PROCESSAMENTO DE LINGUAGEM NATURAL

Redes Neurais Profundas

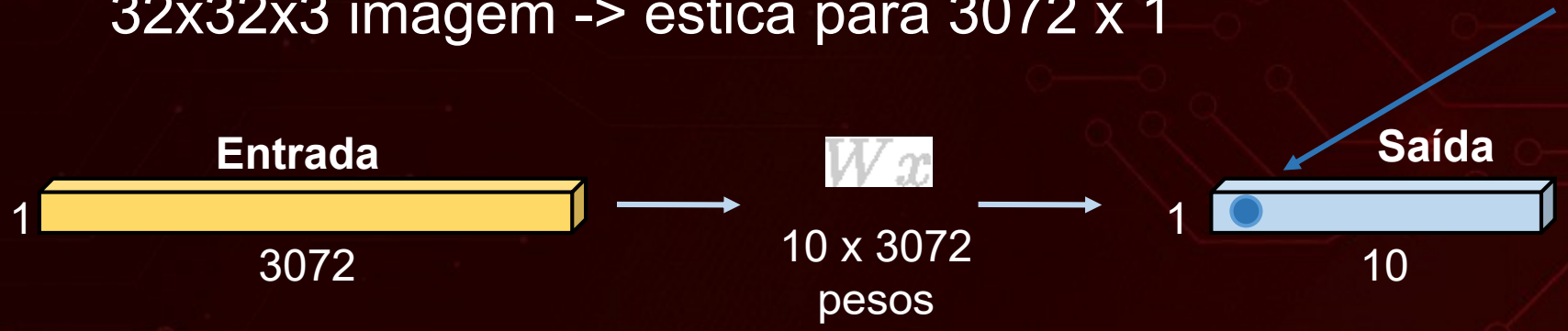
TÓPICOS

1. Introdução
2. Camadas de convolução e *pooling*
3. CNN para classificação de textos

REDES NEURAIS

Camadas totalmente conectadas (Fully-Connected Layers)

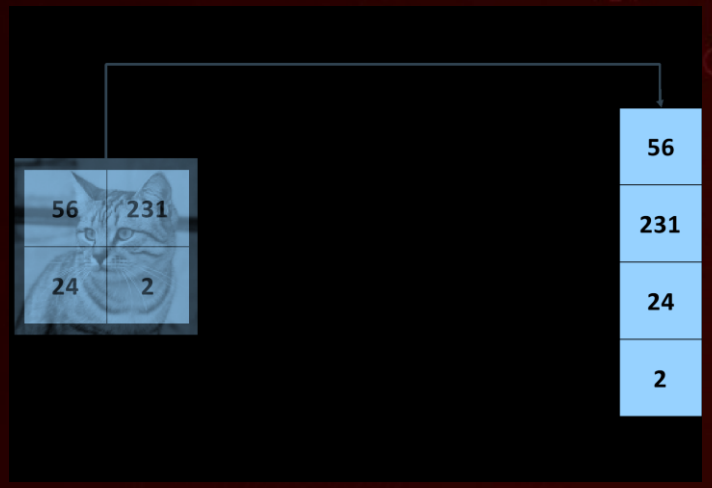
32x32x3 imagem -> estica para 3072 x 1



1 número: **o** resultado do produto interno entre uma linha de W e a entrada (um produto interno de 3072 dimensões)

Problema: os classificadores não respeitam a estrutura espacial de imagens.

Solução: definir novos nós computacionais que operem em imagens.



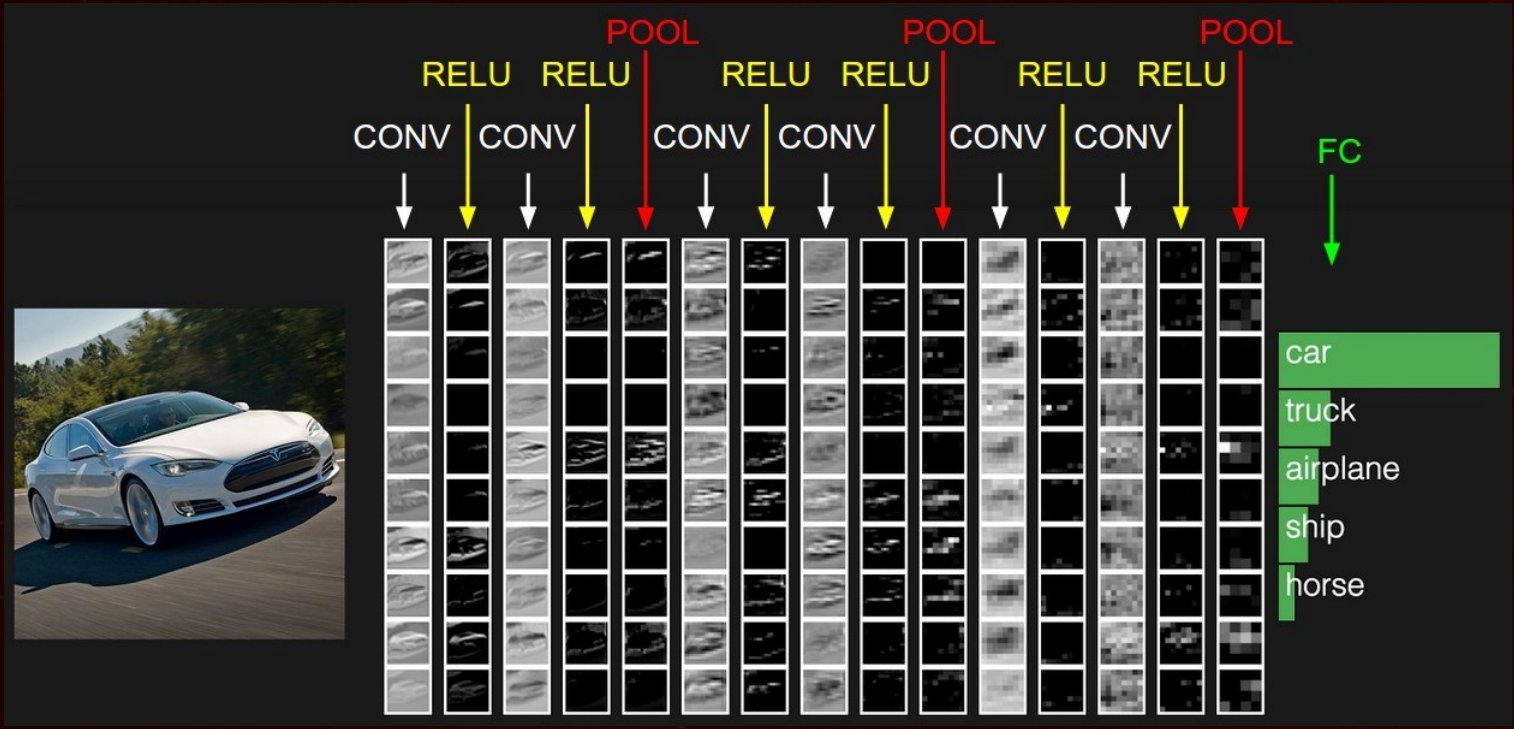
O QUE É RN CONVOLUCIONAL?

- “Rede Neural *feed-forward* cujos neurônios estão dispostos de tal forma a responder por regiões sobrepostas (campos receptores) que preenchem o campo visual”

- “Múltiplas camadas de pequenas coleções de neurônios que processam porções da imagem de entrada (campos receptores)”

- “Sequências de camadas de convolução intercaladas por funções de ativação”

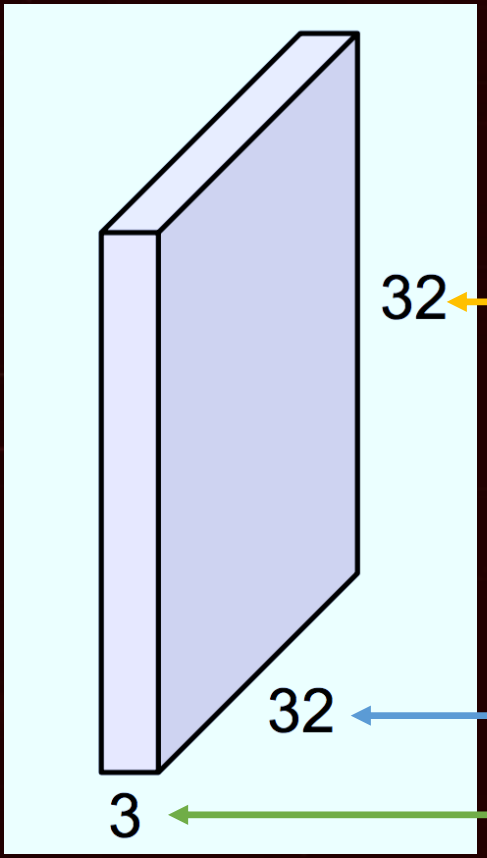
REDE NEURAL CONVOLUCIONAL



Começaremos com a camada de convolução

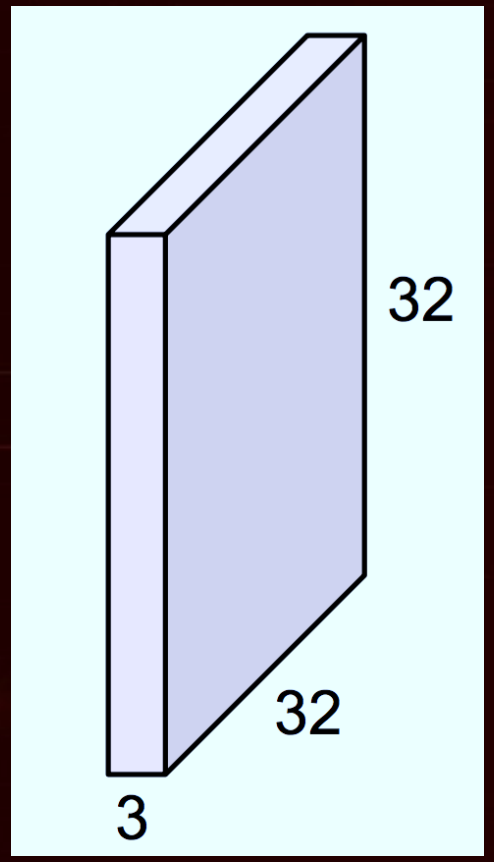
CAMADA DE CONVOLUÇÃO

Imagem $32 \times 32 \times 3$

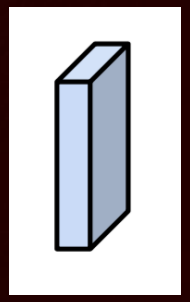


CAMADA DE CONVOLUÇÃO

Imagem $32 \times 32 \times \mathbf{3}$



Filtro $5 \times 5 \times \mathbf{3}$

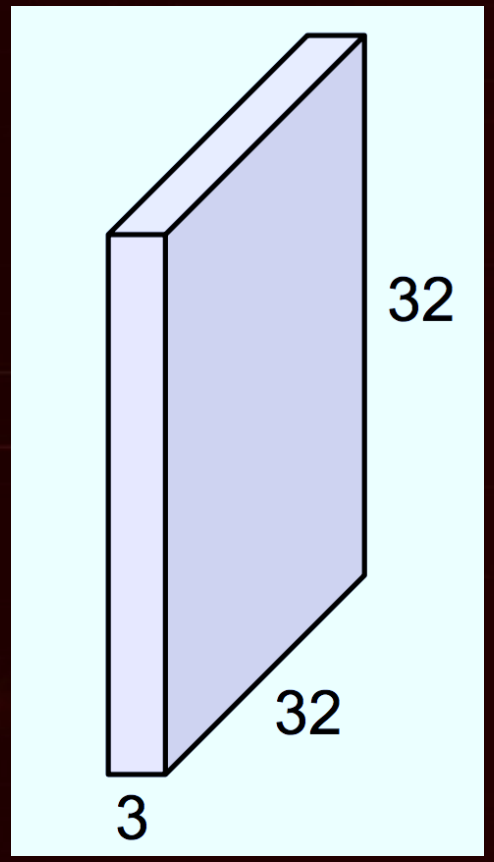


Convoluir o filtro com a imagem

Deslizar espacialmente o filtro pela imagem, computando produtos internos

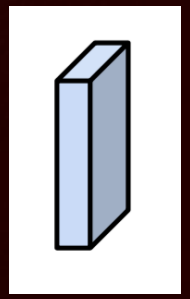
CAMADA DE CONVOLUÇÃO

Imagem $32 \times 32 \times \mathbf{3}$



Filtros sempre compreendem a profundidade inteira do volume de entrada

Filtro $5 \times 5 \times \mathbf{3}$



Convoluir o filtro com a imagem

Deslizar espacialmente o filtro pela imagem, computando produtos internos

CAMADA DE CONVOLUÇÃO

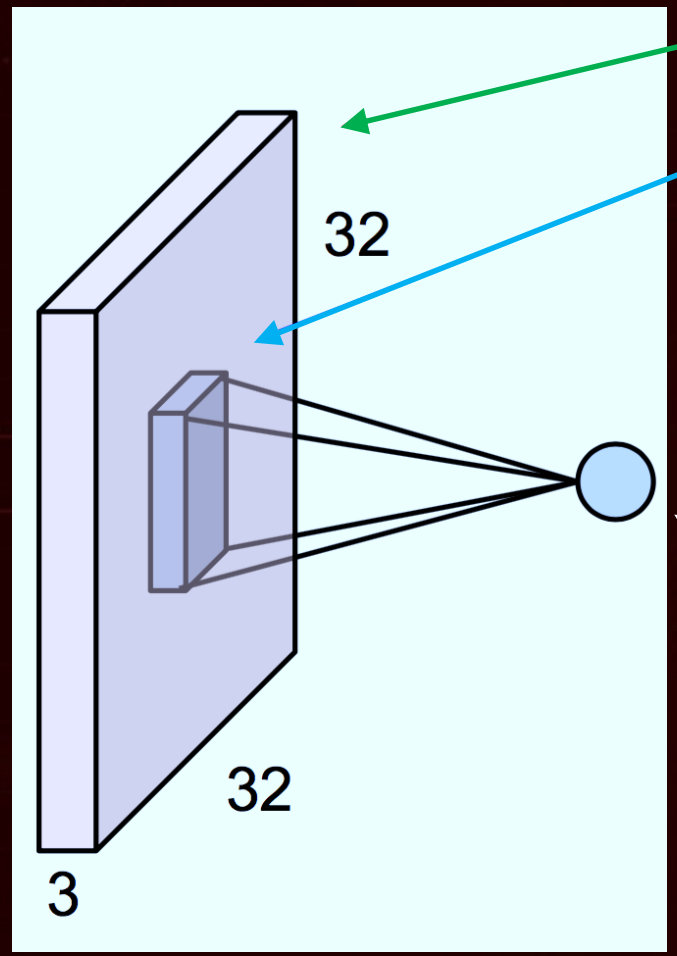
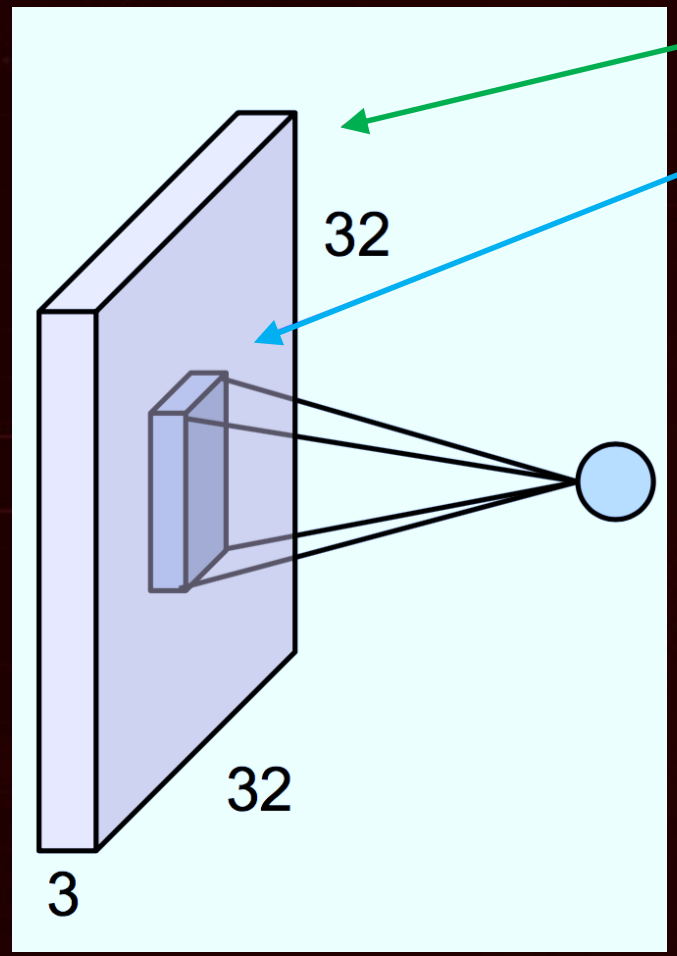


