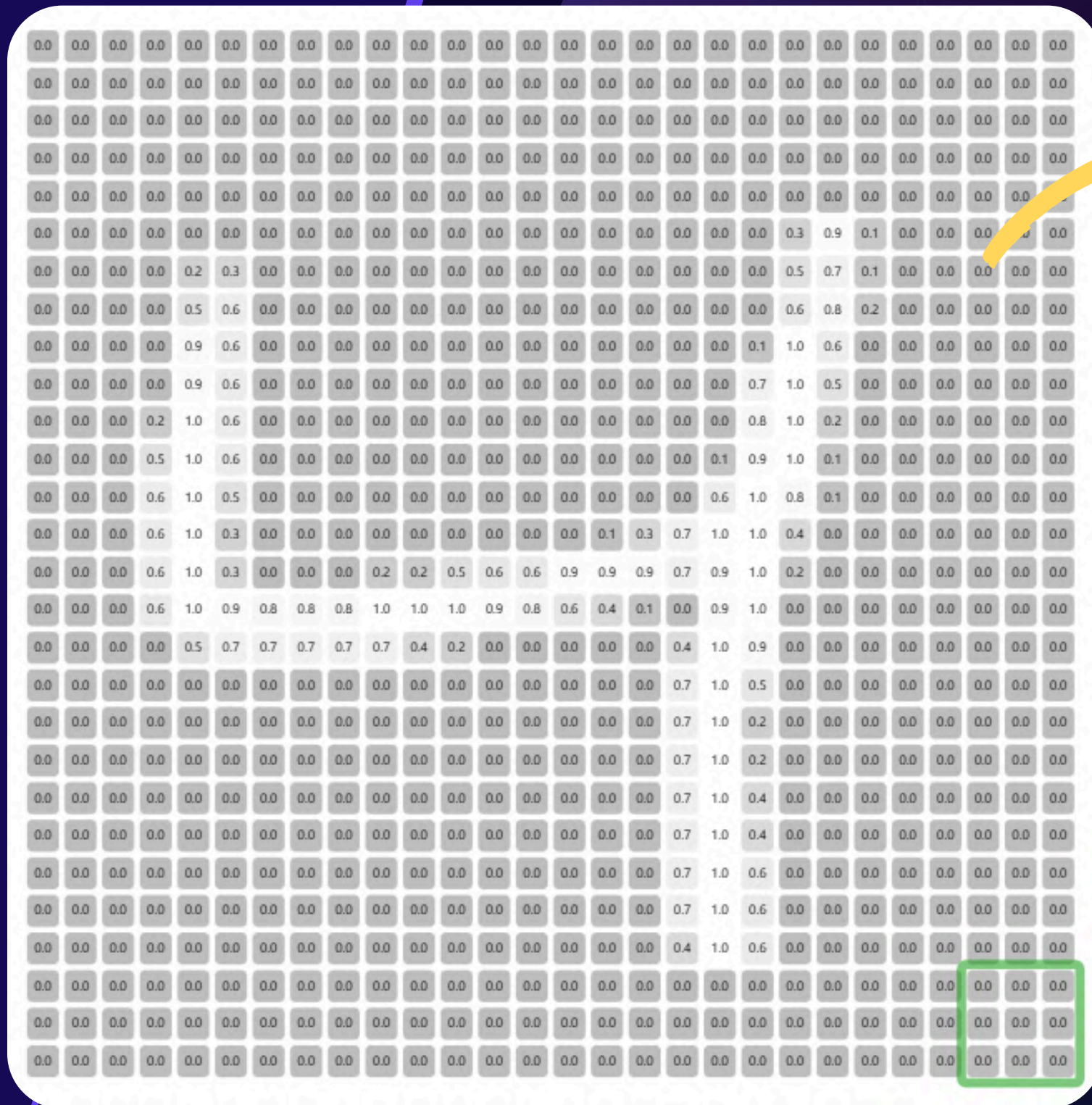


PADDING

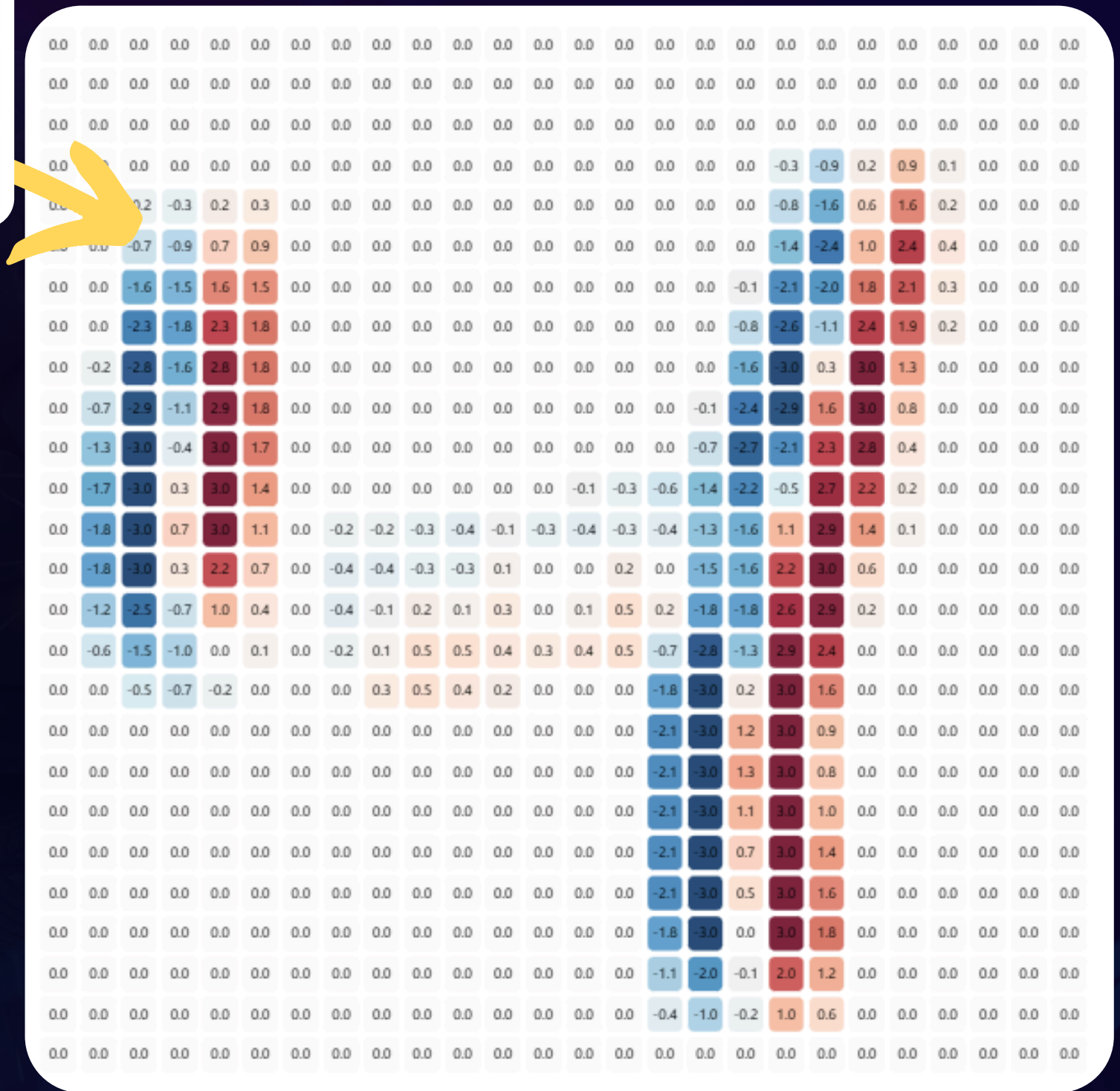
- Padding se refere a adição de linhas ou colunas ao redor da imagem
- Objetivo é ajustar o tamanho da saída e preservar as informações presentes nas bordas



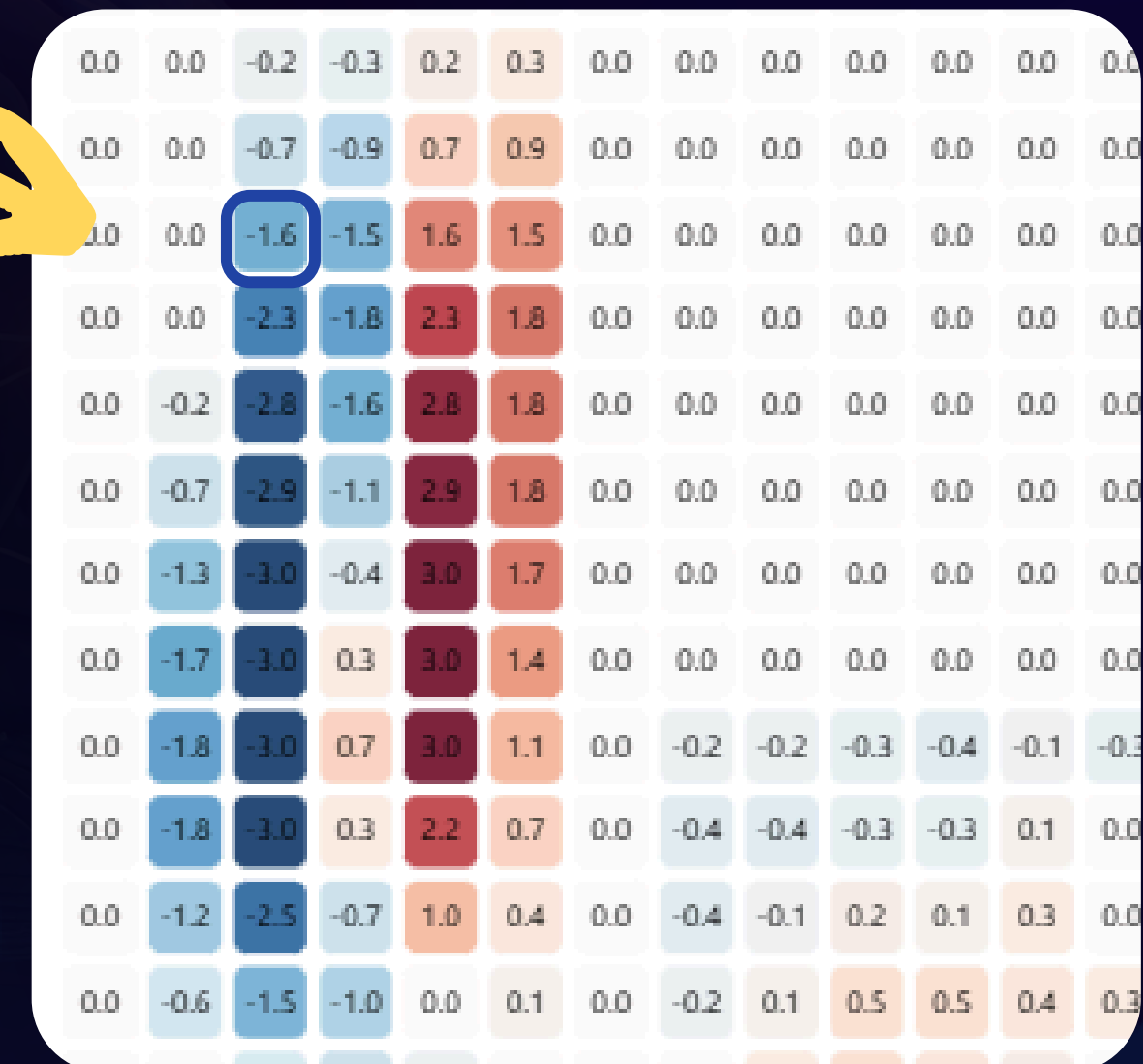
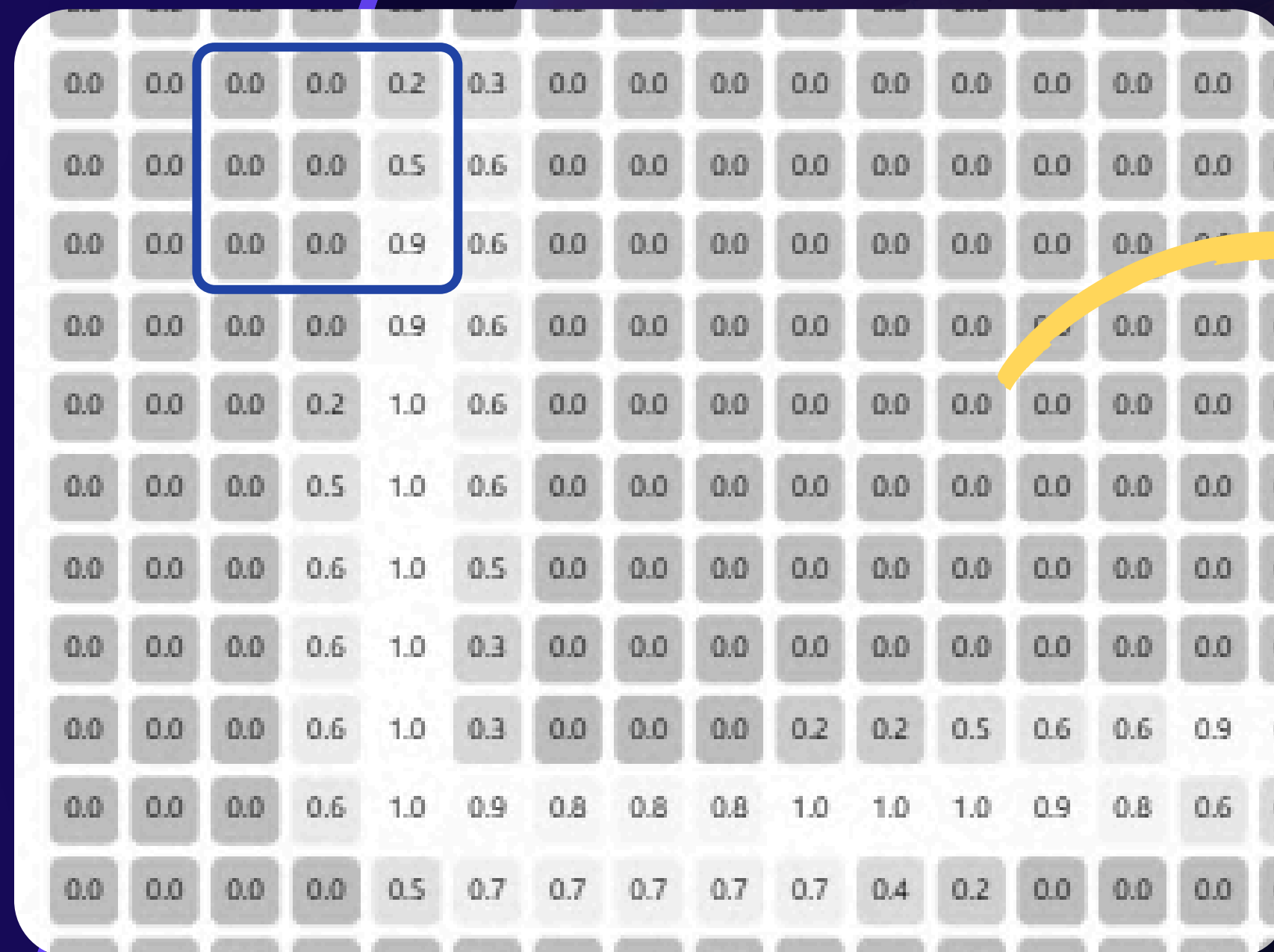
DETECÇÃO DE BORDAS VERTICAIS



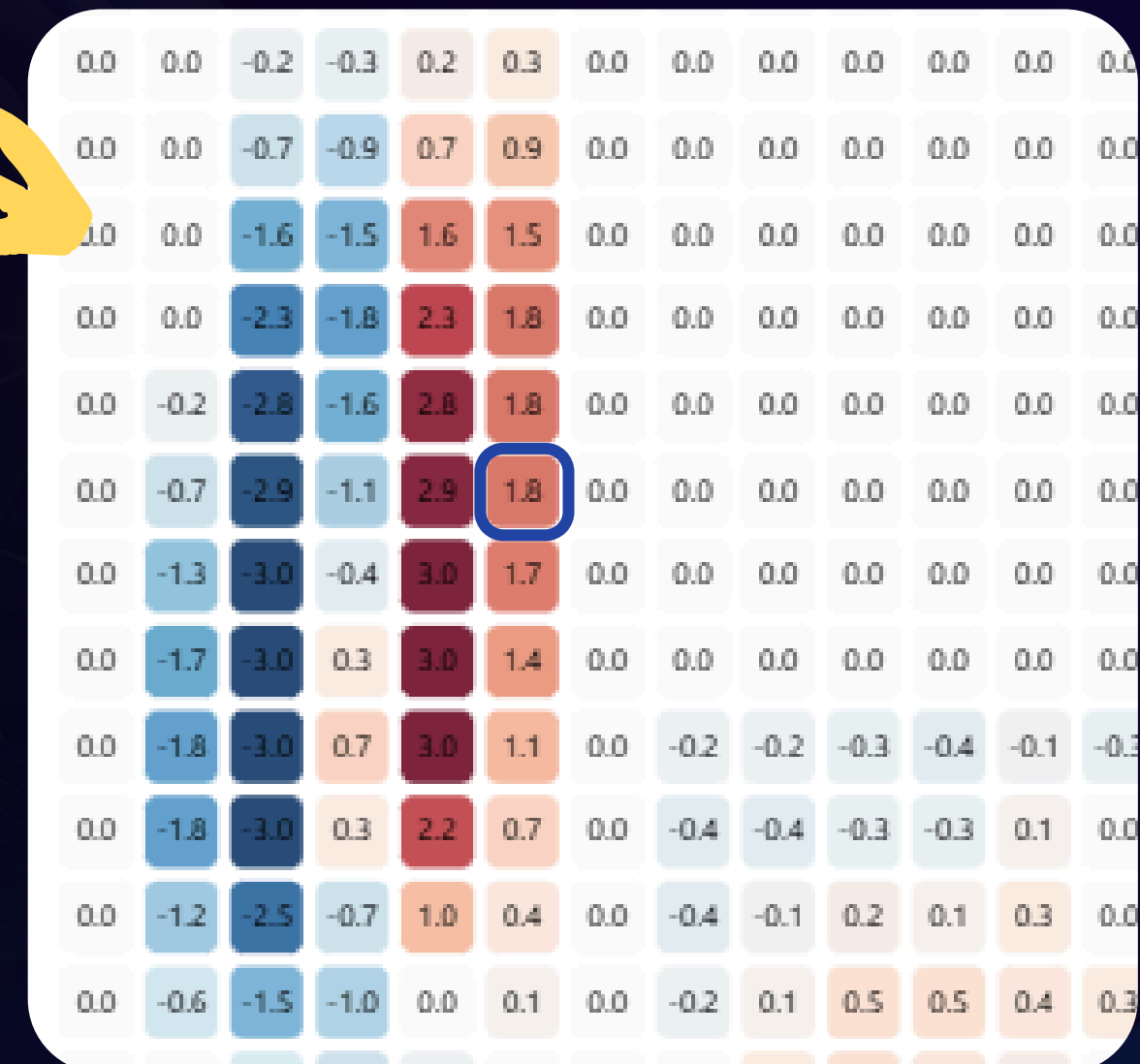
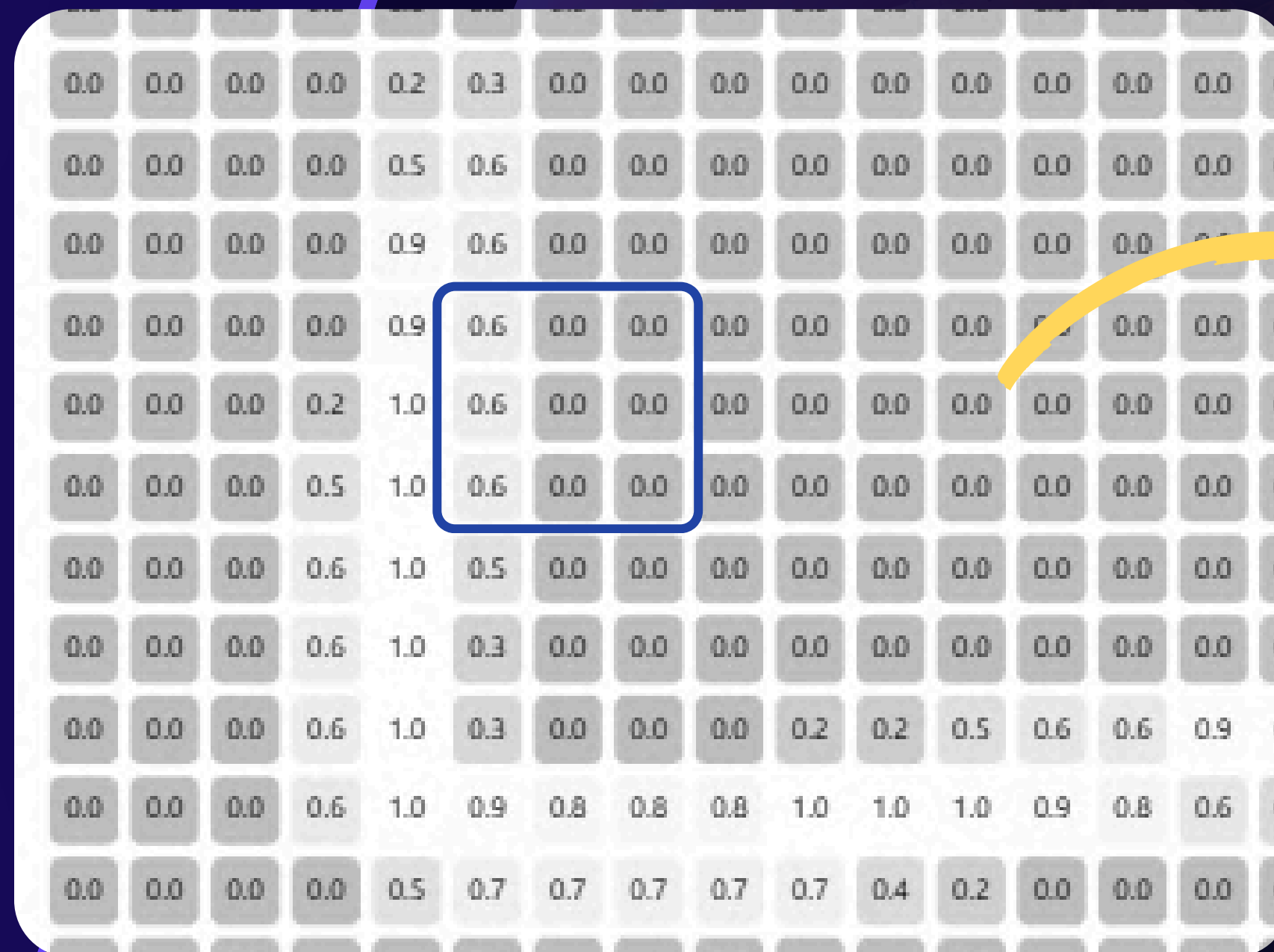
1.0	0.0	-1.0
1.0	0.0	-1.0
1.0	0.0	-1.0