Imagem $(32 \times 32 \times 3)$

Filtro $w(5 \times 5 \times 3)$

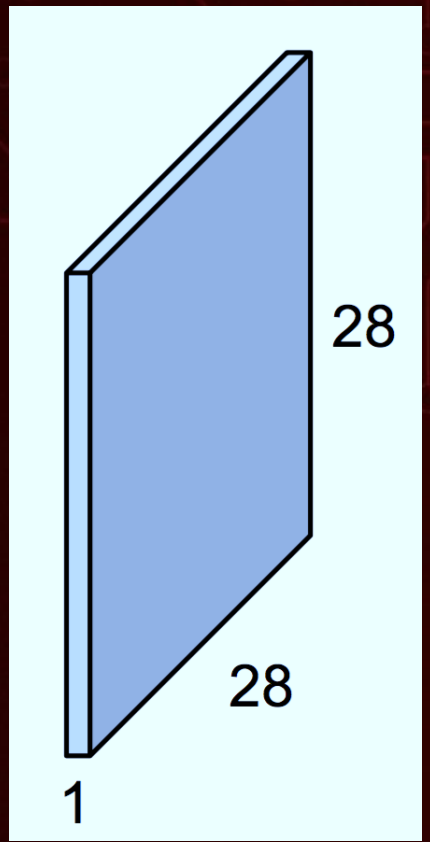
1 número:
resultado do produto interno entre o filtro e uma porção $5 \times 5 \times 3$ da imagem
$(5 \times 5 \times 3 = 75$ parâmetros + bias)

CAMADA DE CONVOLUÇÃO



convoluir (deslizar) sobre todas as regiões espaciais

mapa de ativação



CAMADA DE CONVOLUÇÃO

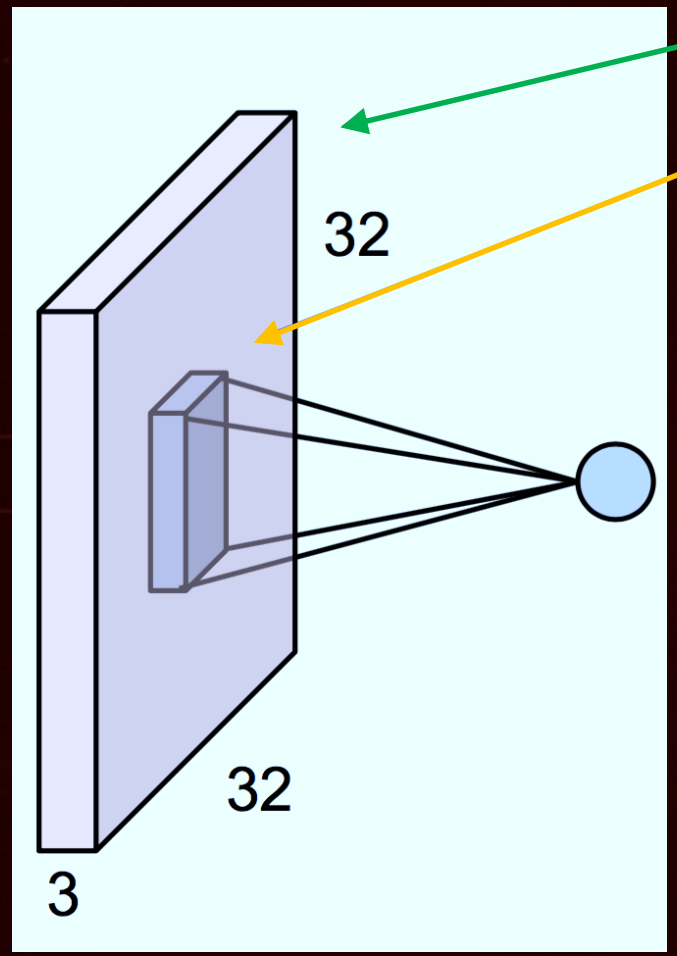
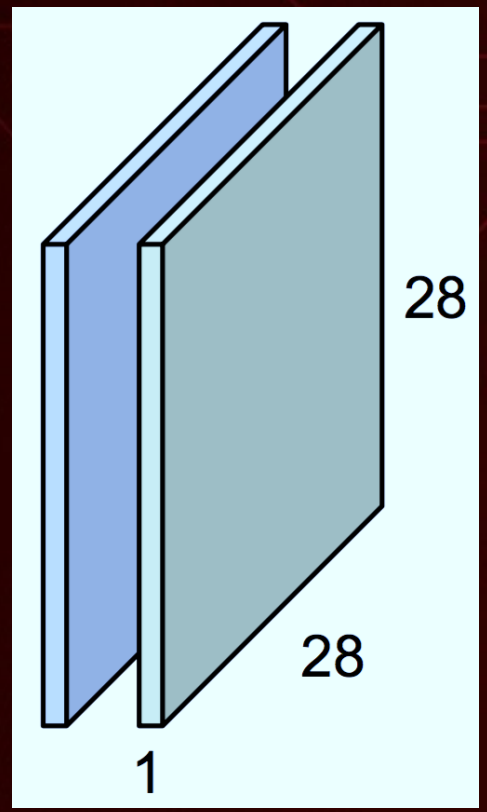


Imagem $(32 \times 32 \times 3)$

Filtro $w_2(5 \times 5 \times 3)$

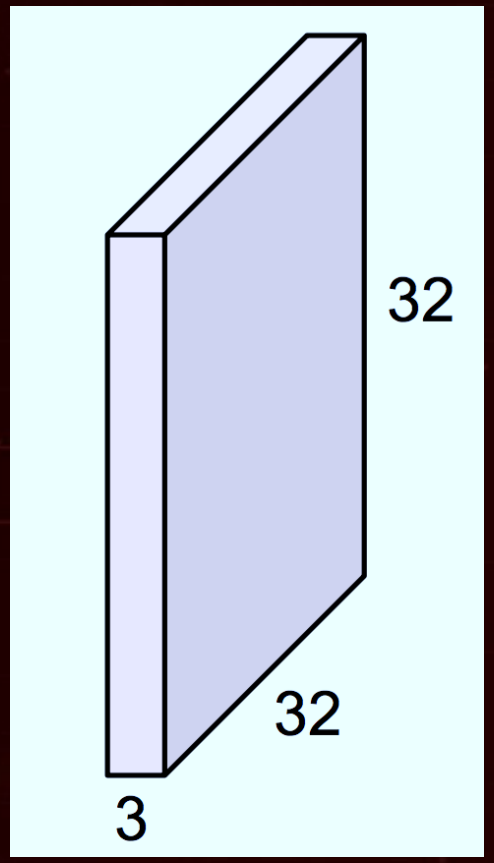
convoluir (deslizar) sobre todas as regiões espaciais

mapa de ativação

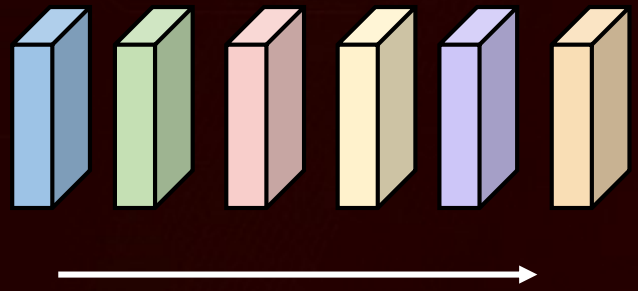


considere a existência de um segundo filtro (**amarelo**)

CAMADA DE CONVOLUÇÃO

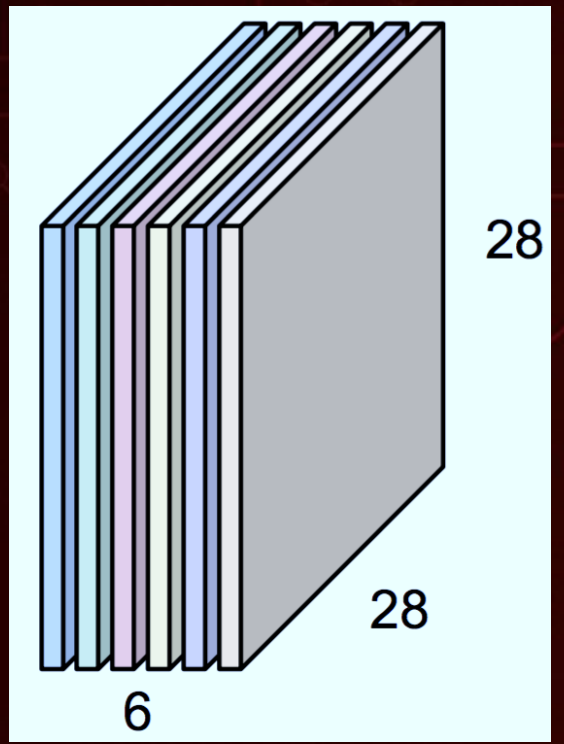


Se tivermos 6 filtros $5 \times 5 \times 3$, teremos 6 *featuremaps* distintos:



Camada de Convolução

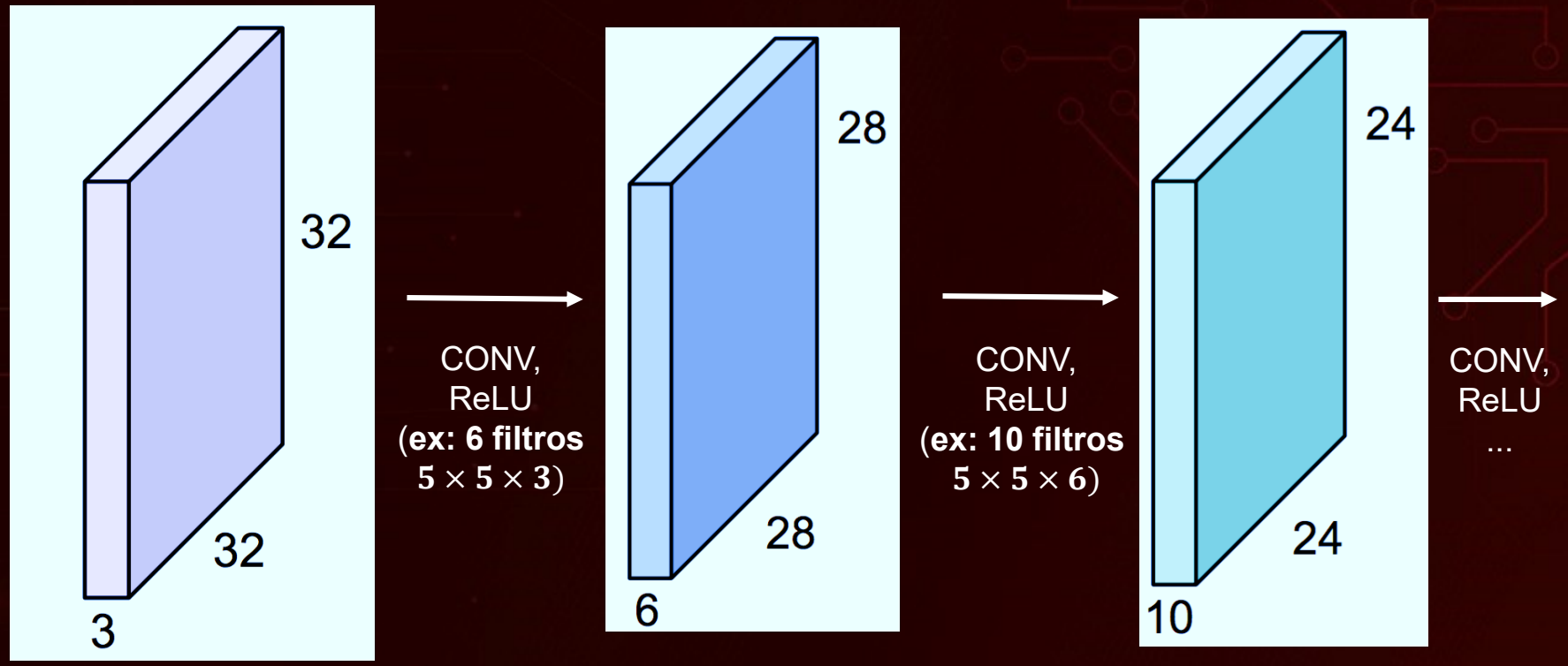
mapas de ativação



Empilhamos o *featuremaps* para gerar uma ”nova imagem” de tamanho $28 \times 28 \times 6$

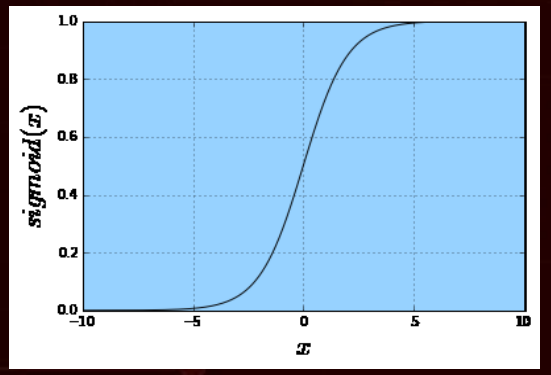
REDE CONVOLUCIONAL

"Sequência de camadas de convolução intercaladas por funções de ativação"



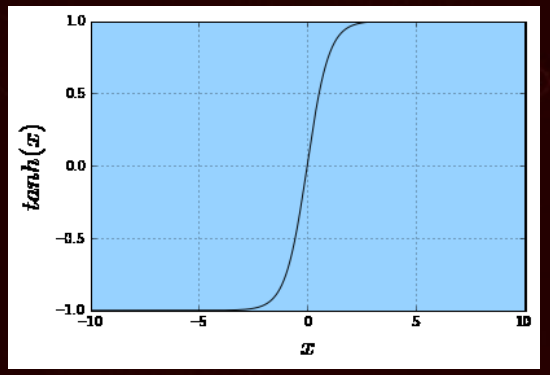
FUNÇÕES DE ATIVAÇÃO

Sigmoid 



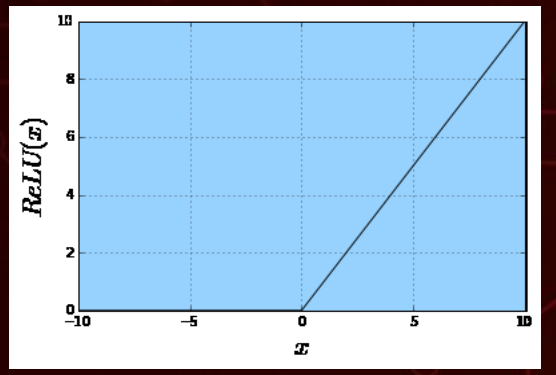
Interpretação da saída por taxa média de disparo
Tende a saturar.
Mata os gradientes.
Não é centrada em zero.
Custosa.
Ruim para redes profundas

Tanh 



Centrada em zero.
Mata os gradientes.
Ruim para redes profundas.