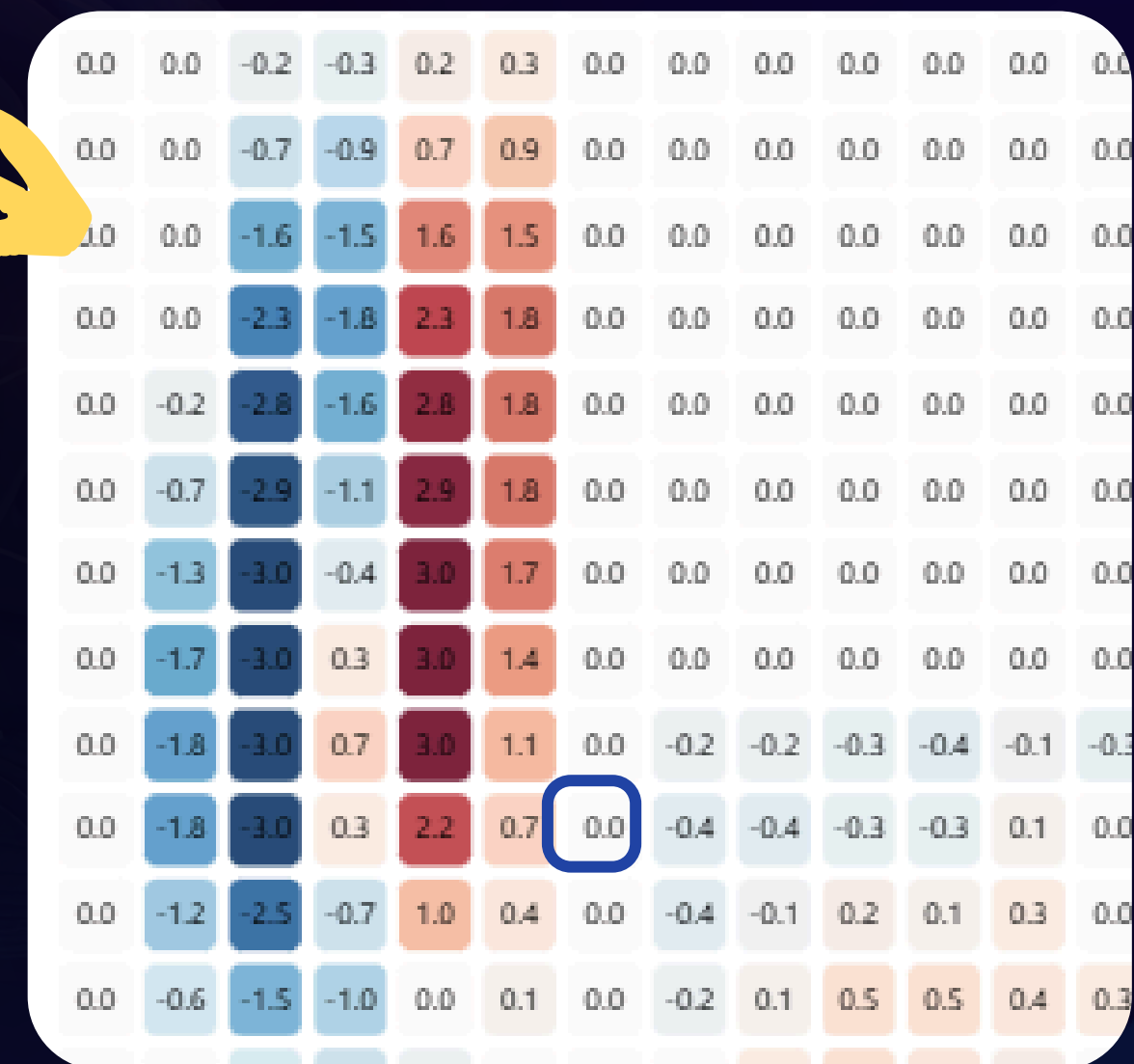
DETECÇÃO DE BORDAS VERTICAIS



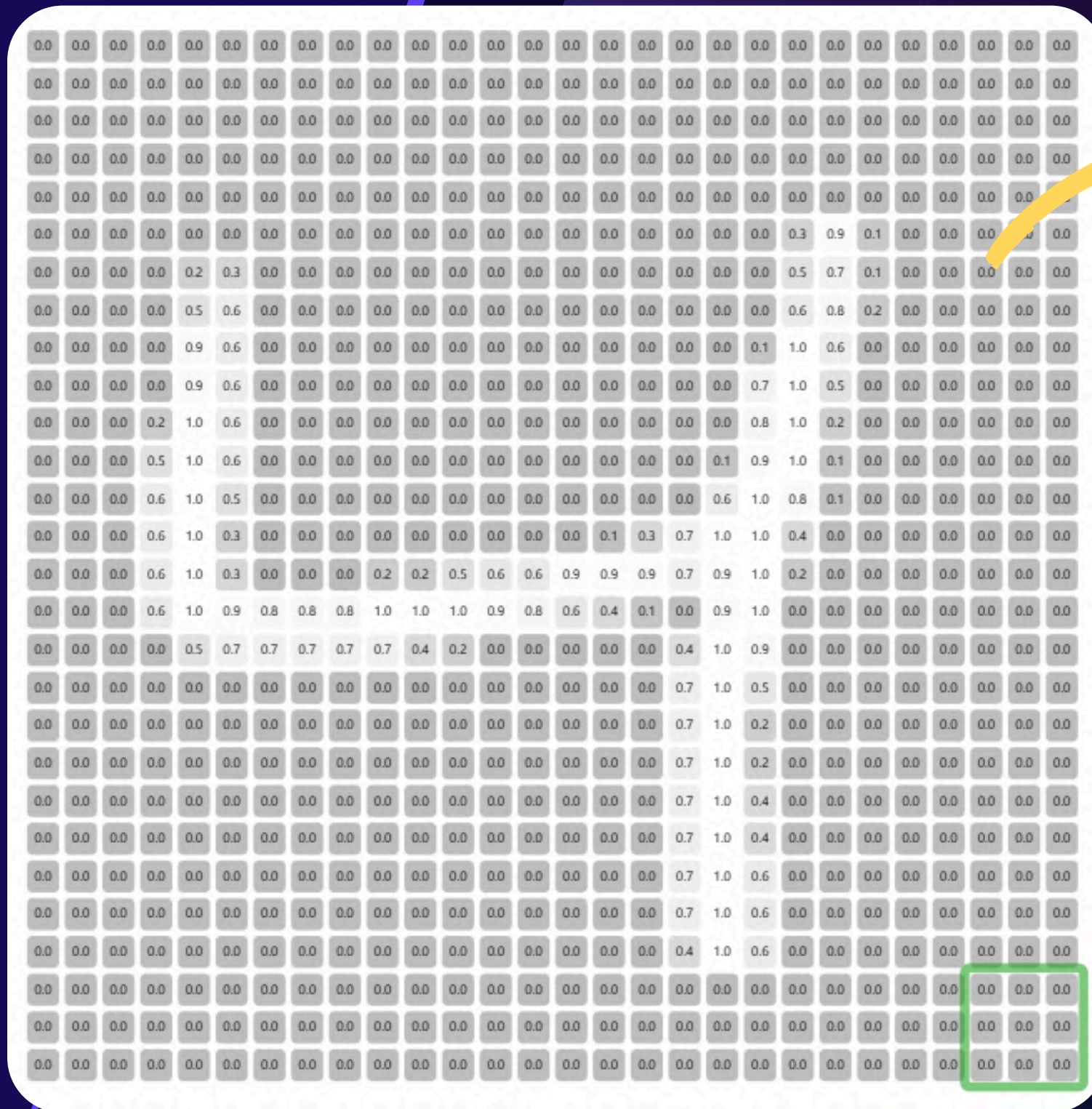
DETECÇÃO DE BORDAS VERTICAIS



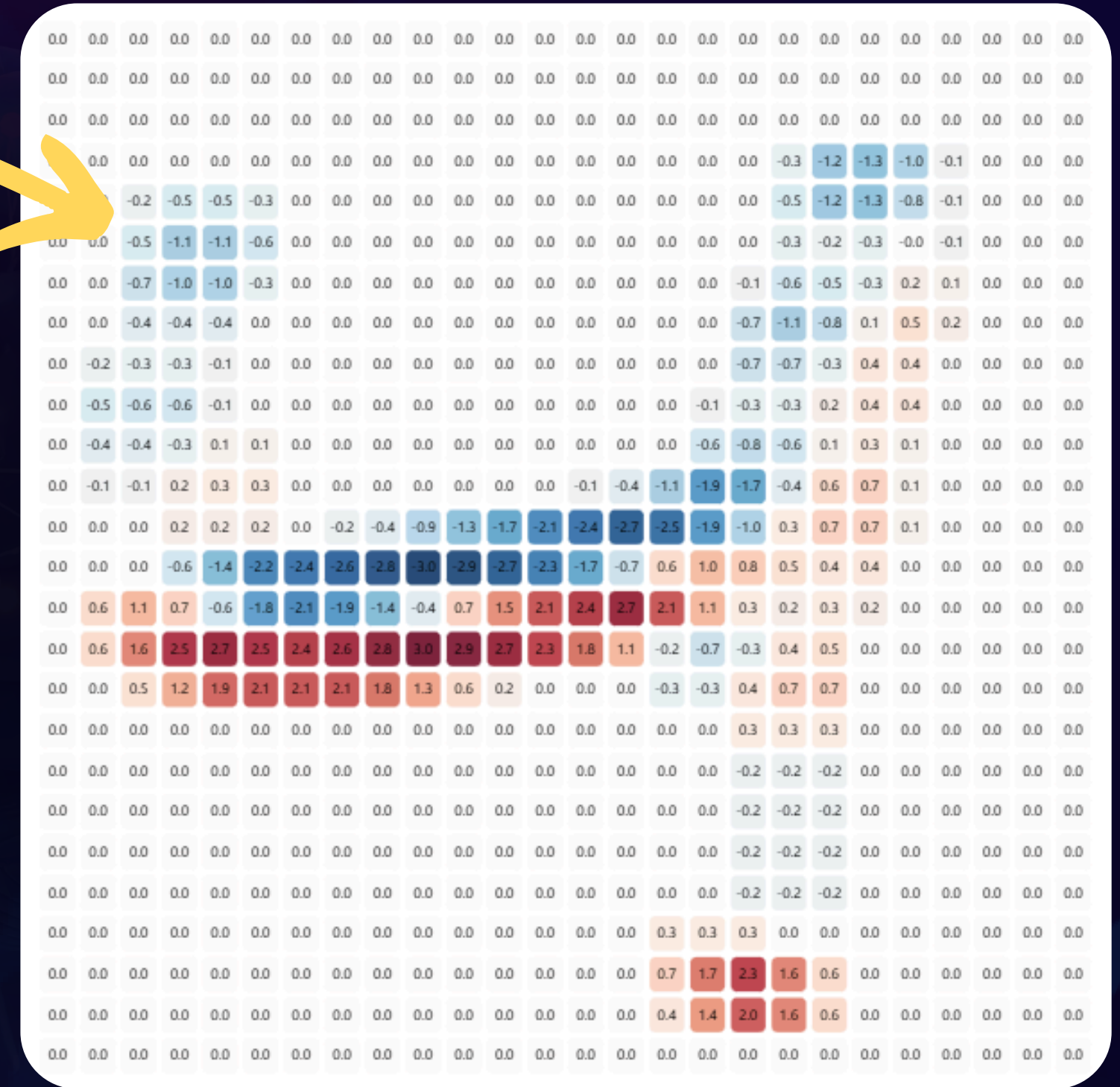
DETECÇÃO DE BORDAS VERTICAIS