ReLU

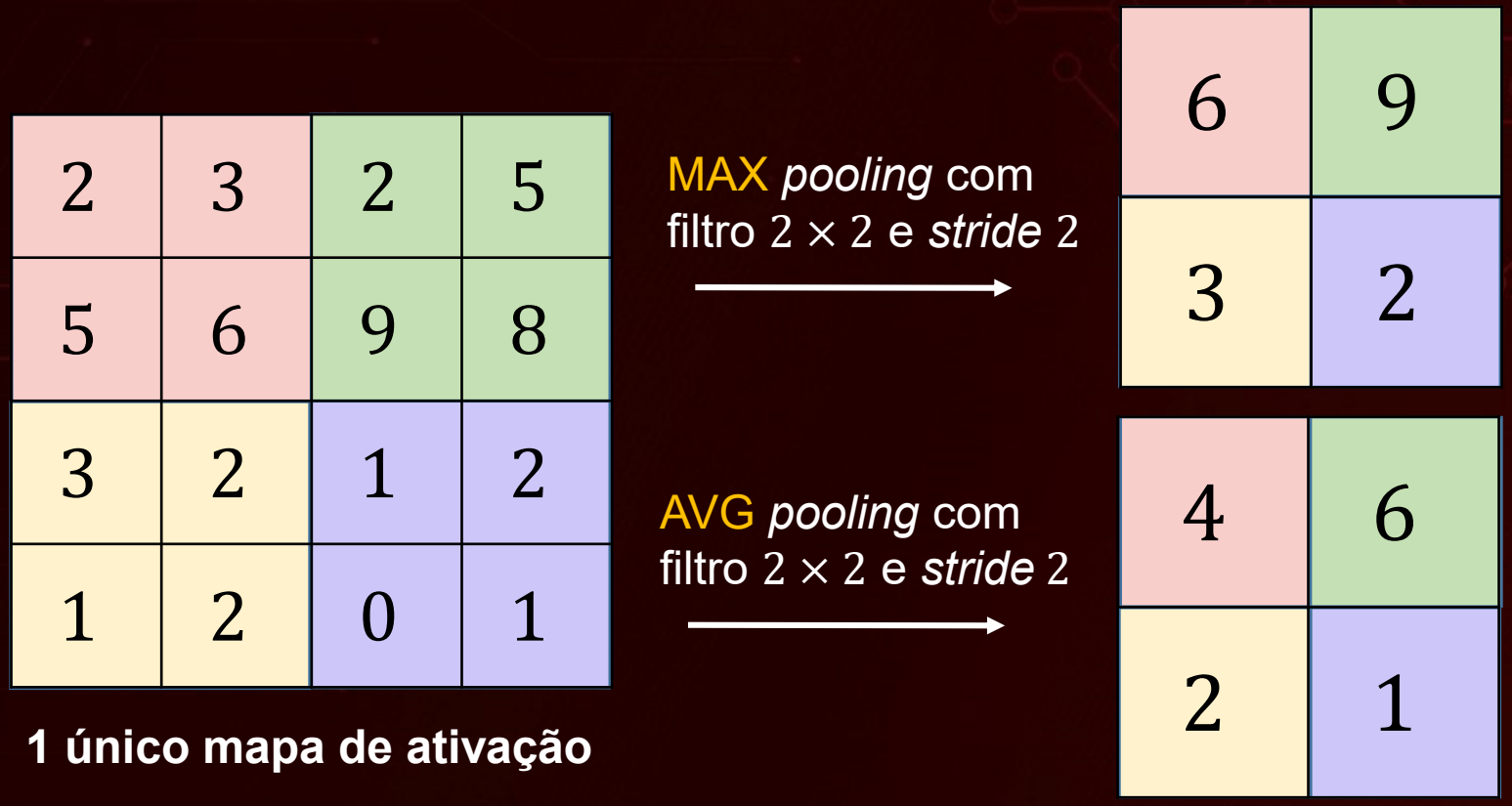


Evita saturação.
Boa para redes profundas.
Convergência rápida.
Eficiente.
Não centrada em zero.

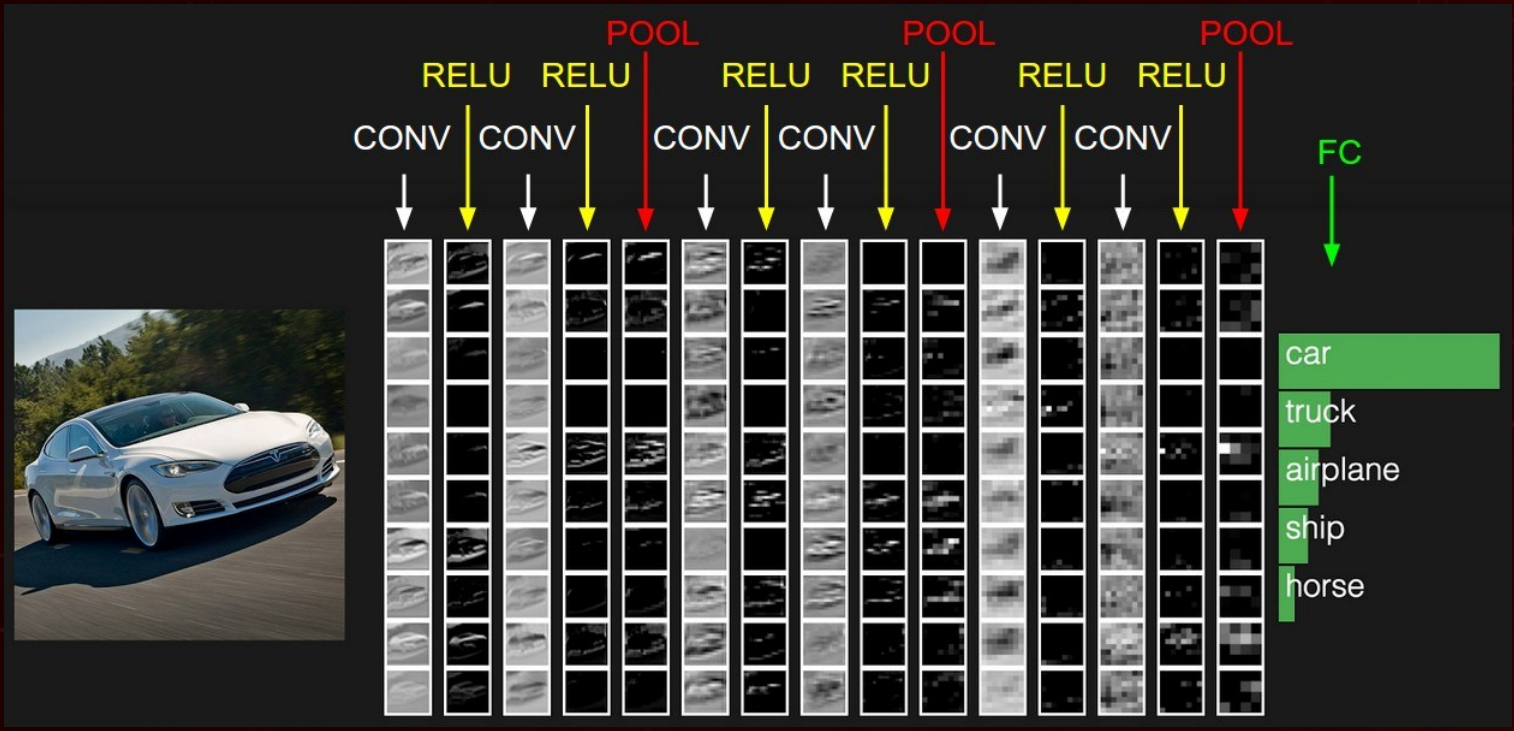
Leaky ReLU
PReLU
ELU

CAMADA DE POOLING

- **Reduz os volumes** para que se tornem gerenciáveis
- Atua sobre cada mapa de ativação de **maneira independente**



REDE NEURAL CONVOLUCIONAL



Uma rede convolucional pode conter uma ou várias camadas totalmente conectadas, e nelas os neurônios estão conectados com toda a entrada (redes neurais tradicionais).

Finalmente: FC (Fully-Connected Layer)

TEXTO COM CNN

- **Embora originalmente concebidas para processamento de imagens, CNNs também são frequentemente utilizadas para processamento de texto...**

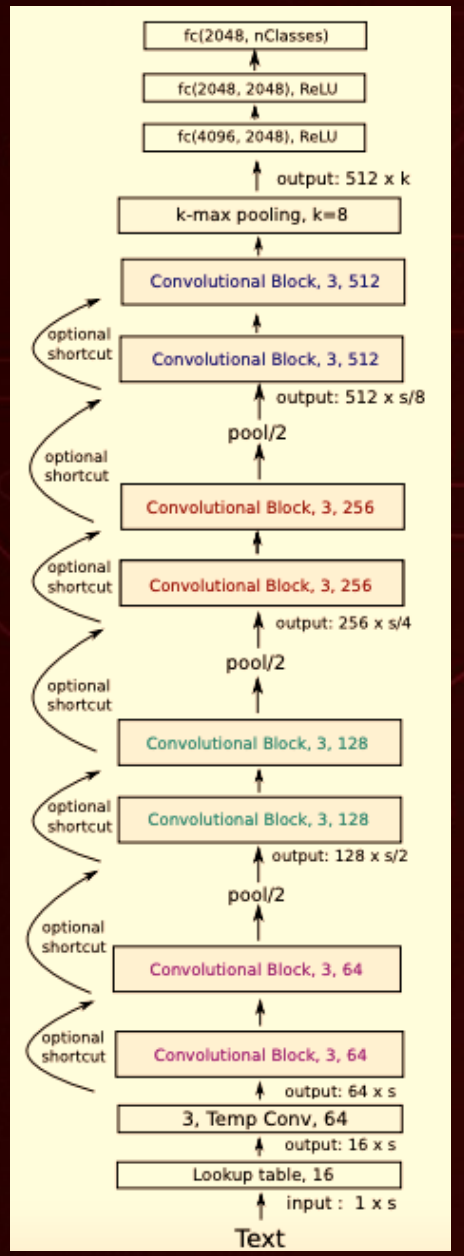


Imagem do paper “Very Deep Convolutional Networks for Text Classification” de Conneau et al., EACL 2017

PLN COM CNN

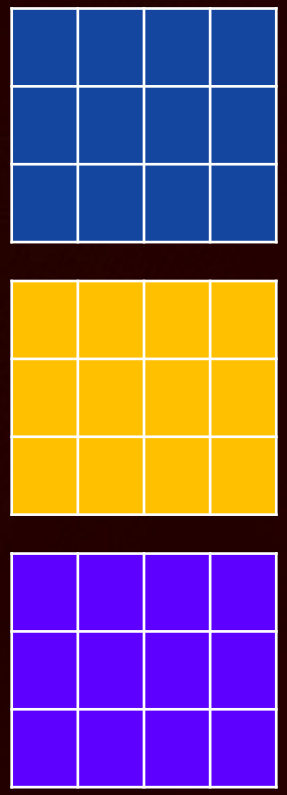
Entrada $t \times d$

0	0	0	0
two			
people			
riding			
skis			
on			
a			
snowy			
slope			
0	0	0	0

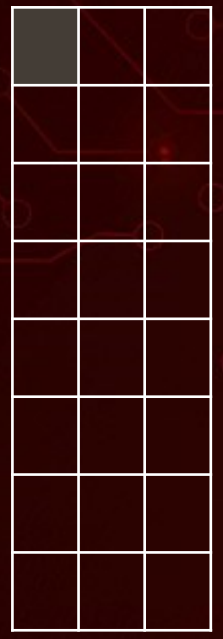
$d = 4$

$t = 8 + 2$ (pad)

3 filtros de tamanho $3 \times d$

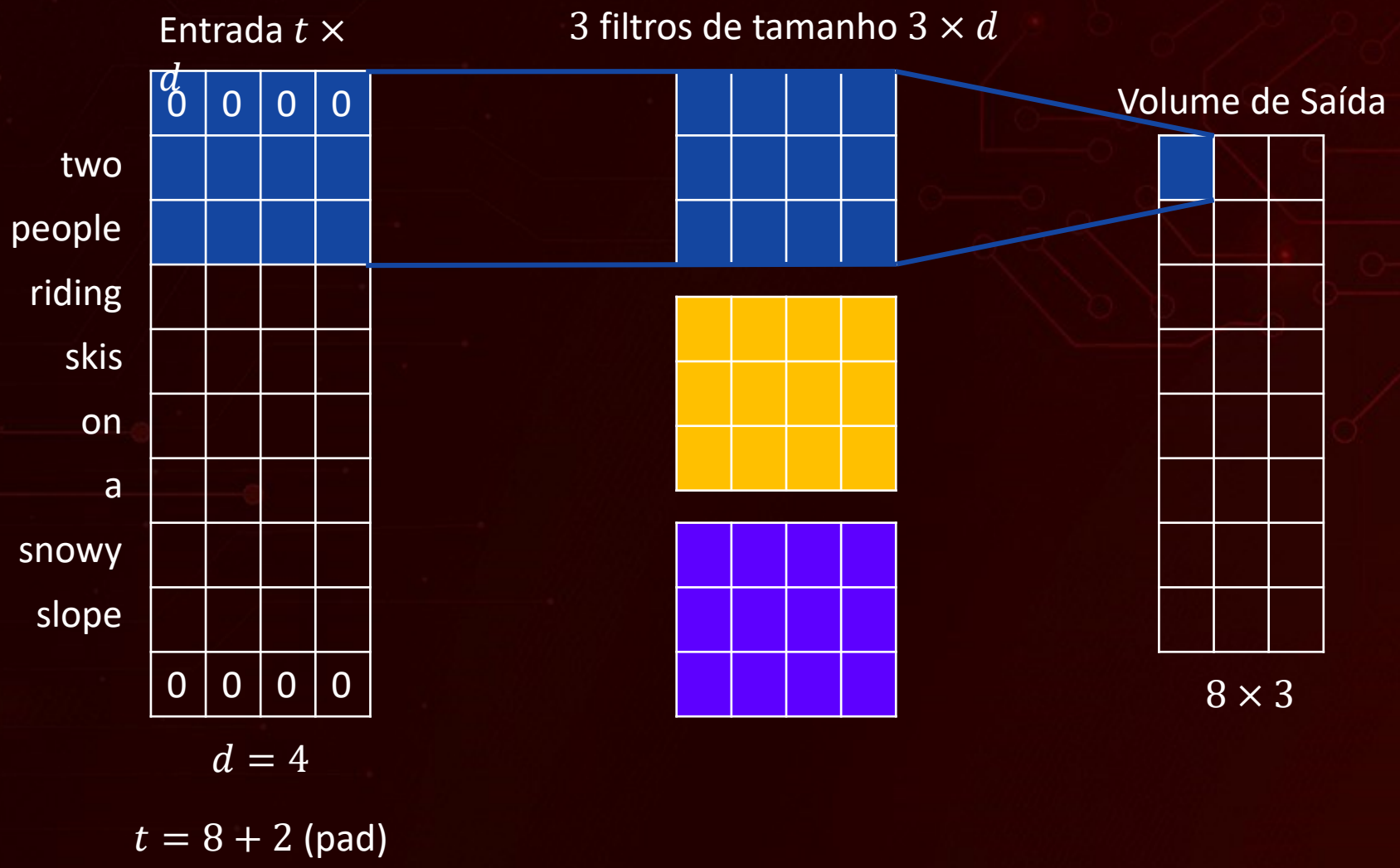


Volume de Saída

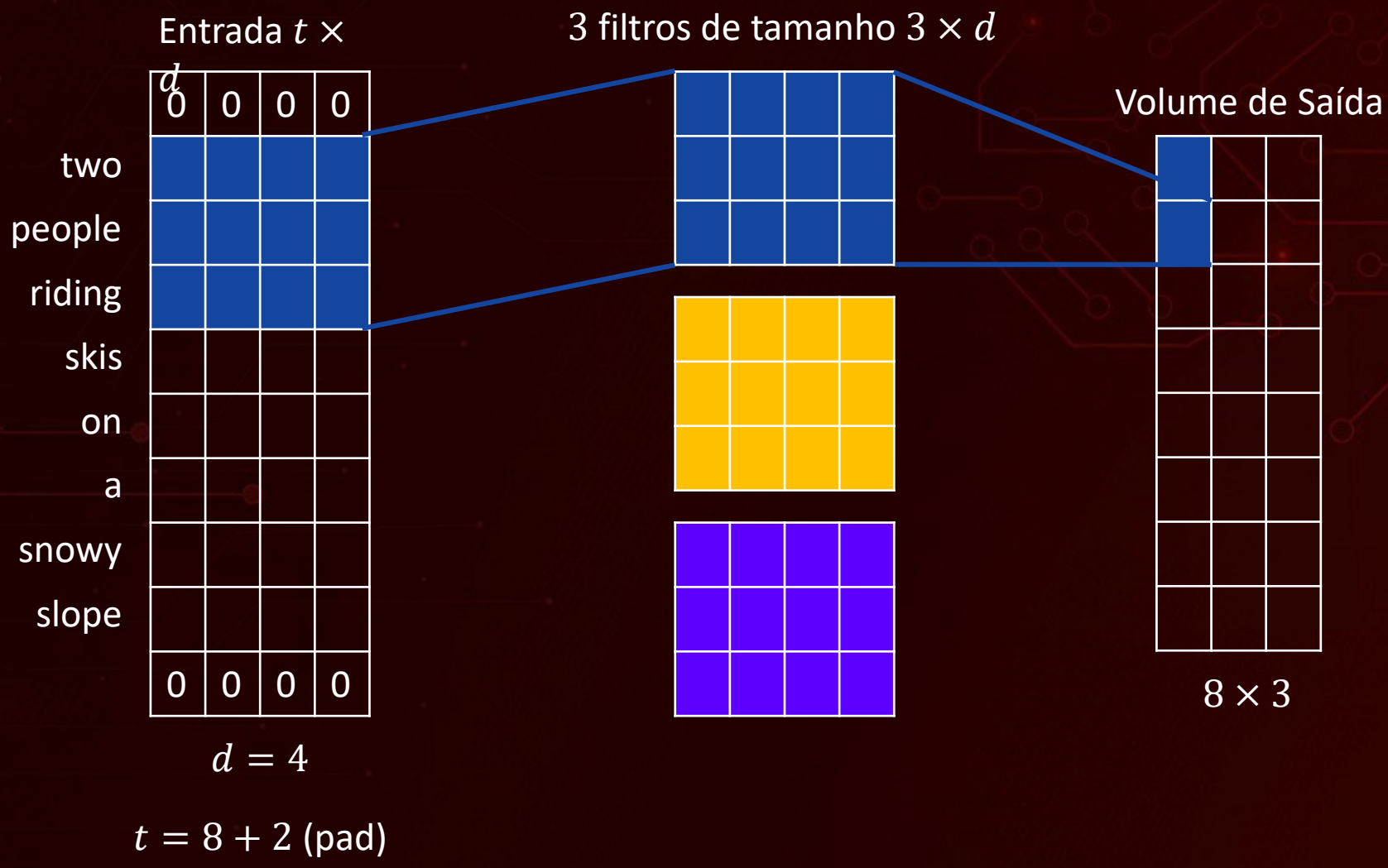


8×3

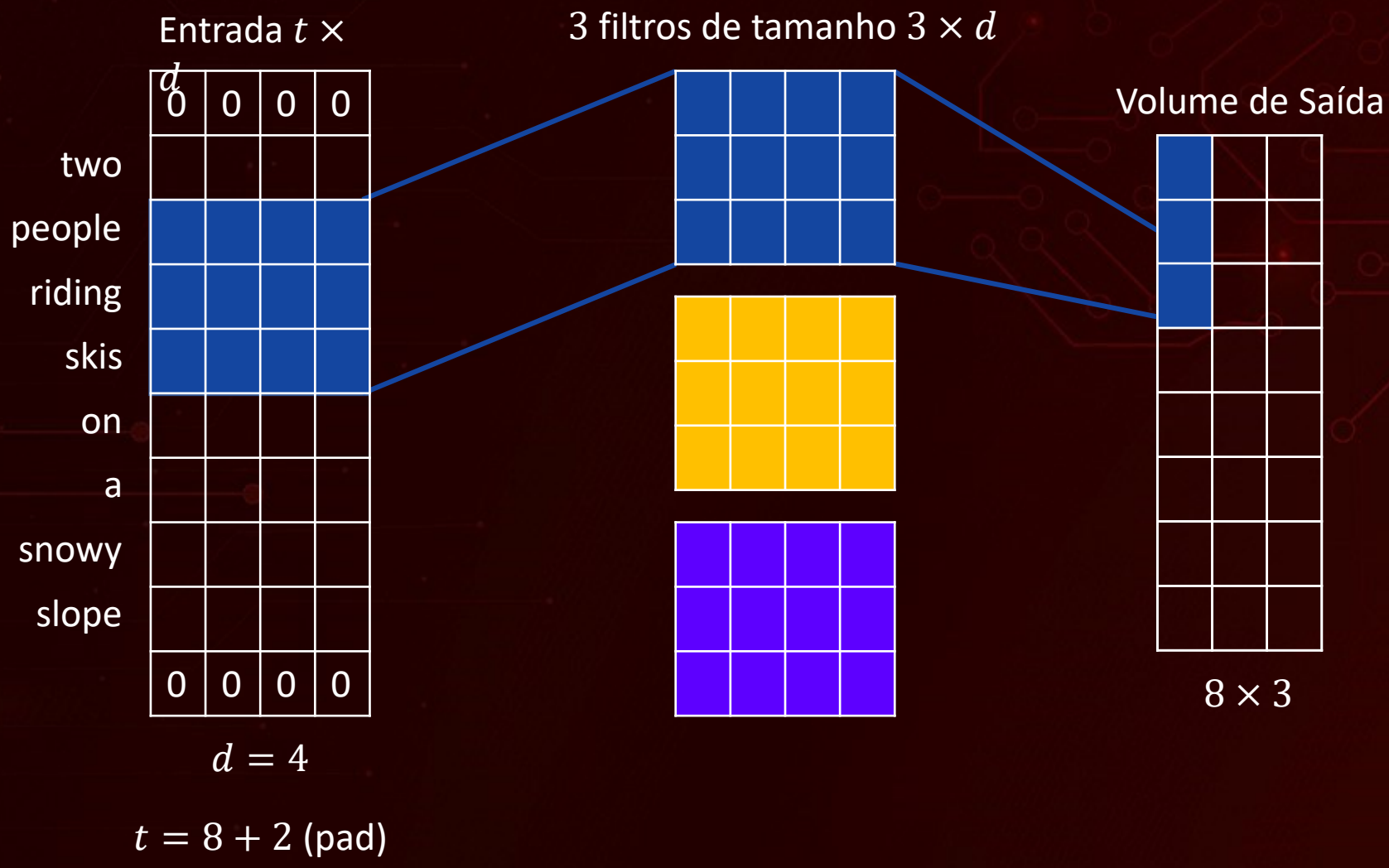
PLN COM CNN


Entrada $t \times d$
3 filtros de tamanho $3 \times d$
Volume de Saída
0 0 0 0
two
people
riding
skis
on
a
snowy
slope
0 0 0 0
$d = 4$
$t = 8 + 2$ (pad)
8×3

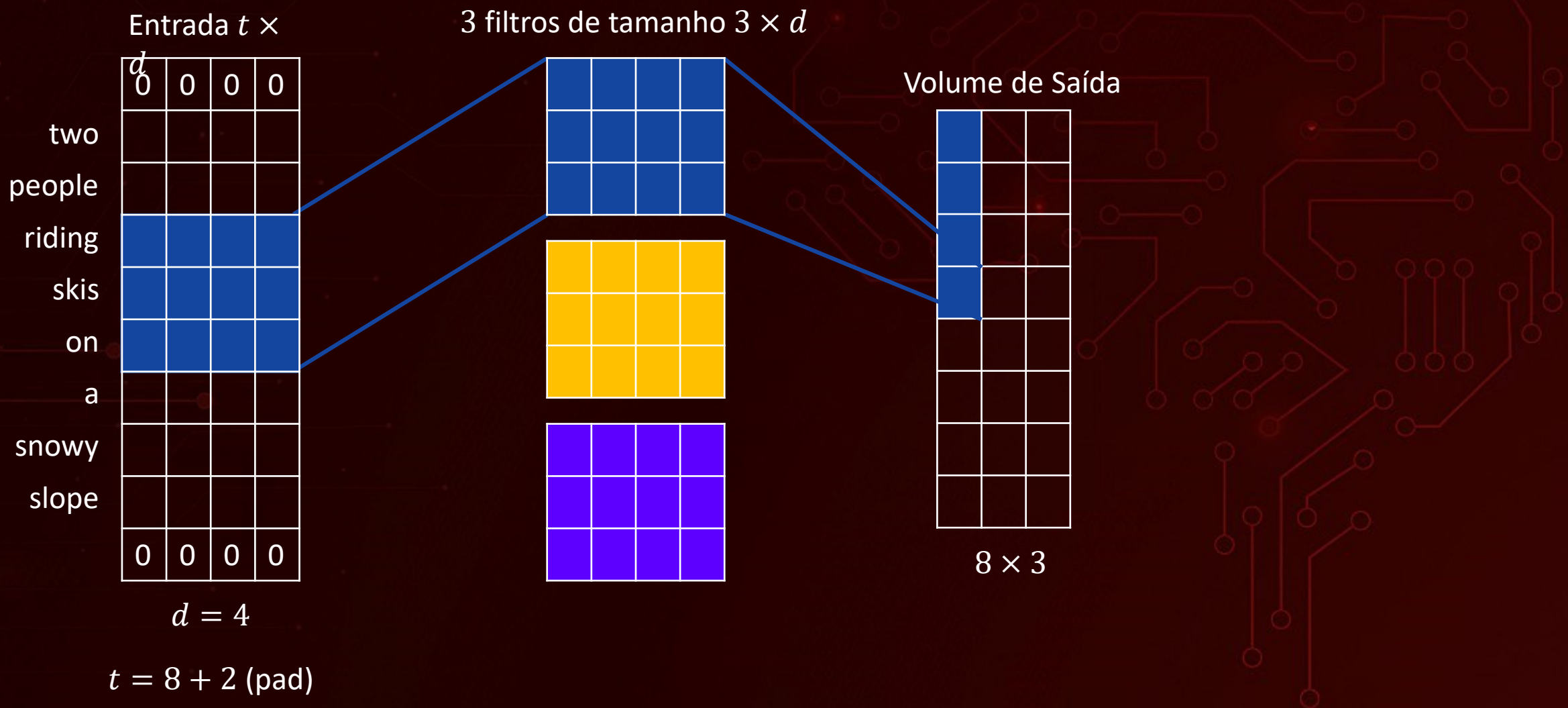
PLN COM CNN


Entrada $t \times d$
0 0 0 0
two
people
riding
skis
on
a
snowy
slope
0 0 0 0
$d = 4$
$t = 8 + 2$ (pad)
3 filtros de tamanho $3 \times d$
Volume de Saída
8×3

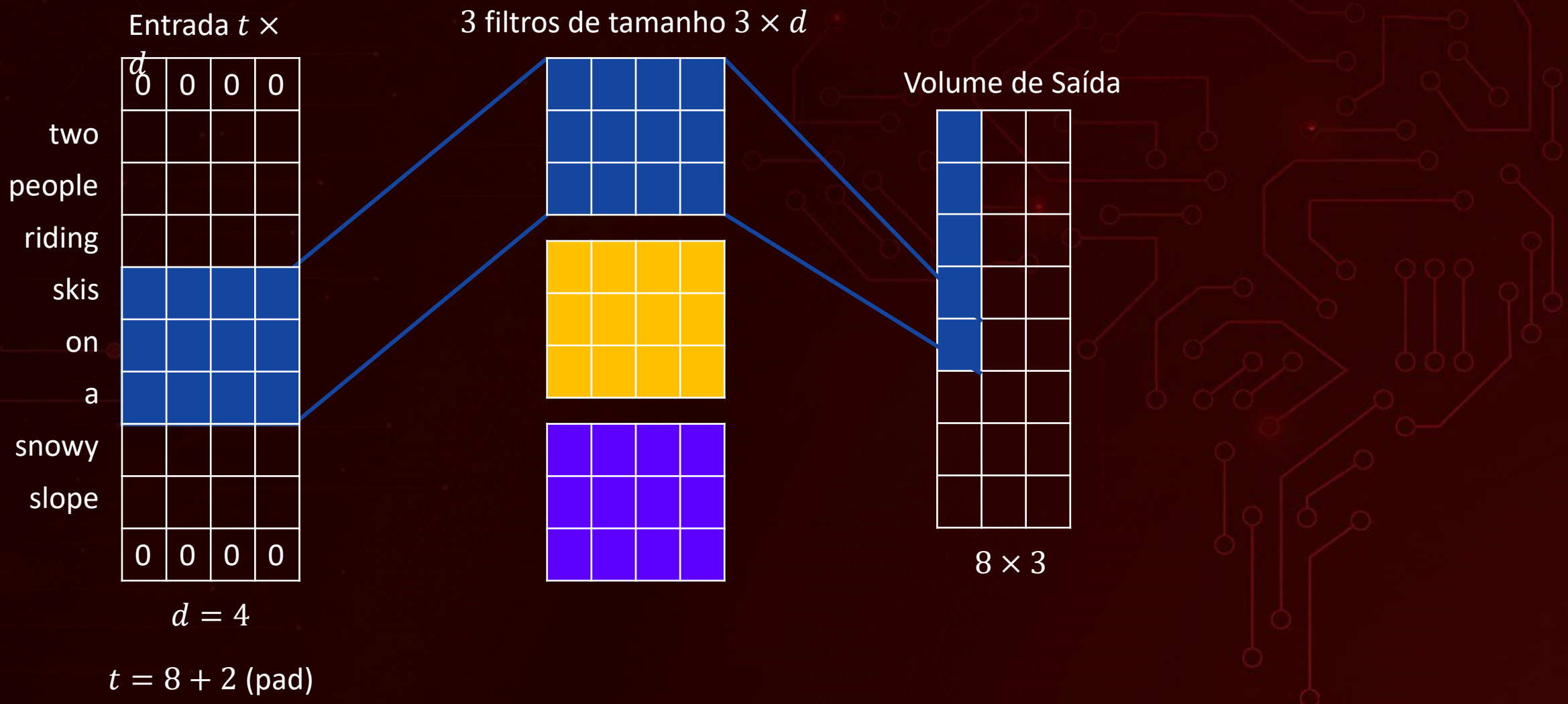
PLN COM CNN


Entrada $t \times d$
0 0 0 0
two
people
riding
skis
on
a
snowy
slope
0 0 0 0
$d = 4$
$t = 8 + 2$ (pad)
3 filtros de tamanho $3 \times d$
Volume de Saída
8×3

PLN COM CNN


Entrada $t \times d$
0 0 0 0
two
people
riding
skis
on
a
snowy
slope
0 0 0 0
$d = 4$
$t = 8 + 2$ (pad)
3 filtros de tamanho $3 \times d$
Volume de Saída
8×3