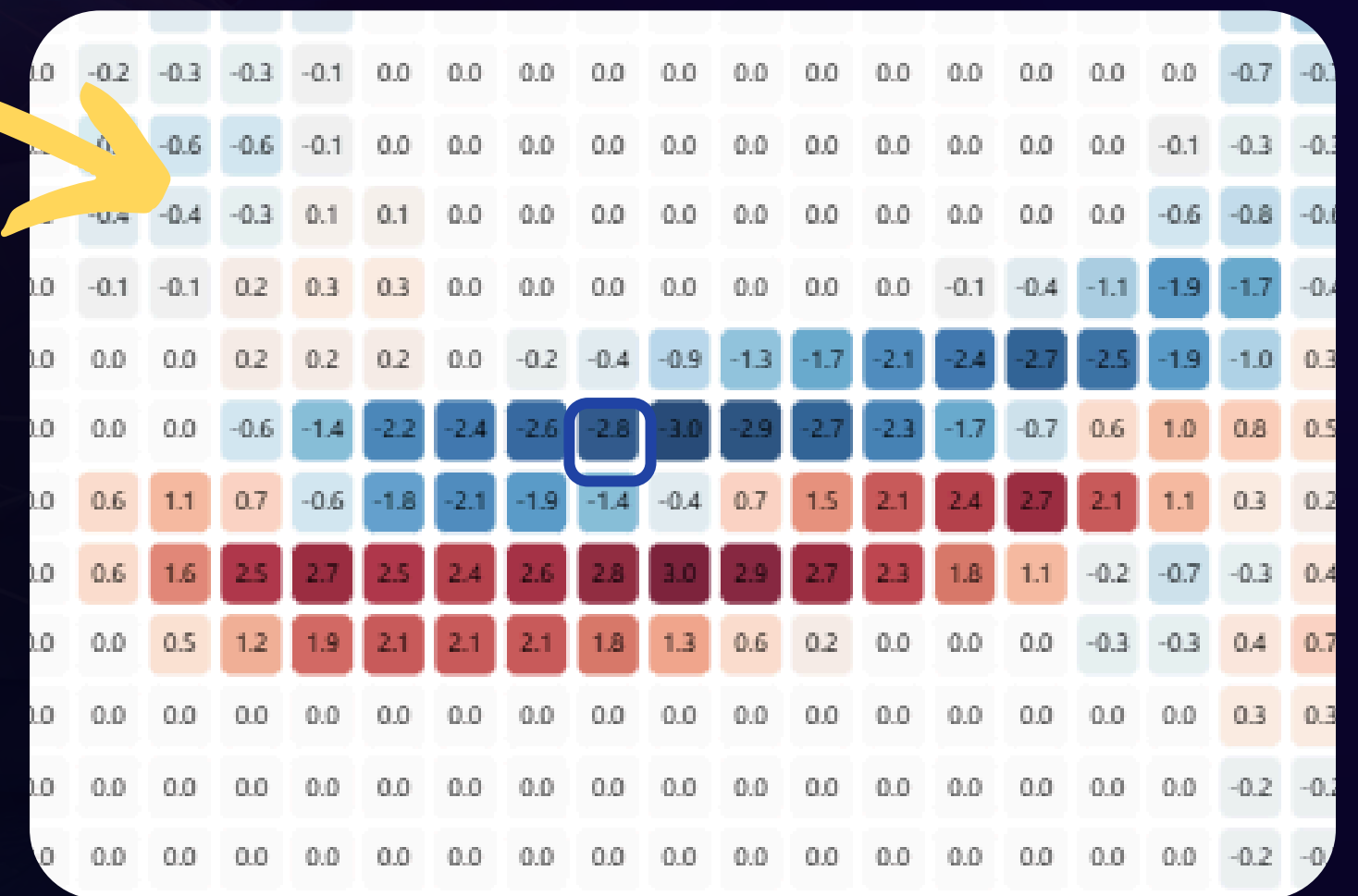
DETECÇÃO DE BORDAS HORIZONTAIS



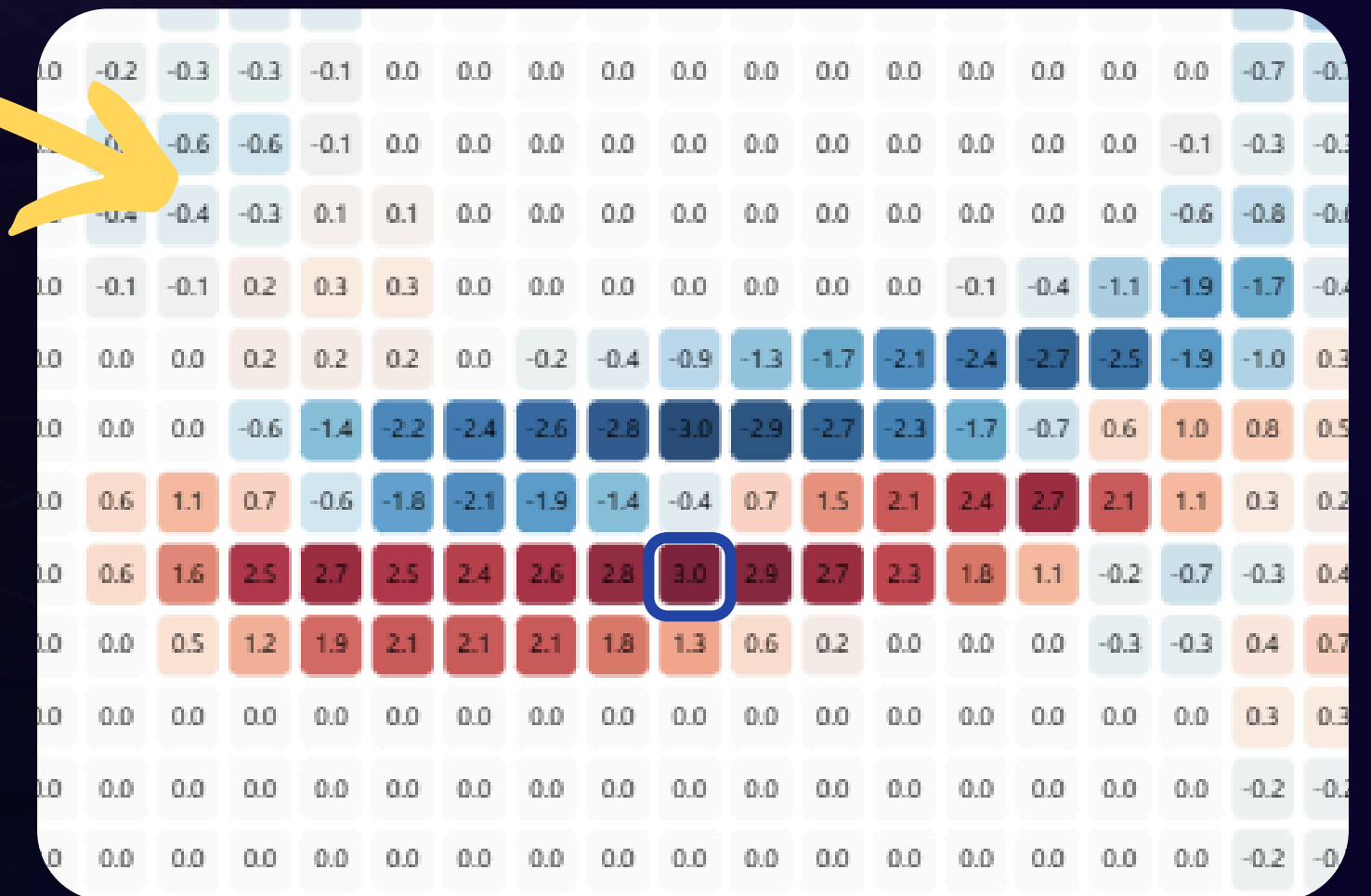
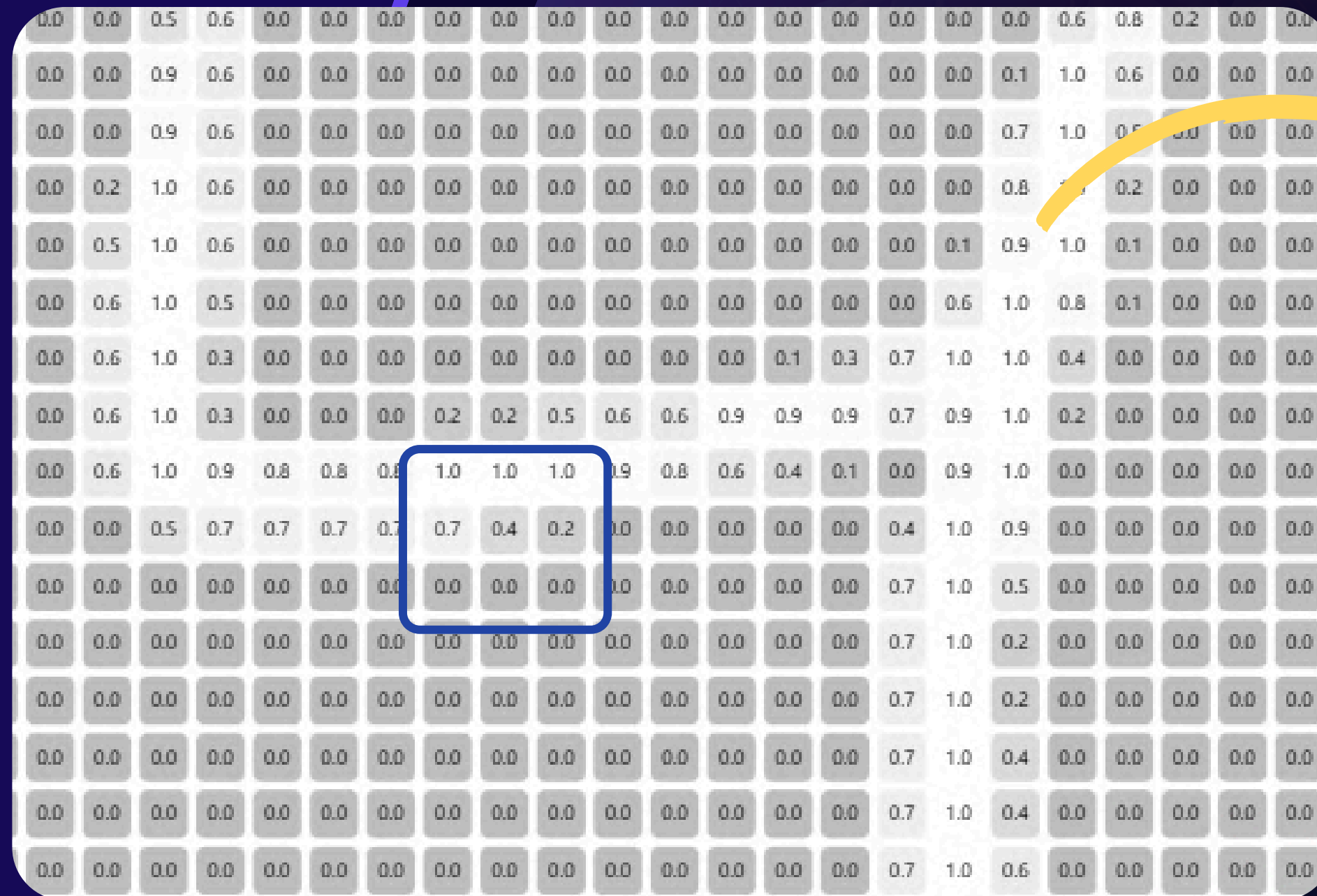
1.0	1.0	1.0
0.0	0.0	0.0
-1.0	-1.0	-1.0