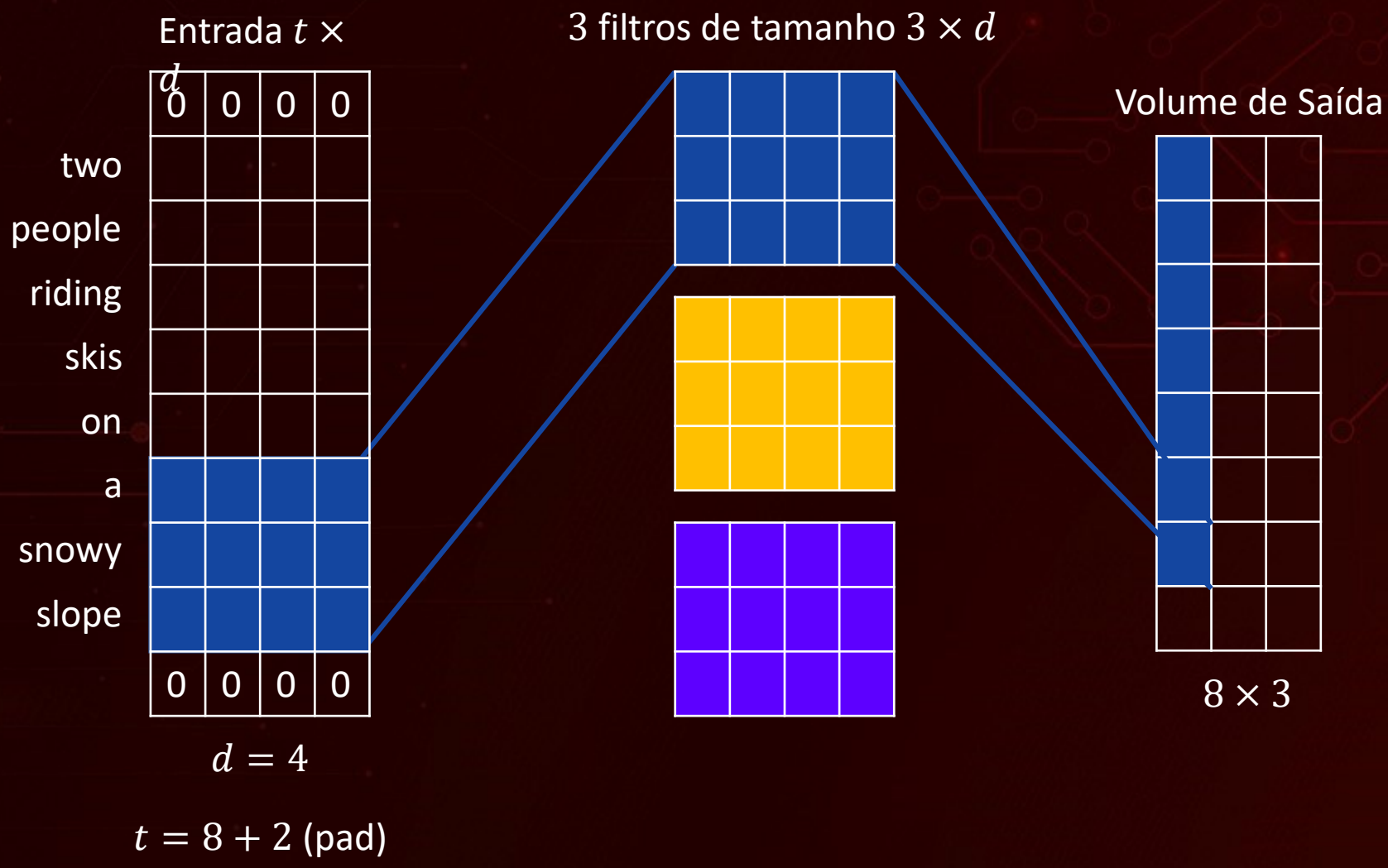
PLN COM CNN


Entrada $t \times d$
0 0 0 0
two
people
riding
skis
on
a
snowy
slope
0 0 0 0
$d = 4$
$t = 8 + 2$ (pad)
3 filtros de tamanho $3 \times d$
Volume de Saída
8×3

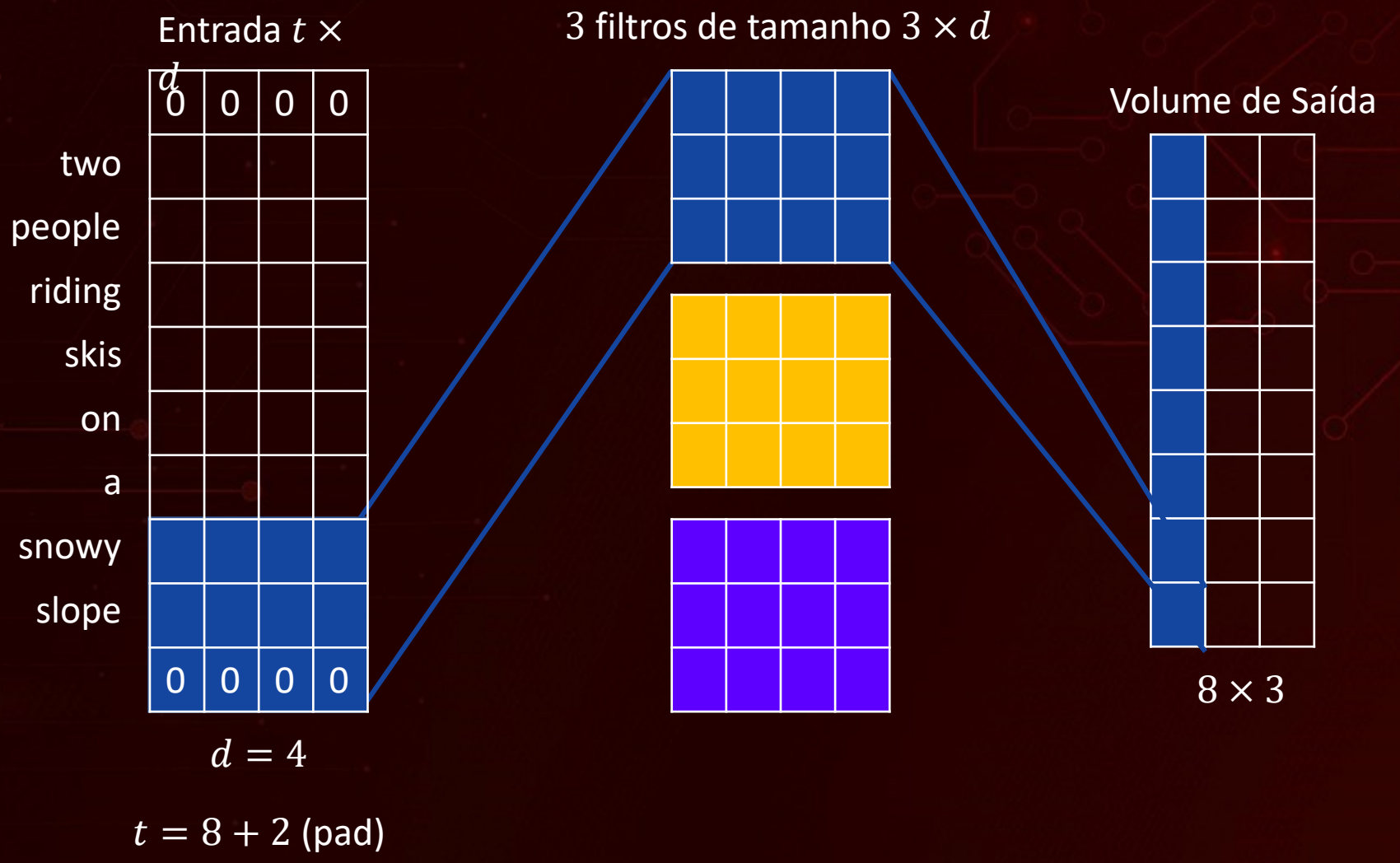
PLN COM CNN


Entrada $t \times d$
0 0 0 0
two
people
riding
skis
on
a
snowy
slope
0 0 0 0
$d = 4$
$t = 8 + 2$ (pad)
3 filtros de tamanho $3 \times d$
Volume de Saída
8×3

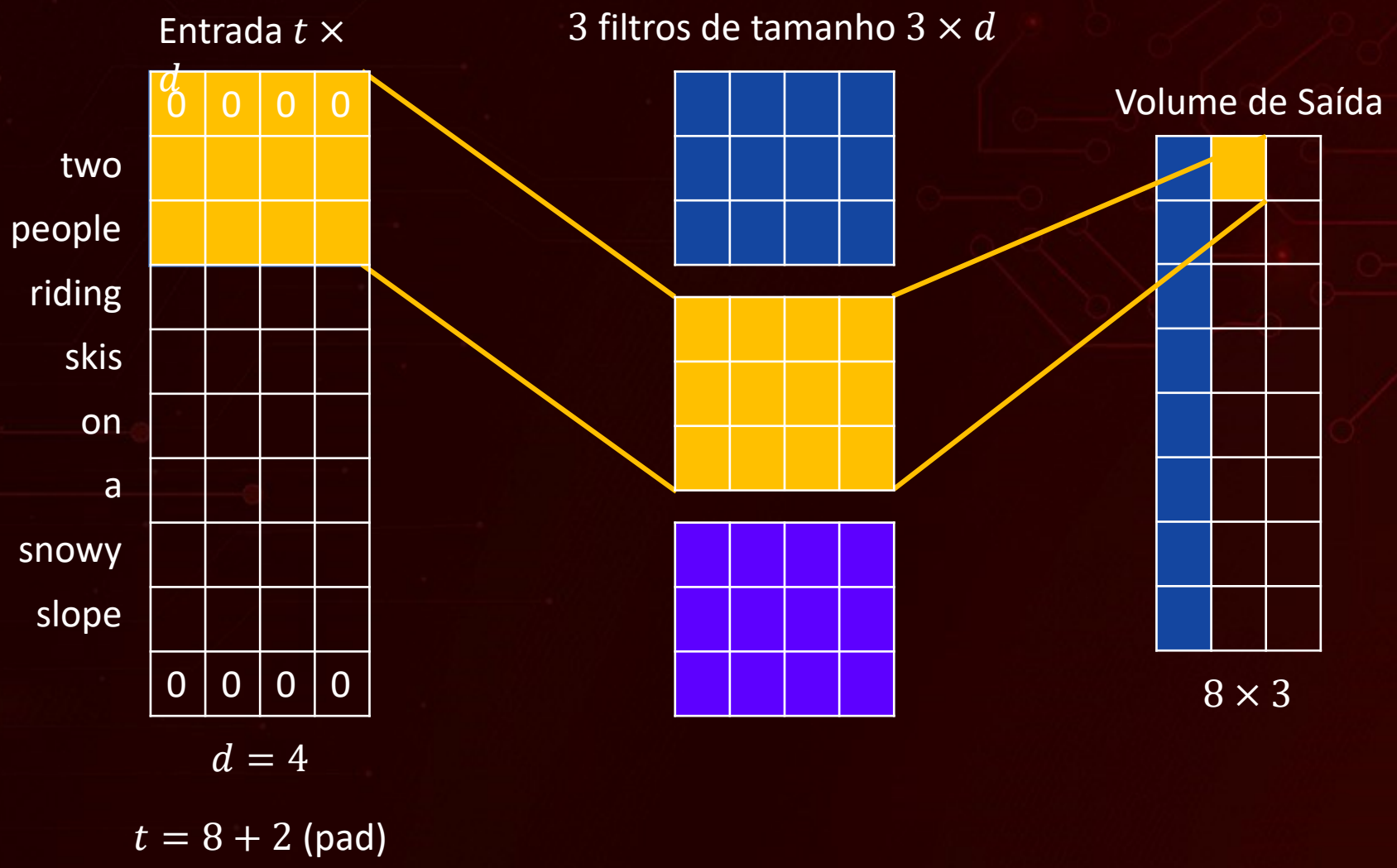
PLN COM CNN


Entrada $t \times d$
0 0 0 0
two
people
riding
skis
on
a
snowy
slope
0 0 0 0
$d = 4$
$t = 8 + 2$ (pad)
3 filtros de tamanho $3 \times d$
Volume de Saída
8×3

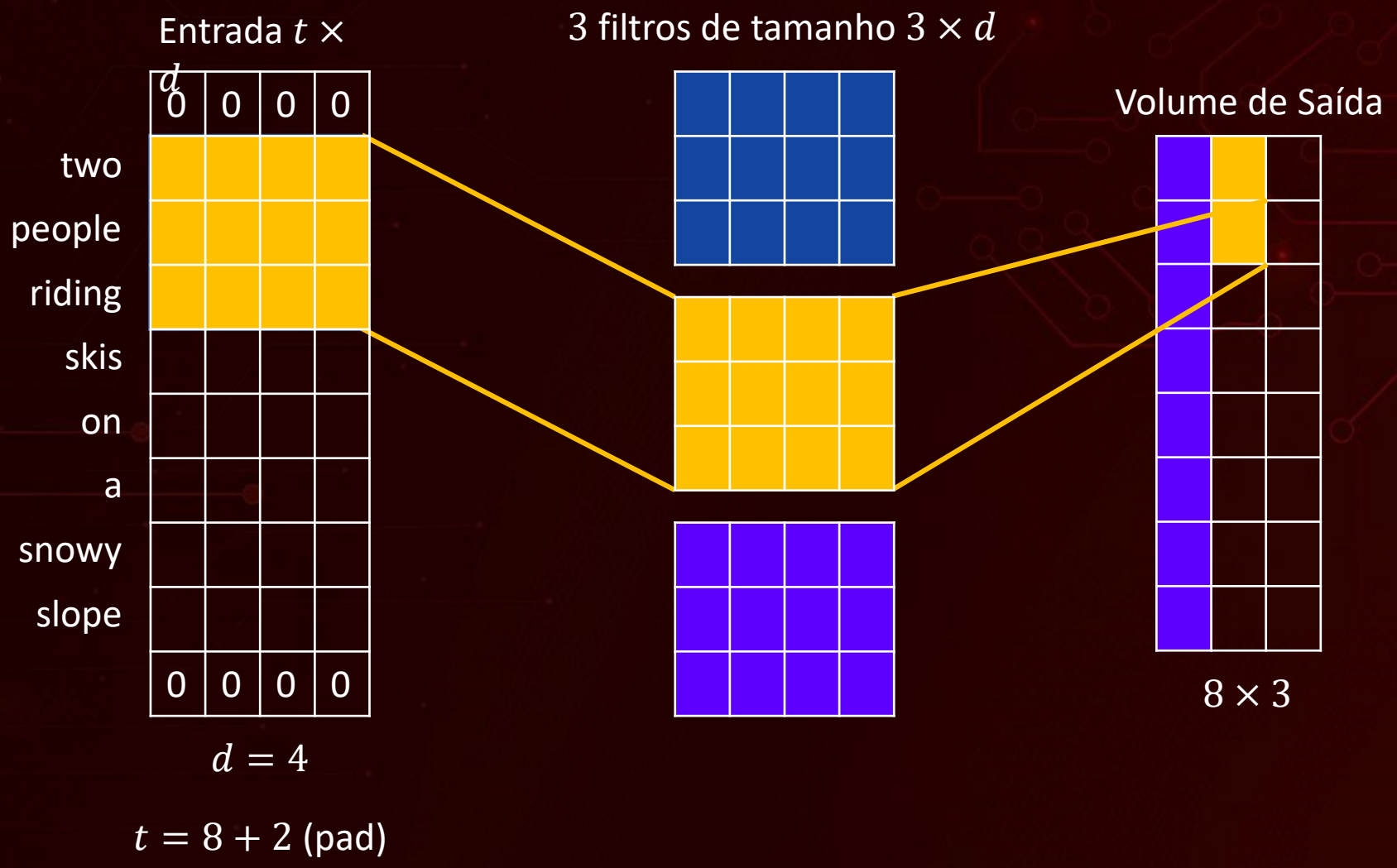
PLN COM CNN


Entrada $t \times d$
0 0 0 0
two
people
riding
skis
on
a
snowy
slope
0 0 0 0
$d = 4$
$t = 8 + 2$ (pad)
3 filtros de tamanho $3 \times d$
Volume de Saída
8×3

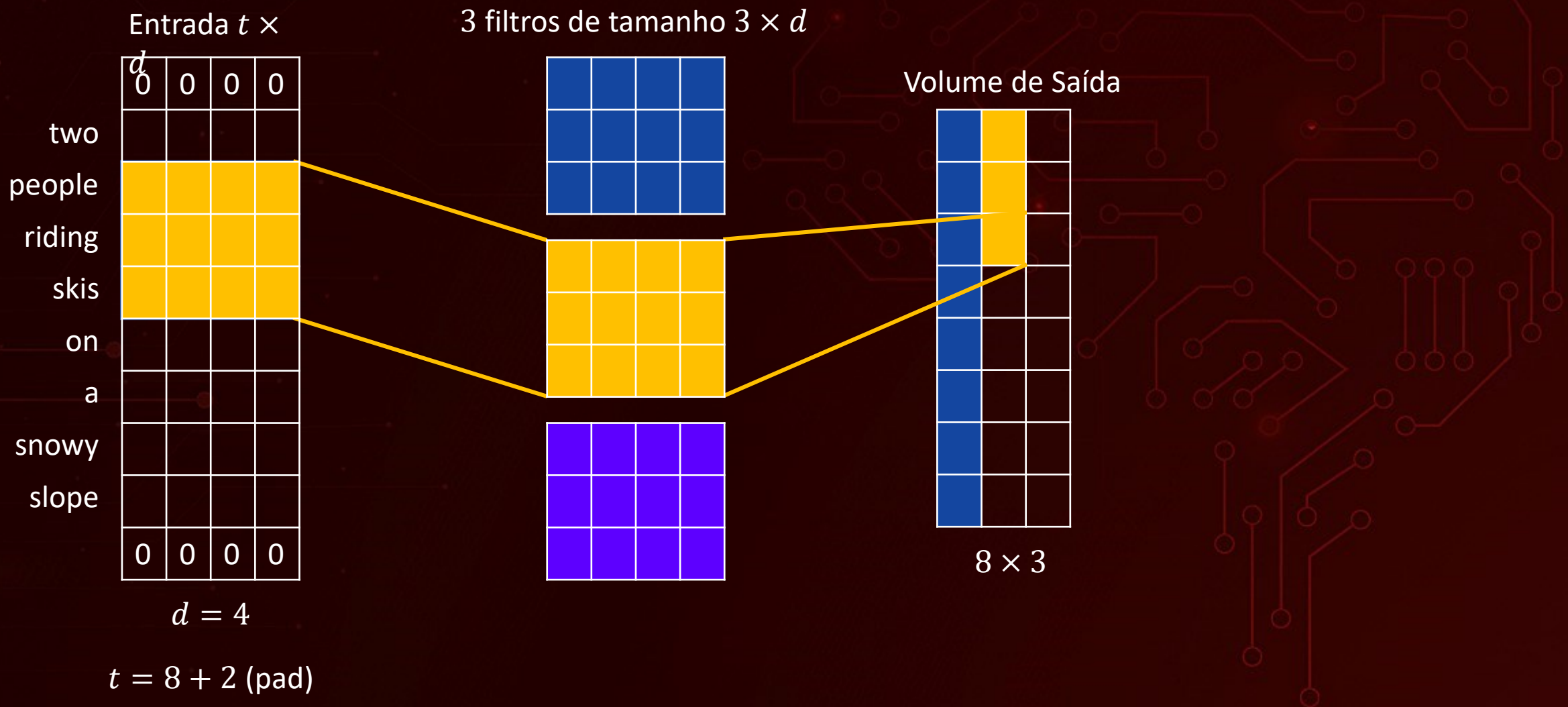
PLN COM CNN


Entrada $t \times d$
3 filtros de tamanho $3 \times d$
Volume de Saída
0 0 0 0
two
people
riding
skis
on
a
snowy
slope
0 0 0 0
$d = 4$
$t = 8 + 2$ (pad)
8×3

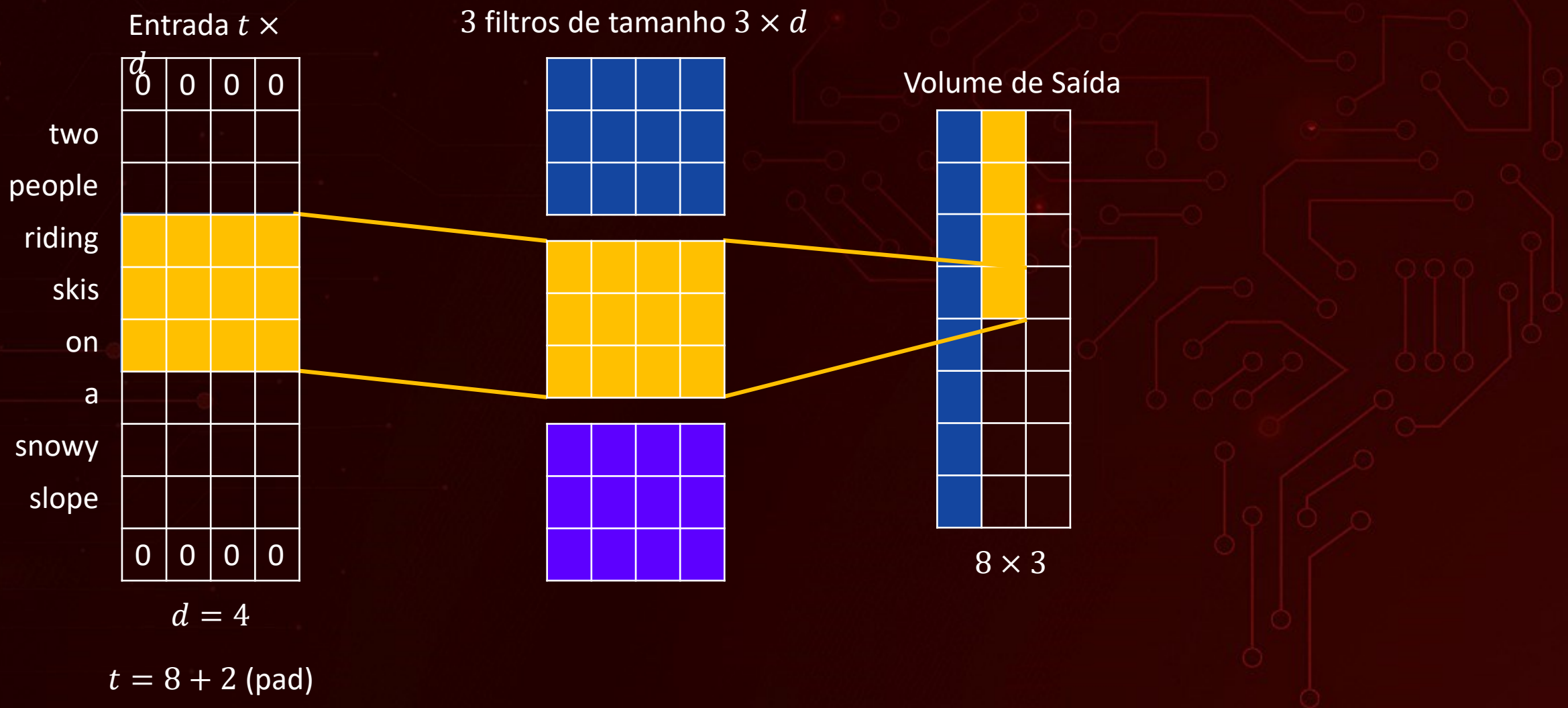
PLN COM CNN


Entrada $t \times d$
3 filtros de tamanho $3 \times d$
Volume de Saída
0 0 0 0
two
people
riding
skis
on
a
snowy
slope
0 0 0 0
$d = 4$
$t = 8 + 2$ (pad)
8×3

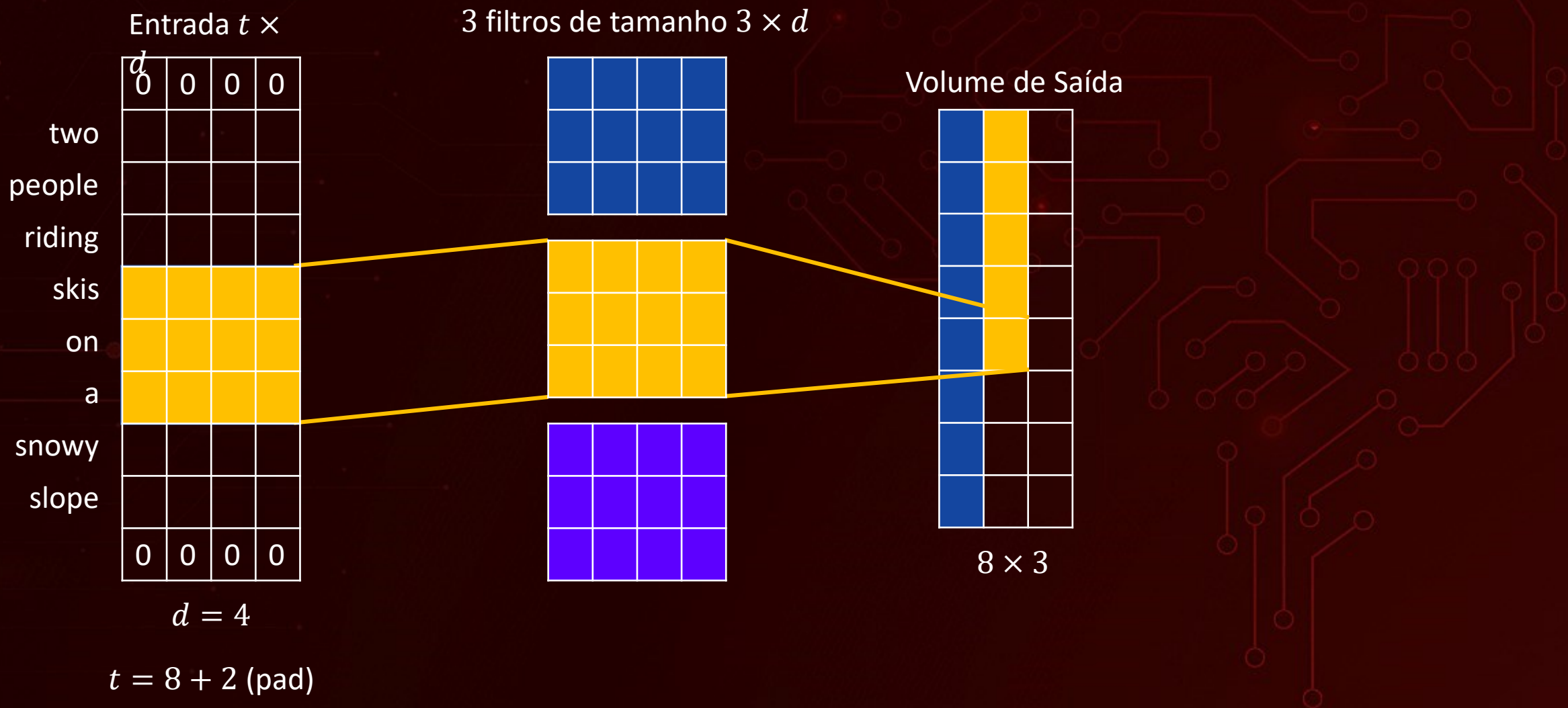
PLN COM CNN


Entrada $t \times d$
0 0 0 0
two
people
riding
skis
on
a
snowy
slope
0 0 0 0
$d = 4$
$t = 8 + 2$ (pad)
3 filtros de tamanho $3 \times d$
Volume de Saída
8×3

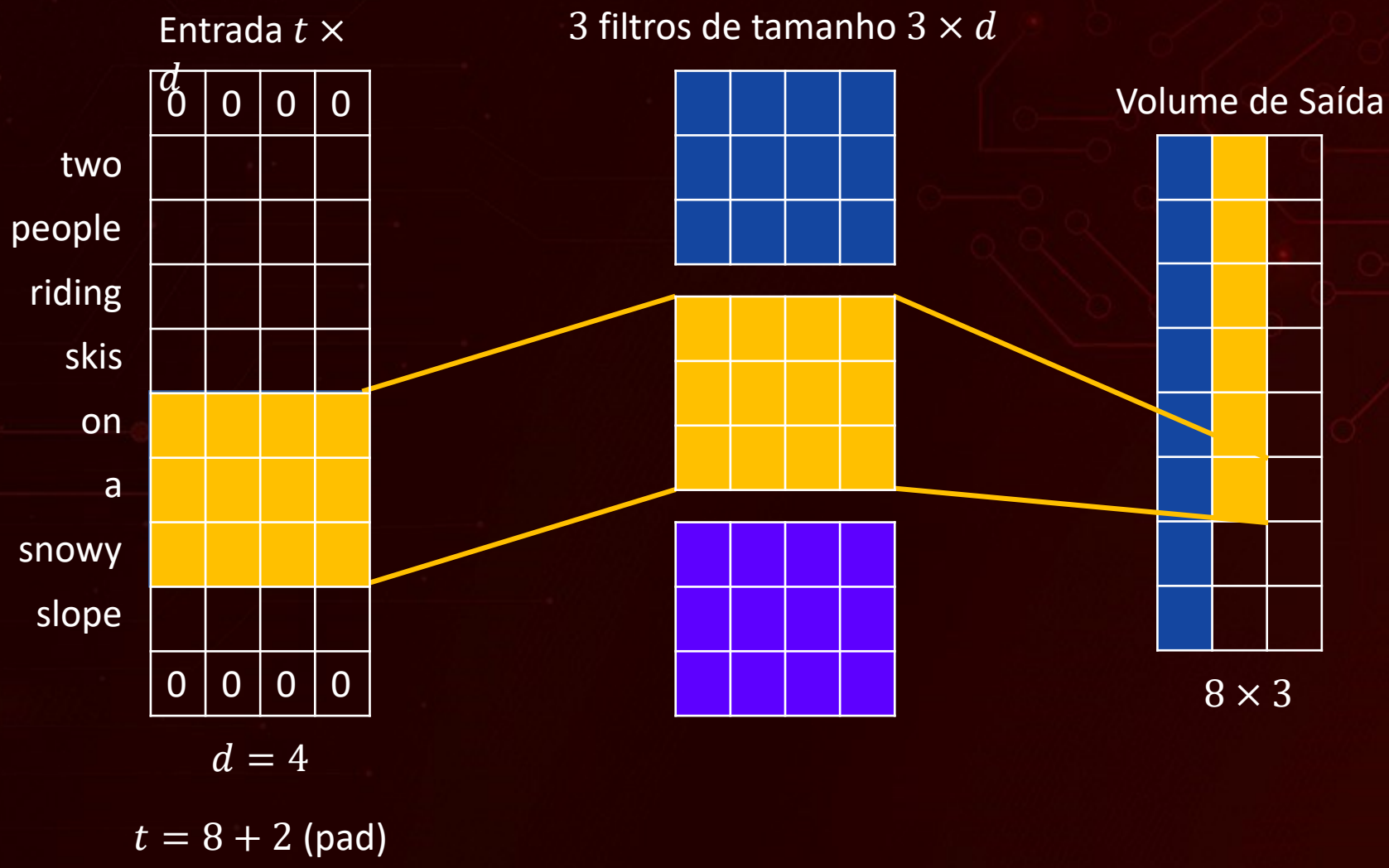
PLN COM CNN


Entrada $t \times d$
0 0 0 0
two
people
riding
skis
on
a
snowy
slope
0 0 0 0
$d = 4$
$t = 8 + 2$ (pad)
3 filtros de tamanho $3 \times d$
Volume de Saída
8×3