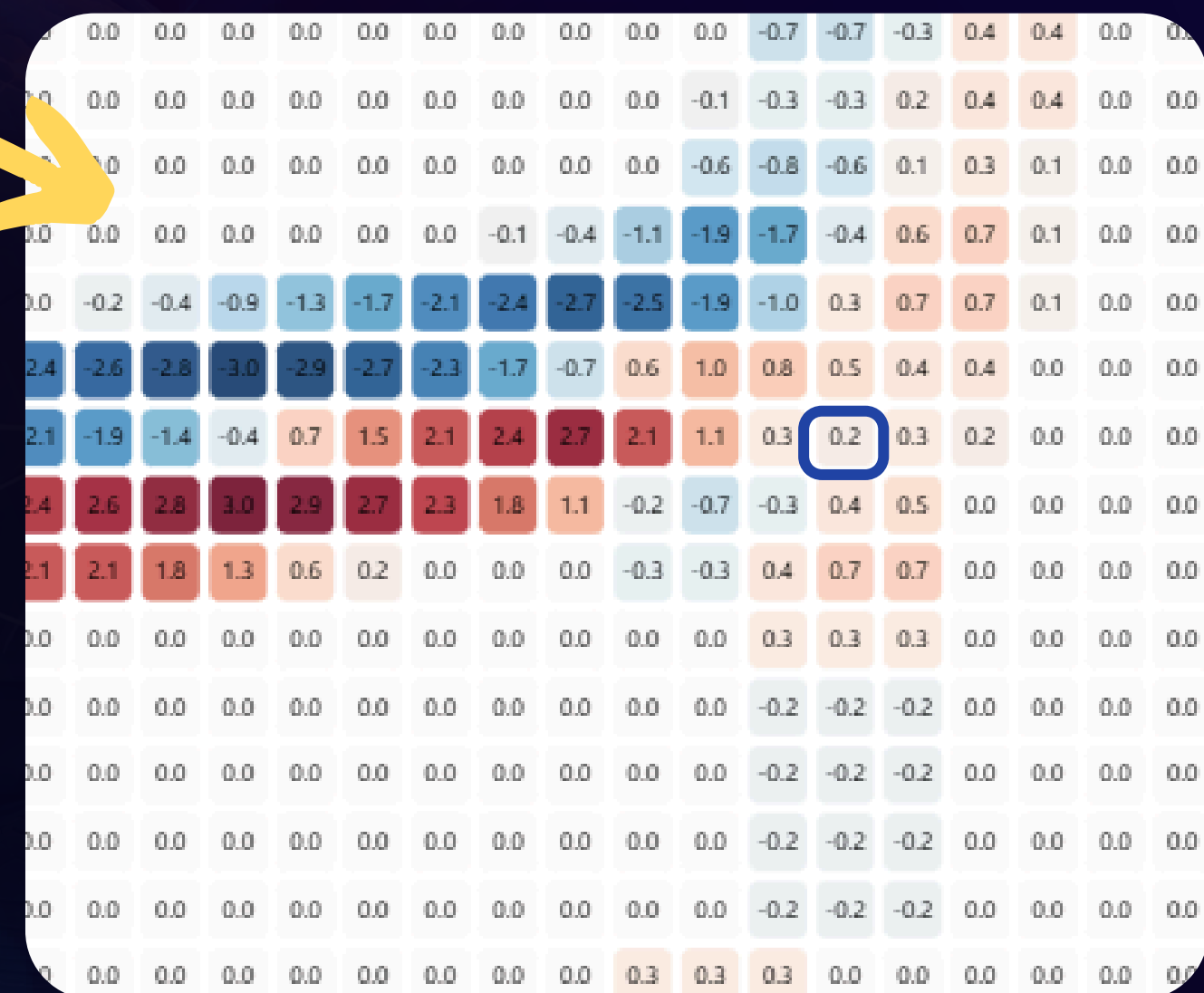
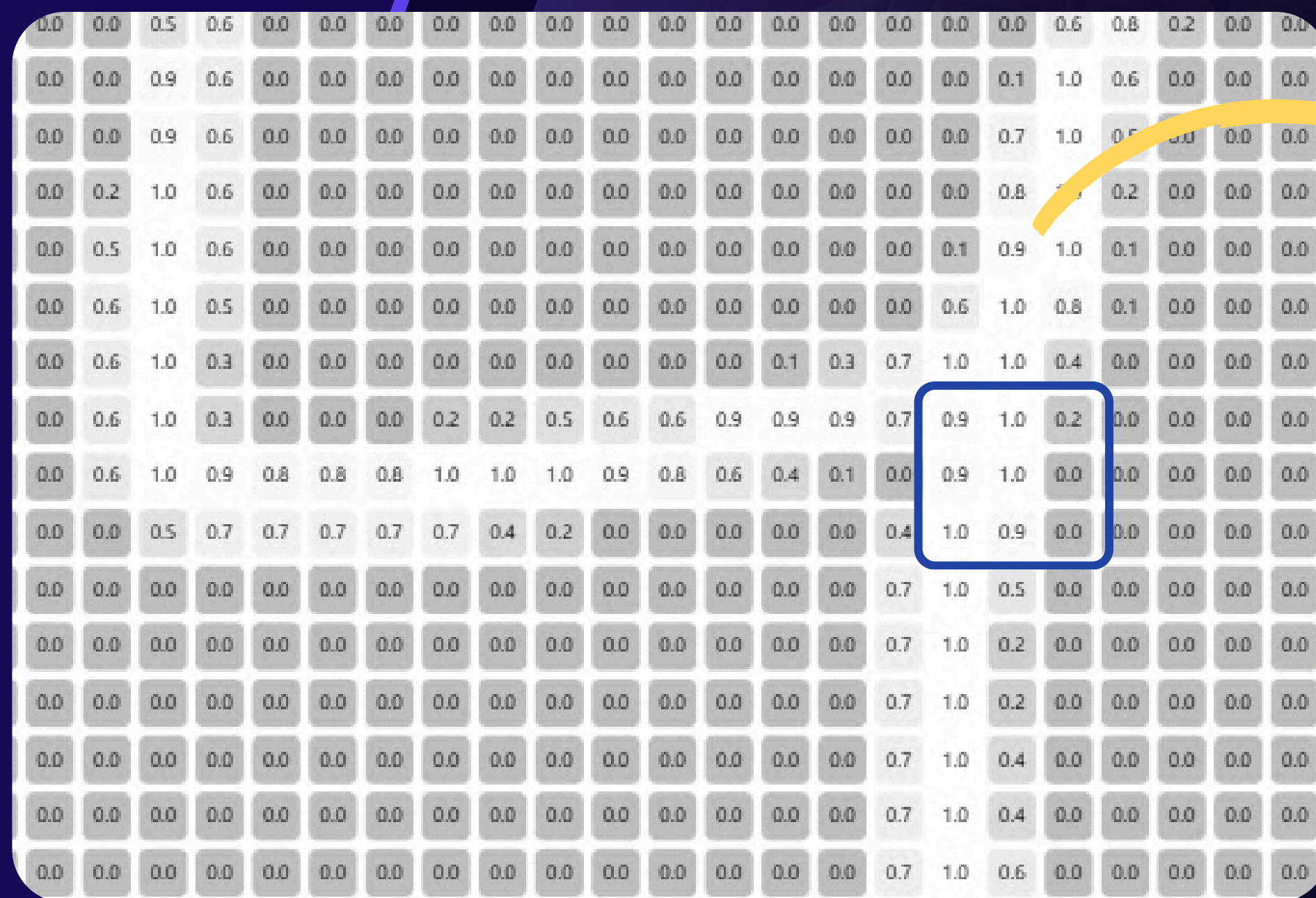
DETECÇÃO DE BORDAS HORIZONTAIS