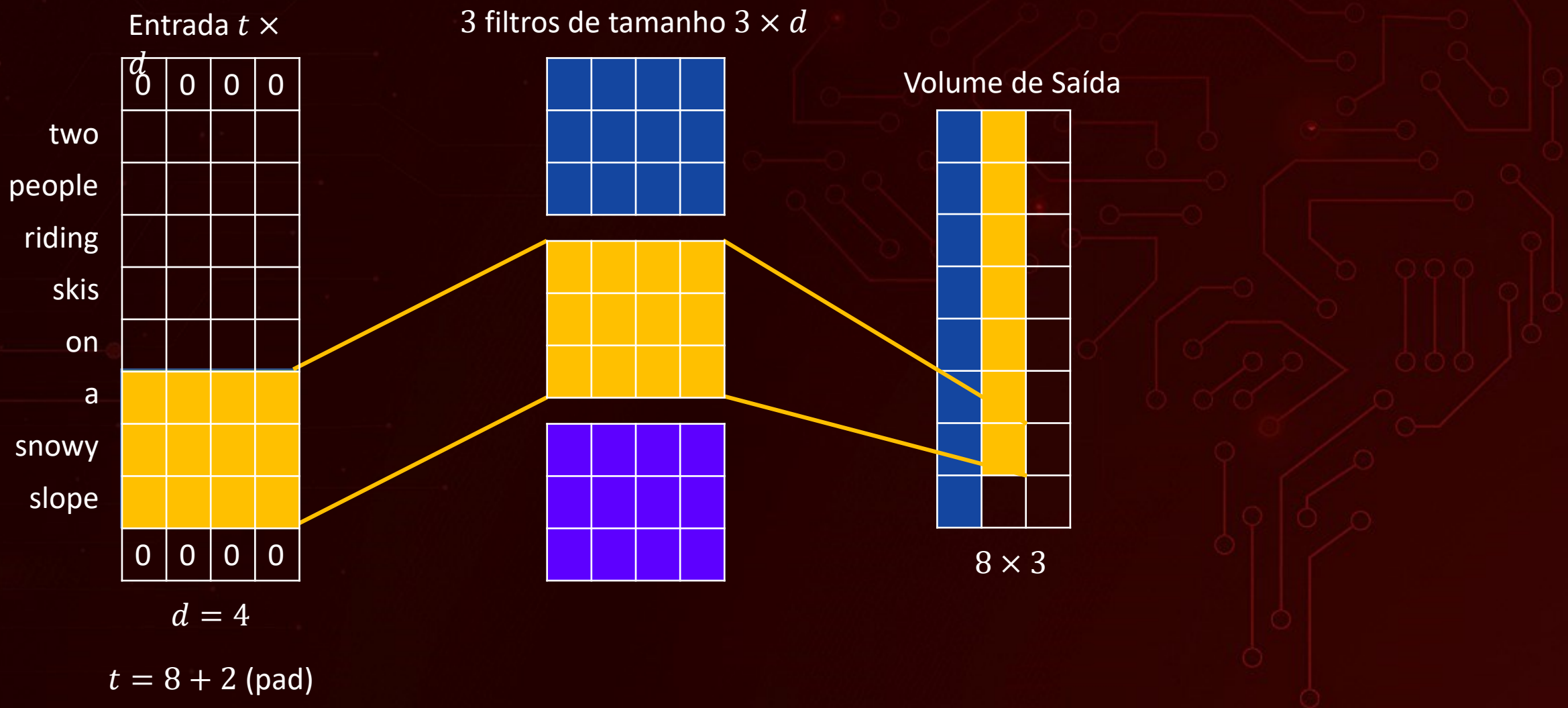
PLN COM CNN


Entrada $t \times d$
0 0 0 0
two
people
riding
skis
on
a
snowy
slope
0 0 0 0
$d = 4$
$t = 8 + 2$ (pad)
3 filtros de tamanho $3 \times d$
Volume de Saída
8×3

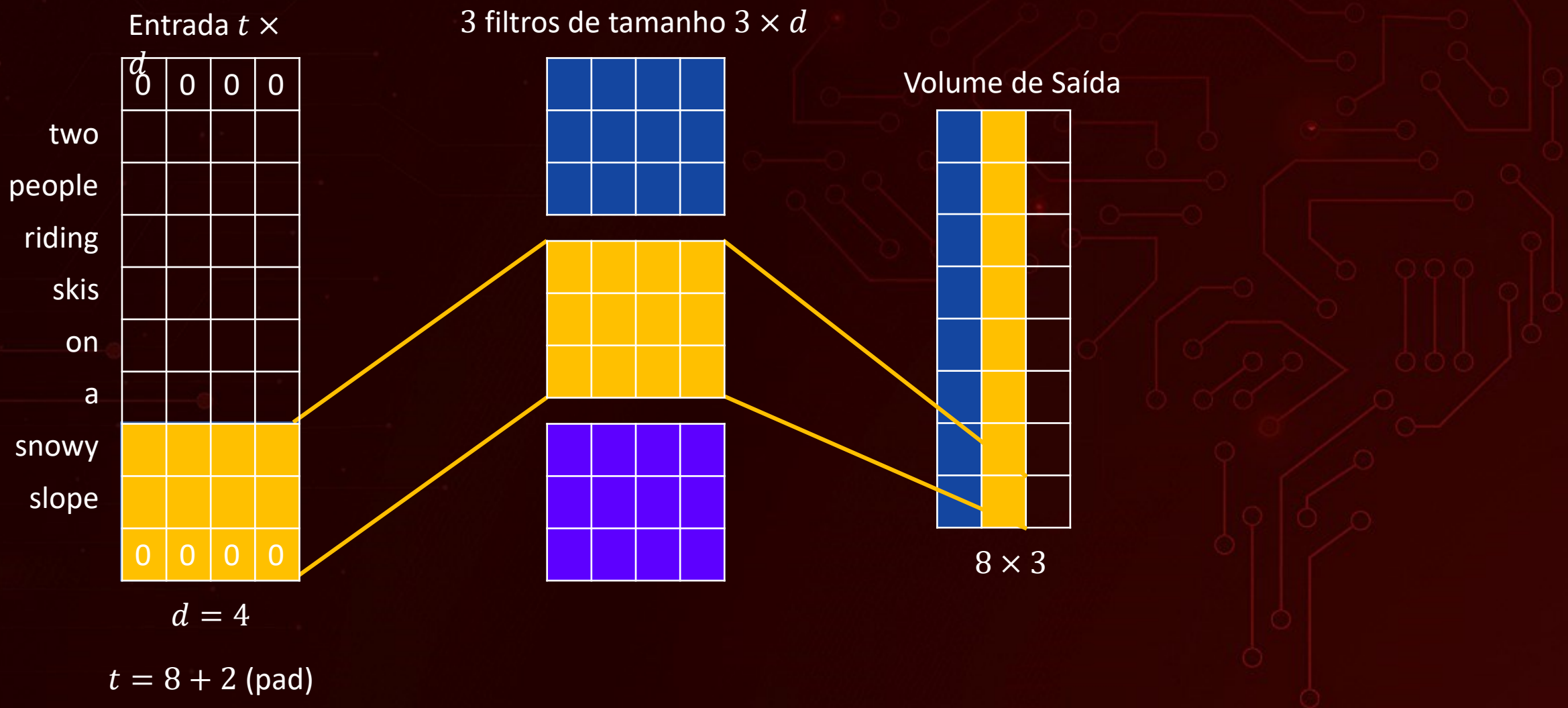
PLN COM CNN


Entrada $t \times d$
0 0 0 0
two
people
riding
skis
on
a
snowy
slope
0 0 0 0
$d = 4$
$t = 8 + 2$ (pad)
3 filtros de tamanho $3 \times d$
Volume de Saída
8×3

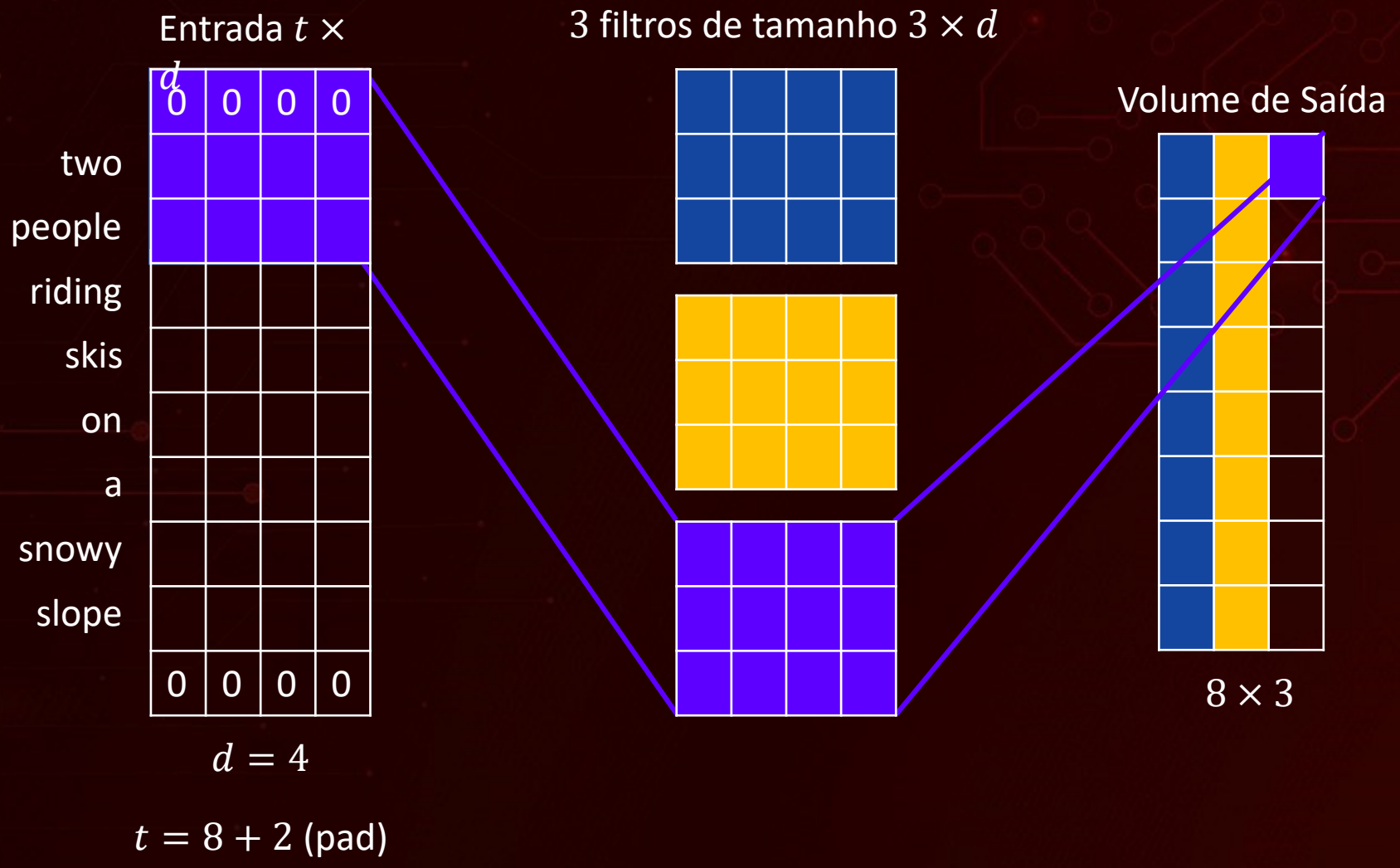
PLN COM CNN


Entrada $t \times d$
0 0 0 0
two
people
riding
skis
on
a
snowy
slope
0 0 0 0
$d = 4$
$t = 8 + 2$ (pad)
3 filtros de tamanho $3 \times d$
Volume de Saída
8×3

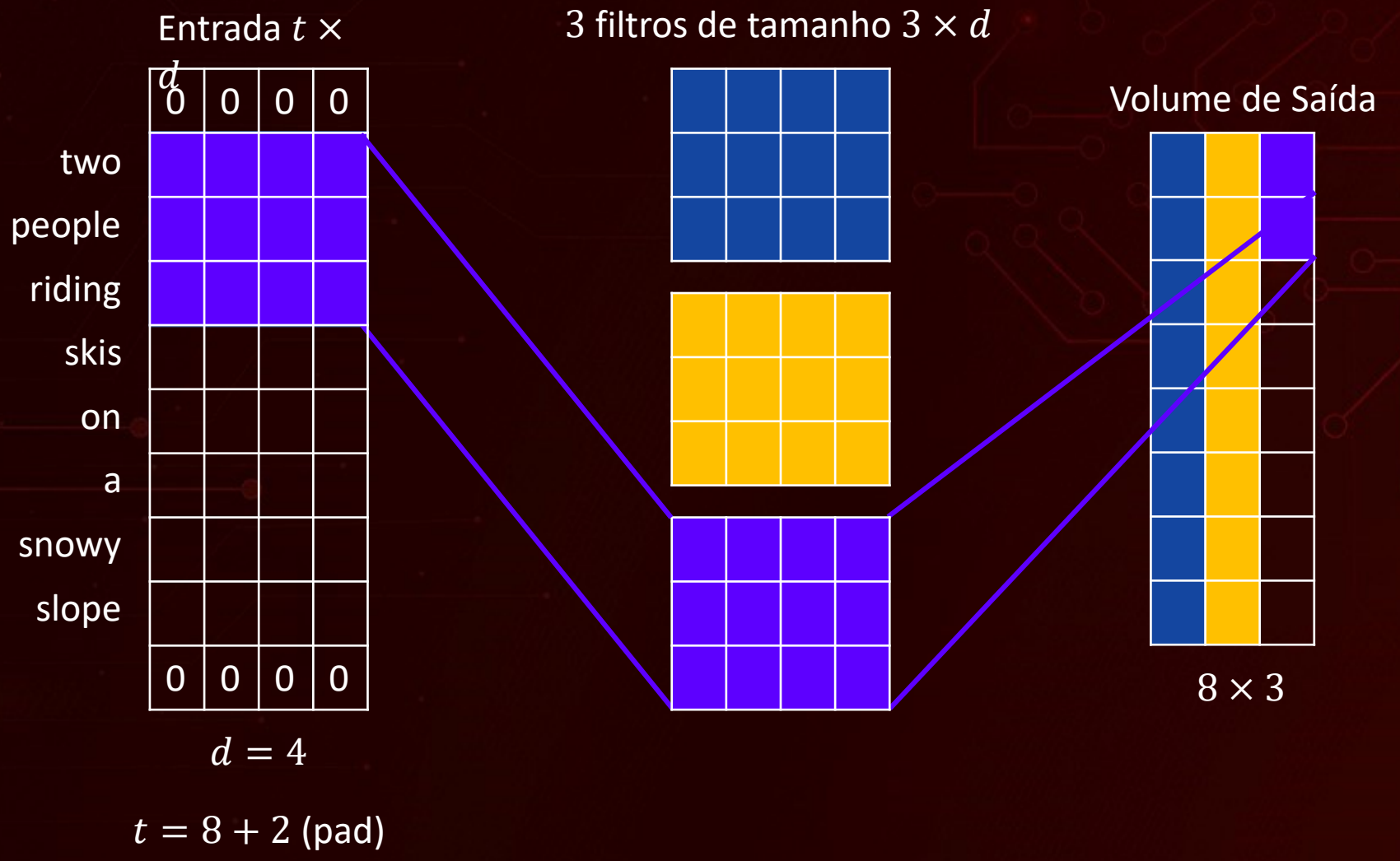
PLN COM CNN


Entrada $t \times d$
0 0 0 0
two
people
riding
skis
on
a
snowy
slope
0 0 0 0
$d = 4$
$t = 8 + 2$ (pad)
3 filtros de tamanho $3 \times d$
Volume de Saída
8×3