DETECÇÃO DE BORDAS HORIZONTAIS



DETECÇÃO DE BORDAS HORIZONTAIS

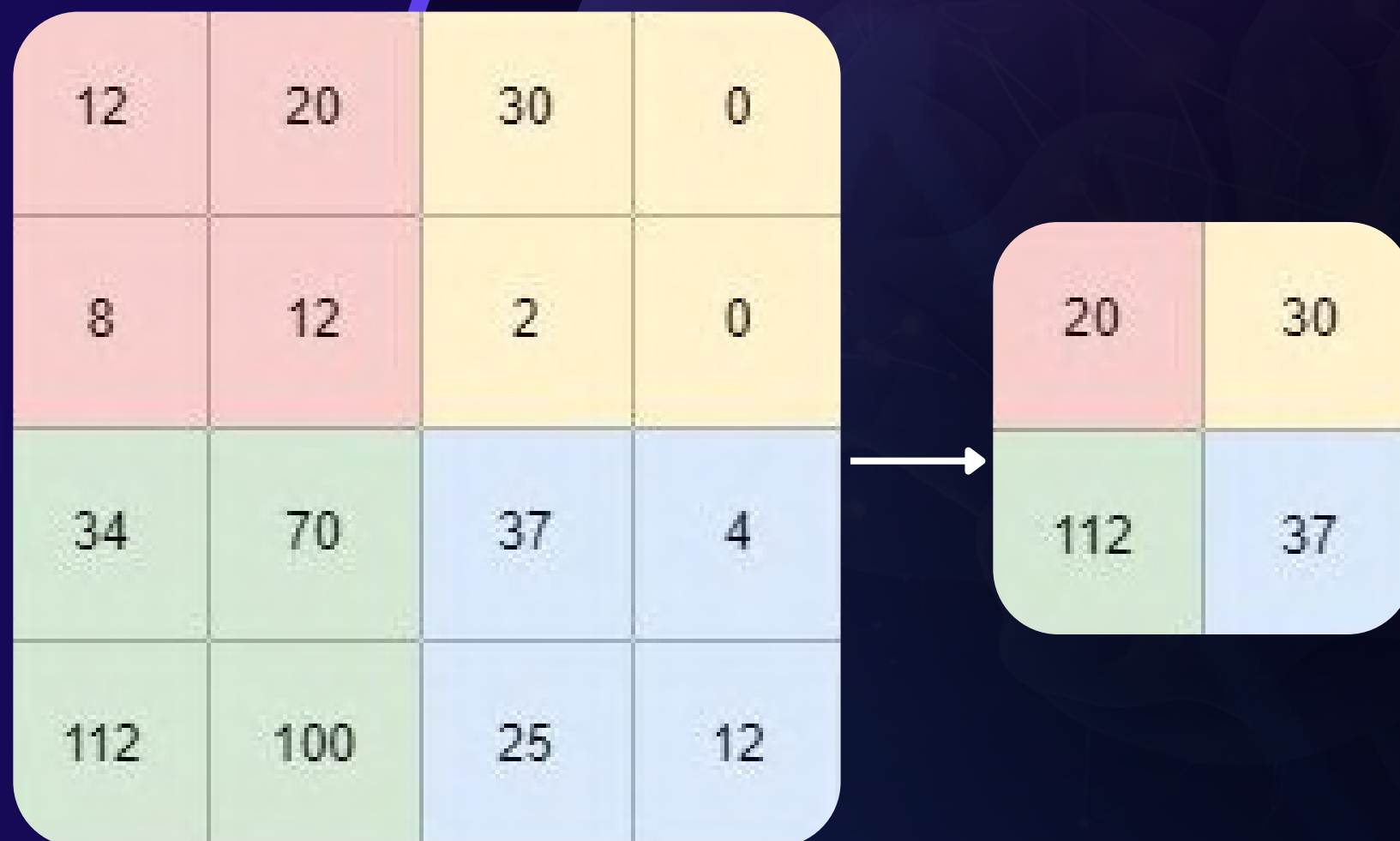


POOLING

- Pooling é uma operação que resume os valores de unidades adjacentes da camada anterior em um único valor, reduzindo dimensionalidade enquanto retém informações
- Semelhante a convolução, é aplicado um kernel sobre o feature map
- Ao contrário da convolução, o objetivo não é extração de padrões e possui kernel fixo
- Importante para:
 - Redução de dimensionalidade do feature map
 - Diminuição do custo computacional
 - Invariância de translação

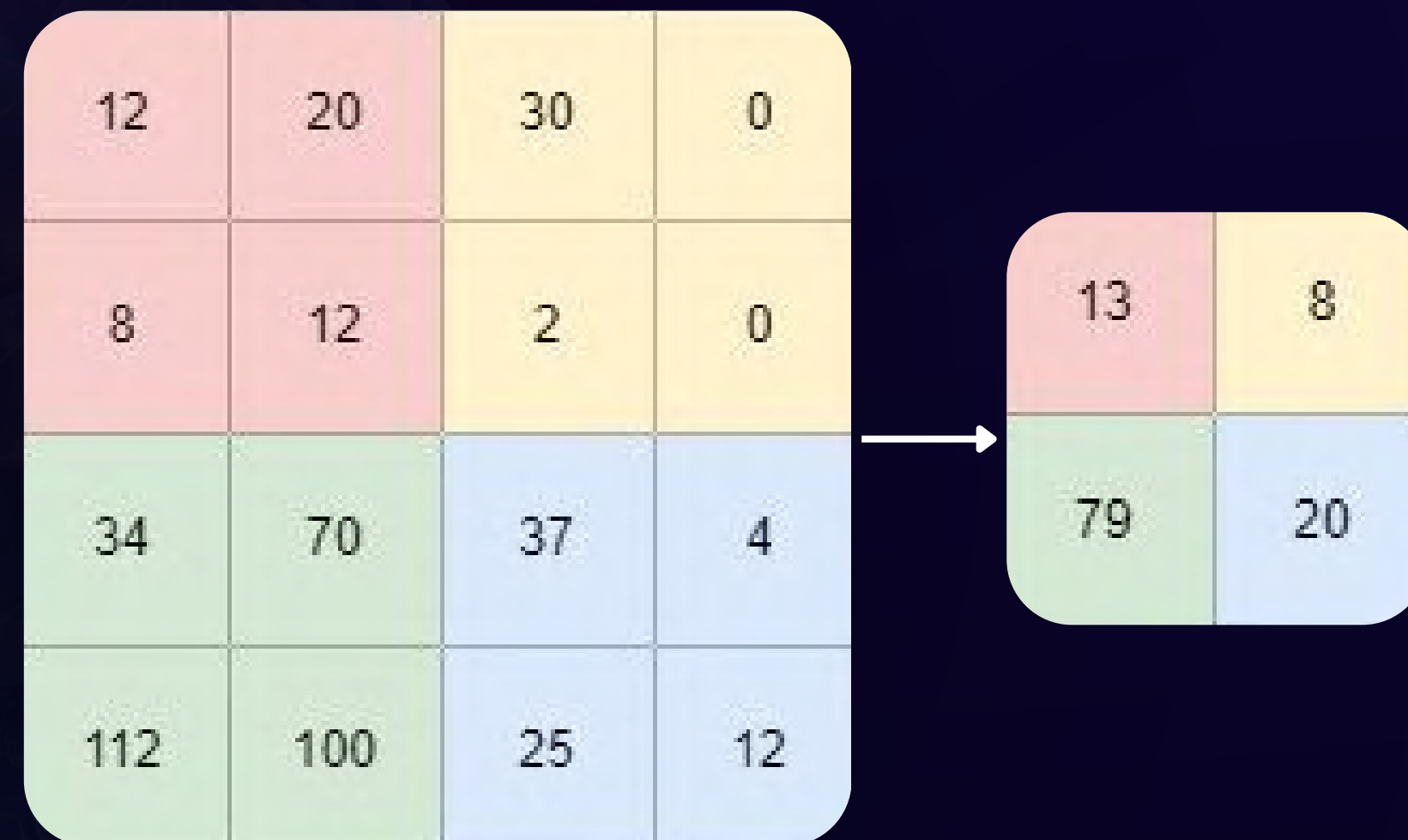
POOLING

MAX POOLING



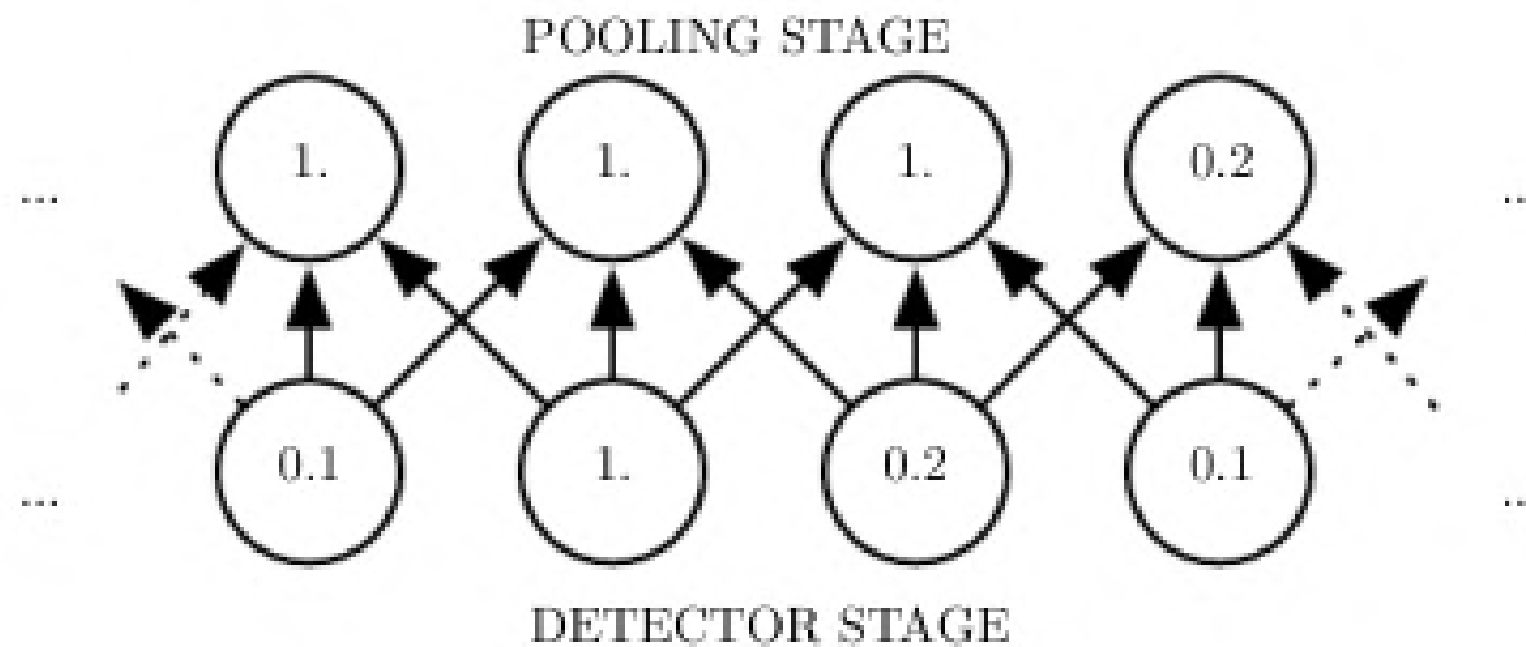
- Preserva maiores ativações, ou seja, o padrão mais presente na região
- Maior contribuição para invariância de translação

AVERAGE POOLING

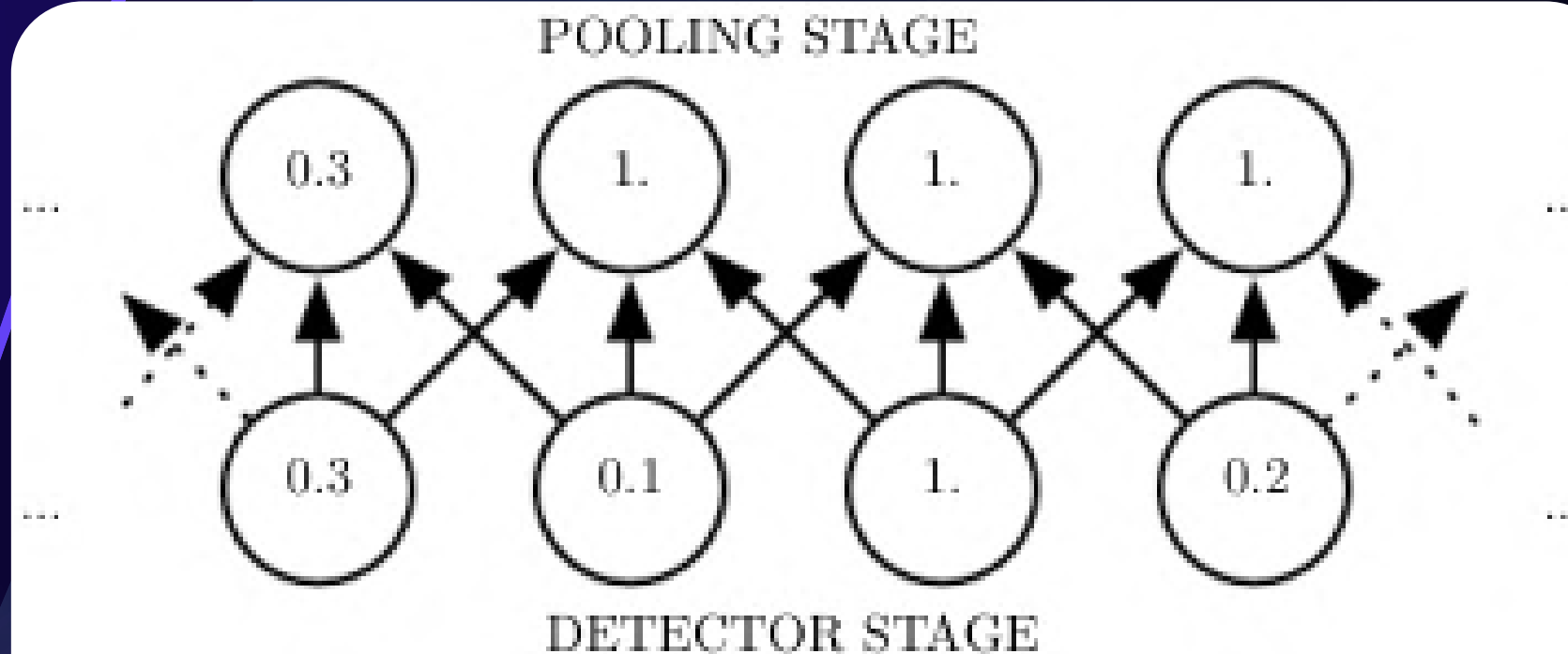


- Fornece representação mais geral da entrada

IMPORTÂNCIA DO MAX POOLING

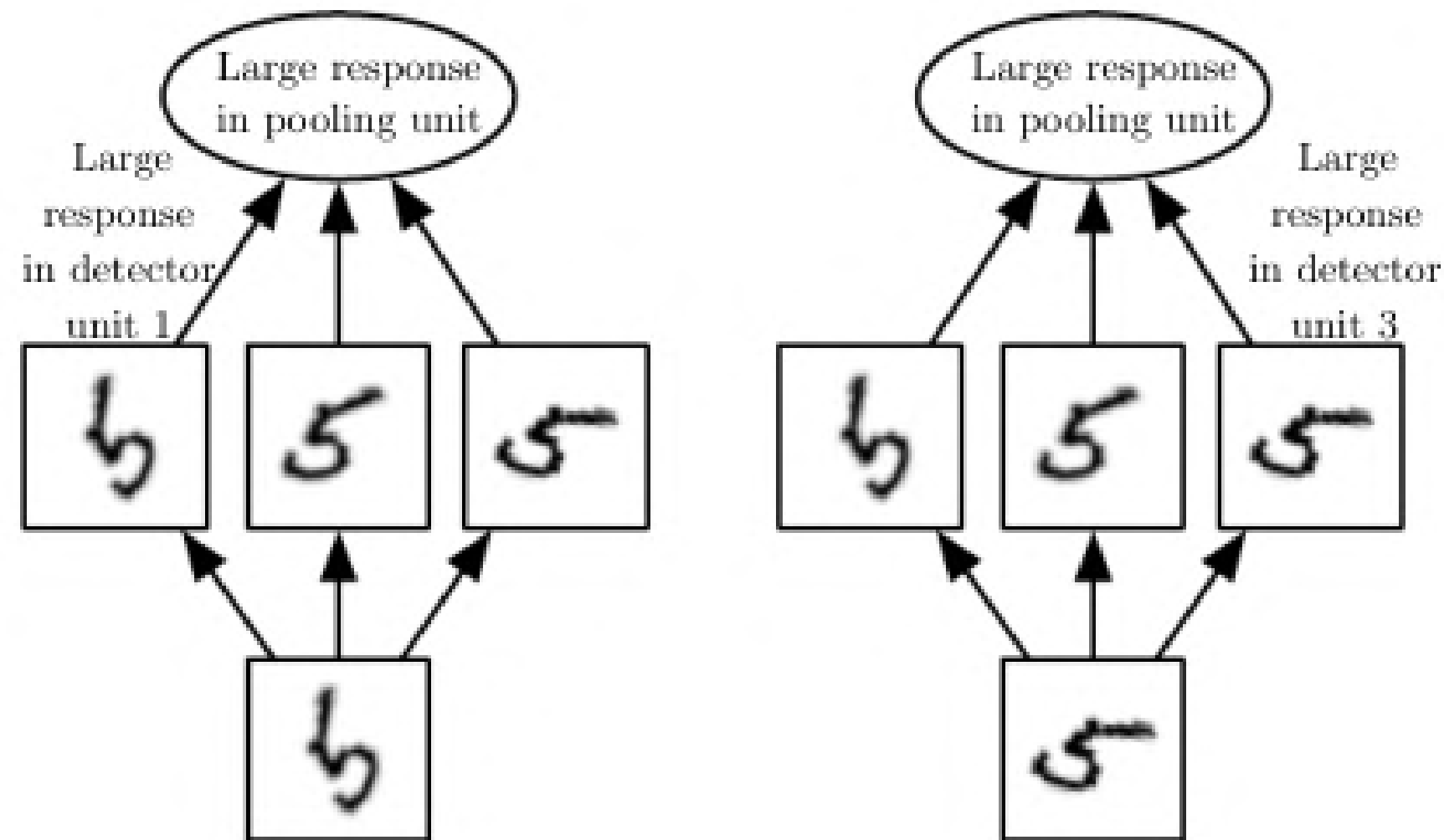


Translação da entrada em um pixel



- Todos os valores resultantes da função de ativação mudaram de posição, mas apenas metade da saída do pooling mudou
- As unidades do max pooling são sensíveis ao máximo valor em uma região, não de uma posição exata

IMPORTÂNCIA DO MAX POOLING



- Todos os 3 filtros foram planejados para detectar o padrão do dígito 5
- Quando a entrada for 5, algum filtro vai identificar e causar uma alta ativação
- Uma vez que o max pooling depende do maior valor em uma região, a saída é alta independentemente da **posição** da unidade ativada

BACKPROPAGATION

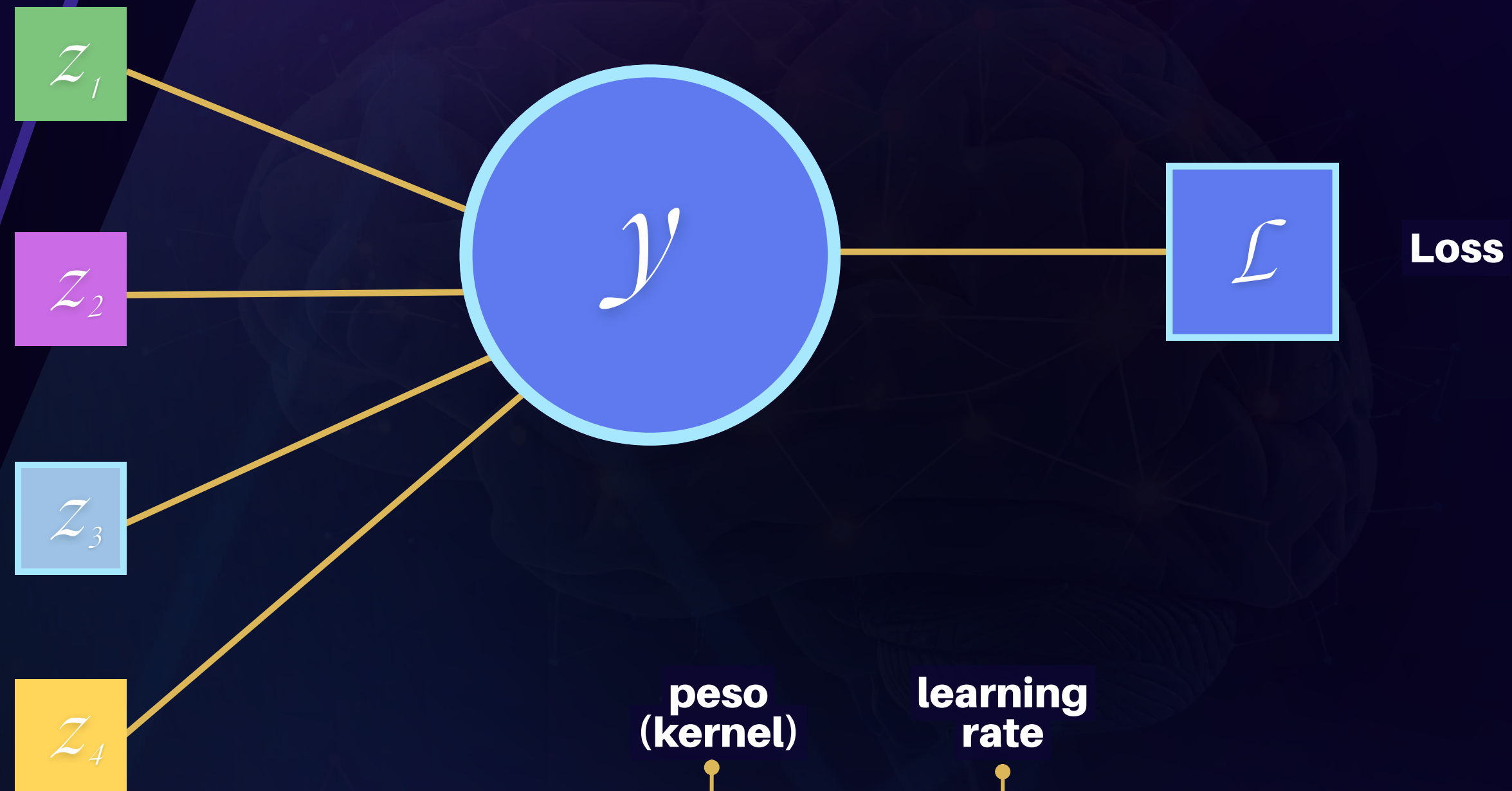
a1	a2	a3	a4	a5
a6	a7	a8	a9	a10
a11	a12	a13	a14	a15
a16	a17	a18	a19	a20
a21	a22	a23	a24	a25

w1	w2	w3
w4	w5	w6
w7	w8	w9

z1	z2
z3	z4

$$z_1 = w_1 \times a_1 + w_2 \times a_2 + w_3 \times a_3 + w_4 \times a_6 \dots w_9 \times a_{13}$$

BACKPROPAGATION



$$w_i^* = \overset{\substack{\text{peso} \\ \text{(kernel)}}}{w_i} - \overset{\substack{\text{learning} \\ \text{rate}}}{\alpha} \times \frac{\partial L}{\partial w_i}$$

BACKPROPAGATION

$w1'$	$w2'$	$w3'$
$w4'$	$w5'$	$w6'$
$w7'$	$w8'$	$w9'$

novo kernel



$w1$	$w2$	$w3$
$w4$	$w5$	$w6$
$w7$	$w8$	$w9$

kernel

learning
rate

$-\alpha \times$

$\frac{\partial L}{\partial w_1}$	$\frac{\partial L}{\partial w_2}$	$\frac{\partial L}{\partial w_3}$
$\frac{\partial L}{\partial w_4}$	$\frac{\partial L}{\partial w_5}$	$\frac{\partial L}{\partial w_6}$
$\frac{\partial L}{\partial w_7}$	$\frac{\partial L}{\partial w_8}$	$\frac{\partial L}{\partial w_9}$

BACKPROPAGATION

w1	w2	w3
w4	w5	w6
w7	w8	w9

z1	z2
z3	z4



$$\frac{\partial L}{\partial w_1} = \frac{\partial z_1}{\partial w_1} \frac{\partial \hat{y}}{\partial z_1} \frac{\partial L}{\partial \hat{y}} + \frac{\partial z_2}{\partial w_1} \frac{\partial L}{\partial z_2} + \frac{\partial z_3}{\partial w_1} \frac{\partial L}{\partial z_3} + \frac{\partial z_4}{\partial w_1} \frac{\partial L}{\partial z_4}$$

BACKPROPAGATION


$$\frac{\partial L}{\partial w_1} = \frac{\partial z_1}{\partial w_1} \frac{\partial L}{\partial z_1} + \frac{\partial z_2}{\partial w_1} \frac{\partial L}{\partial z_2} + \frac{\partial z_3}{\partial w_1} \frac{\partial L}{\partial z_3} + \frac{\partial z_4}{\partial w_1} \frac{\partial L}{\partial z_4}$$

$$z_1 = w_1 \times a_1 + w_2 \times a_2 + w_3 \times a_3 + w_4 \times a_6 \dots w_9 \times a_{13}$$

$$z_2 = w_1 \times a_3 + w_2 \times a_4 + w_3 \times a_5 + w_4 \times a_8 \dots w_9 \times a_{15}$$

$$z_3 = w_1 \times a_{11} + w_2 \times a_{12} + w_3 \times a_{13} + w_4 \times a_{16} \dots w_9 \times a_{23}$$

$$z_4 = w_1 \times a_{13} + w_2 \times a_{14} + w_3 \times a_{15} + w_4 \times a_{18} \dots w_9 \times a_{25}$$

$$\frac{\partial L}{\partial w_1} = a_1 \frac{\partial L}{\partial z_1} + a_3 \frac{\partial L}{\partial z_2} + a_{11} \frac{\partial L}{\partial z_3} + a_{13} \frac{\partial L}{\partial z_4}$$


BACKPROPAGATION

a1	a2	a3	a4	a5
a6	a7	a8	a9	a10
a11	a12	a13	a14	a15
a16	a17	a18	a19	a20
a21	a22	a23	a24	a25

$\frac{\partial L}{\partial z_1}$	$\frac{\partial L}{\partial z_2}$
$\frac{\partial L}{\partial z_3}$	$\frac{\partial L}{\partial z_4}$

$\frac{\partial L}{\partial w_1}$	$\frac{\partial L}{\partial w_2}$	$\frac{\partial L}{\partial w_3}$
$\frac{\partial L}{\partial w_4}$	$\frac{\partial L}{\partial w_5}$	$\frac{\partial L}{\partial w_6}$
$\frac{\partial L}{\partial w_7}$	$\frac{\partial L}{\partial w_8}$	$\frac{\partial L}{\partial w_9}$

$$\begin{aligned}\frac{\partial L}{\partial w_1} &= a_1 \frac{\partial L}{\partial z_1} + a_3 \frac{\partial L}{\partial z_2} + a_{11} \frac{\partial L}{\partial z_3} + a_{13} \frac{\partial L}{\partial z_4} \\ \frac{\partial L}{\partial w_2} &= a_2 \frac{\partial L}{\partial z_1} + a_4 \frac{\partial L}{\partial z_2} + a_{12} \frac{\partial L}{\partial z_3} + a_{14} \frac{\partial L}{\partial z_4} \\ \frac{\partial L}{\partial w_3} &= a_3 \frac{\partial L}{\partial z_1} + a_5 \frac{\partial L}{\partial z_2} + a_{13} \frac{\partial L}{\partial z_3} + a_{15} \frac{\partial L}{\partial z_4} \\ \frac{\partial L}{\partial w_4} &= a_6 \frac{\partial L}{\partial z_1} + a_8 \frac{\partial L}{\partial z_2} + a_{16} \frac{\partial L}{\partial z_3} + a_{18} \frac{\partial L}{\partial z_4} \\ \frac{\partial L}{\partial w_5} &= a_7 \frac{\partial L}{\partial z_1} + a_9 \frac{\partial L}{\partial z_2} + a_{17} \frac{\partial L}{\partial z_3} + a_{19} \frac{\partial L}{\partial z_4} \\ \frac{\partial L}{\partial w_6} &= a_8 \frac{\partial L}{\partial z_1} + a_{10} \frac{\partial L}{\partial z_2} + a_{18} \frac{\partial L}{\partial z_3} + a_{20} \frac{\partial L}{\partial z_4} \\ \frac{\partial L}{\partial w_7} &= a_{11} \frac{\partial L}{\partial z_1} + a_{13} \frac{\partial L}{\partial z_2} + a_{21} \frac{\partial L}{\partial z_3} + a_{23} \frac{\partial L}{\partial z_4} \\ \frac{\partial L}{\partial w_8} &= a_{12} \frac{\partial L}{\partial z_1} + a_{14} \frac{\partial L}{\partial z_2} + a_{22} \frac{\partial L}{\partial z_3} + a_{24} \frac{\partial L}{\partial z_4} \\ \frac{\partial L}{\partial w_9} &= a_{13} \frac{\partial L}{\partial z_1} + a_{15} \frac{\partial L}{\partial z_2} + a_{23} \frac{\partial L}{\partial z_3} + a_{25} \frac{\partial L}{\partial z_4}\end{aligned}$$

Medidas de desempenho

- **Classe positiva**

Representa o resultado de interesse, e responde a pergunta: Esse é o resultado que eu procurava?

- **Classe negativa**

É a ausência da classe negativa (ex: o paciente não está doente)

- **Legenda**

- VP (Verdadeiro Positivo): Acertos corretos da classe positiva.
- VN (Verdadeiro Negativo): Acertos corretos da classe negativa.
- FP (Falso Positivo): Erros onde o modelo previu positivo, mas era negativo.
- FN (Falso Negativo): Erros onde o modelo previu negativo, mas era positivo.

Medidas de desempenho

- **Acurácia**

Proporção total de predições corretas (acertos gerais) sobre o número total de observações.

$$Acurácia = \frac{VP + VN}{VP + VN + FP + FN}$$

- **Precisão**

Responde à pergunta: "Das vezes que o modelo previu positivo, quantas vezes ele acertou?"

$$Precisão = \frac{VP}{VP + FP}$$

- **Recall**

Responde à pergunta: "De todos os casos positivos reais, quantos o modelo conseguiu encontrar?"

$$Recall = \frac{VP}{VP + FN}$$

- **F1 score**

É a média harmônica da Precisão e do Recall. Oferece um equilíbrio entre essas duas métricas em um único valor.

$$F1-Score = 2 \cdot \frac{Precisão \cdot Recall}{Precisão + Recall}$$

IMPLEMENTAÇÃO



REDES CONVOLUCIONAIS