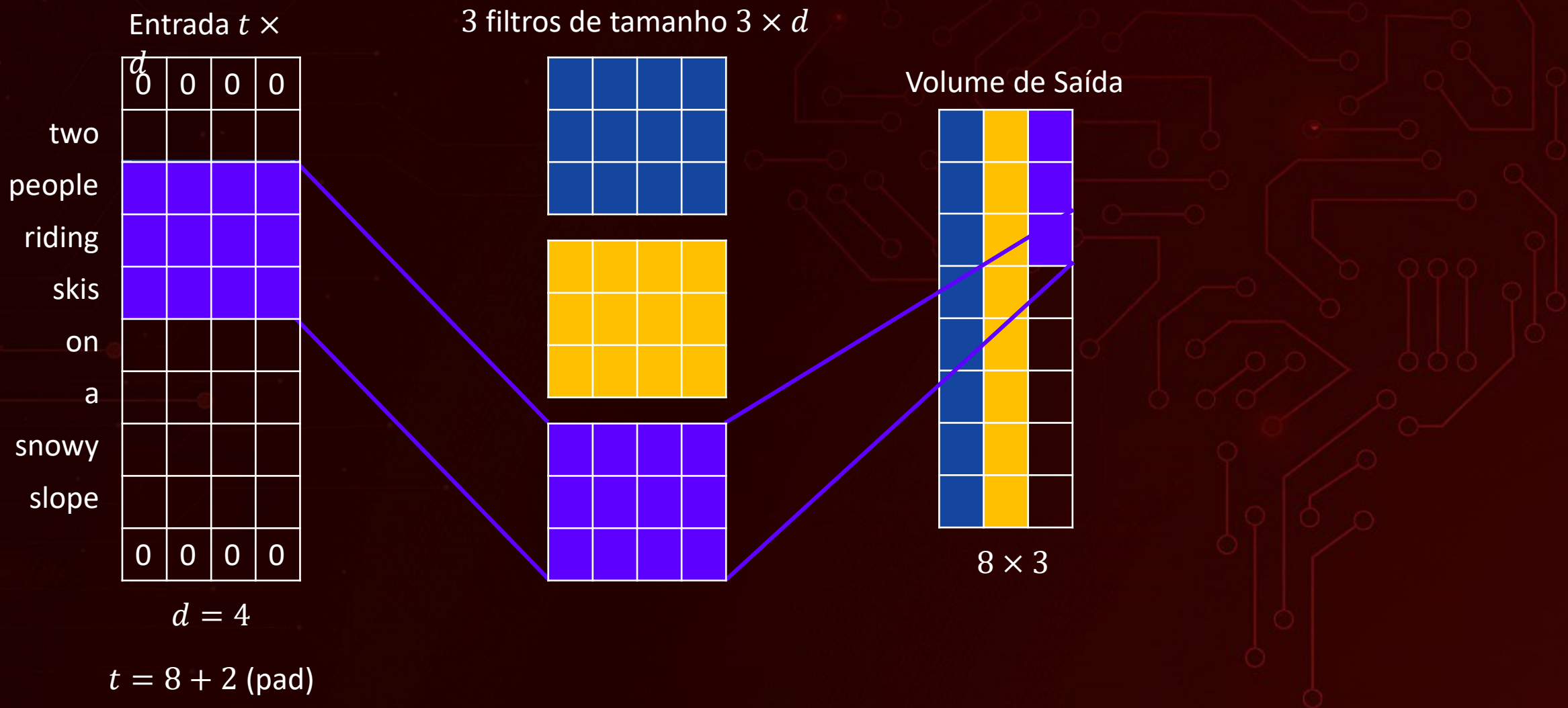
PLN COM CNN


Entrada $t \times d$
0 0 0 0
two
people
riding
skis
on
a
snowy
slope
0 0 0 0
$d = 4$
$t = 8 + 2$ (pad)
3 filtros de tamanho $3 \times d$
Volume de Saída
8×3

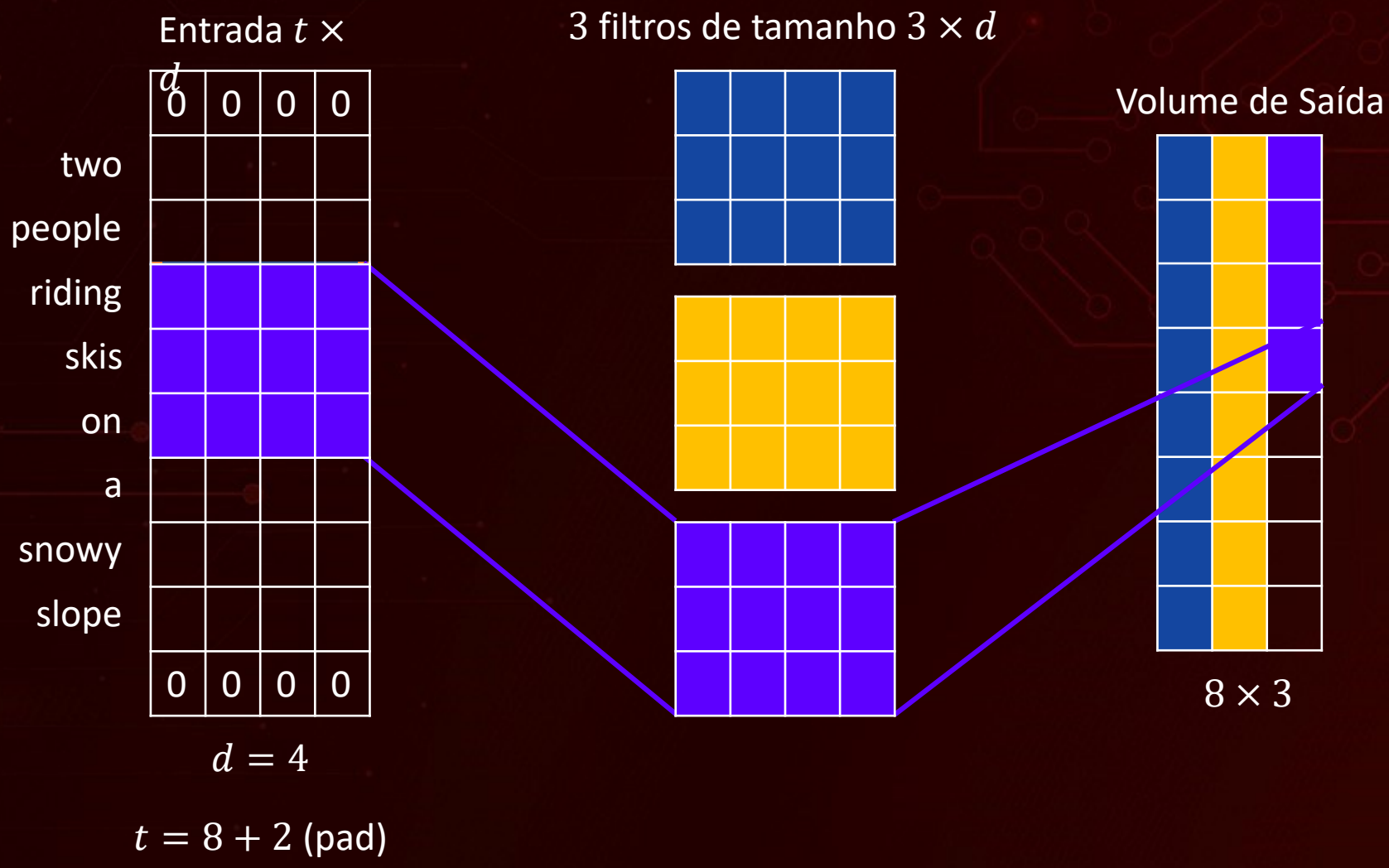
PLN COM CNN


Entrada $t \times d$
0 0 0 0
two
people
riding
skis
on
a
snowy
slope
0 0 0 0
$d = 4$
$t = 8 + 2$ (pad)
3 filtros de tamanho $3 \times d$
Volume de Saída
8×3

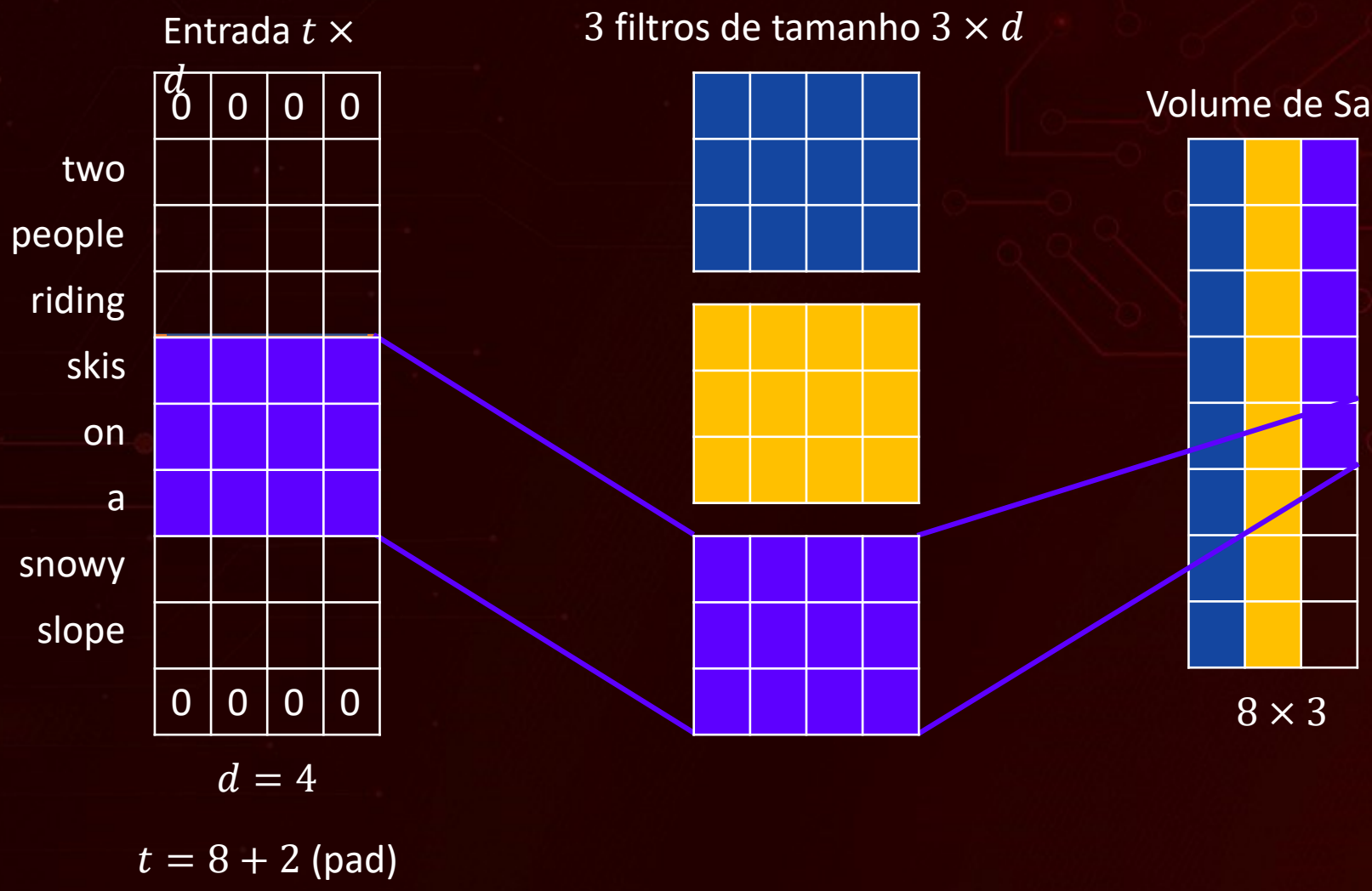
PLN COM CNN


Entrada $t \times d$
0 0 0 0
two
people
riding
skis
on
a
snowy
slope
0 0 0 0
$d = 4$
$t = 8 + 2$ (pad)
3 filtros de tamanho $3 \times d$
Volume de Saída
8×3

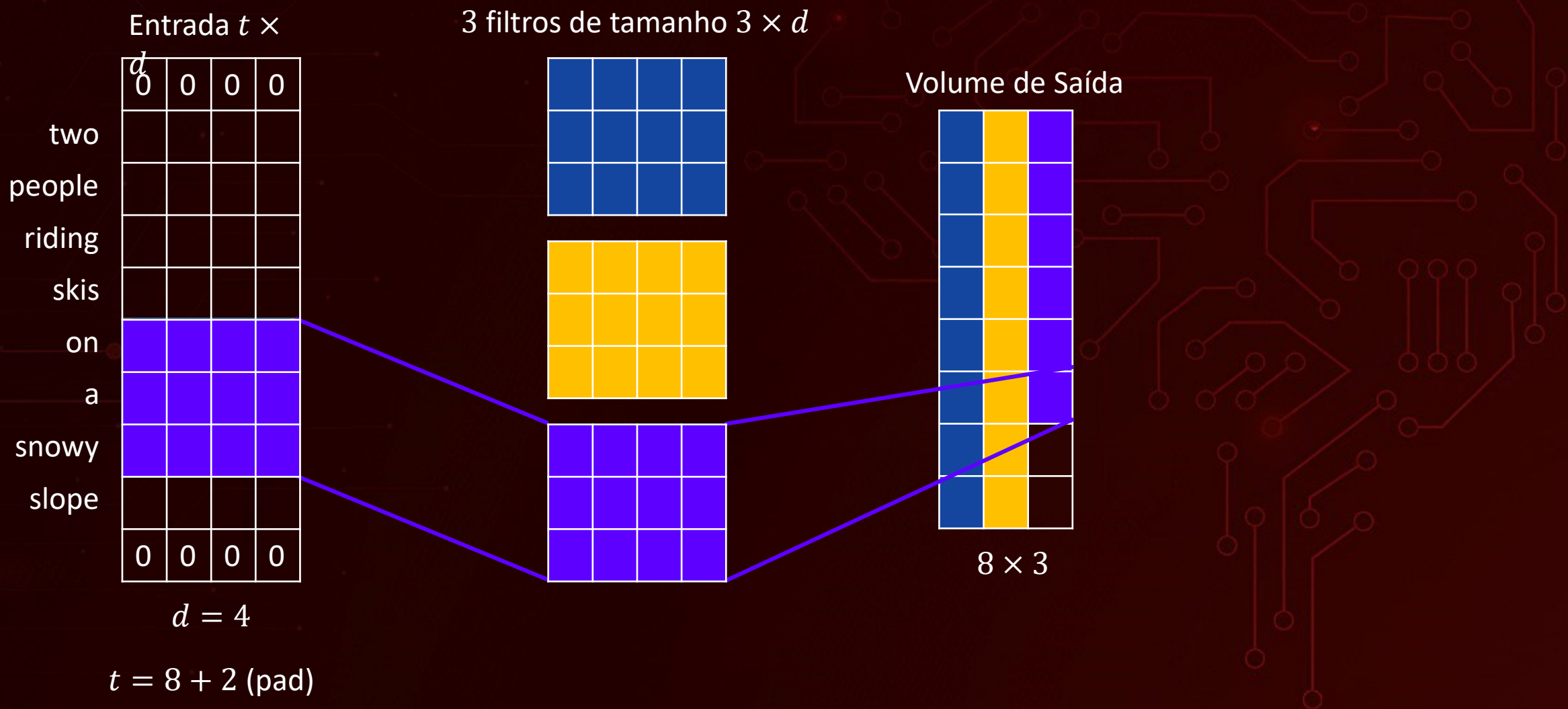
PLN COM CNN


Entrada $t \times d$
0 0 0 0
two
people
riding
skis
on
a
snowy
slope
0 0 0 0
$d = 4$
$t = 8 + 2$ (pad)
3 filtros de tamanho $3 \times d$
Volume de Saída
8×3

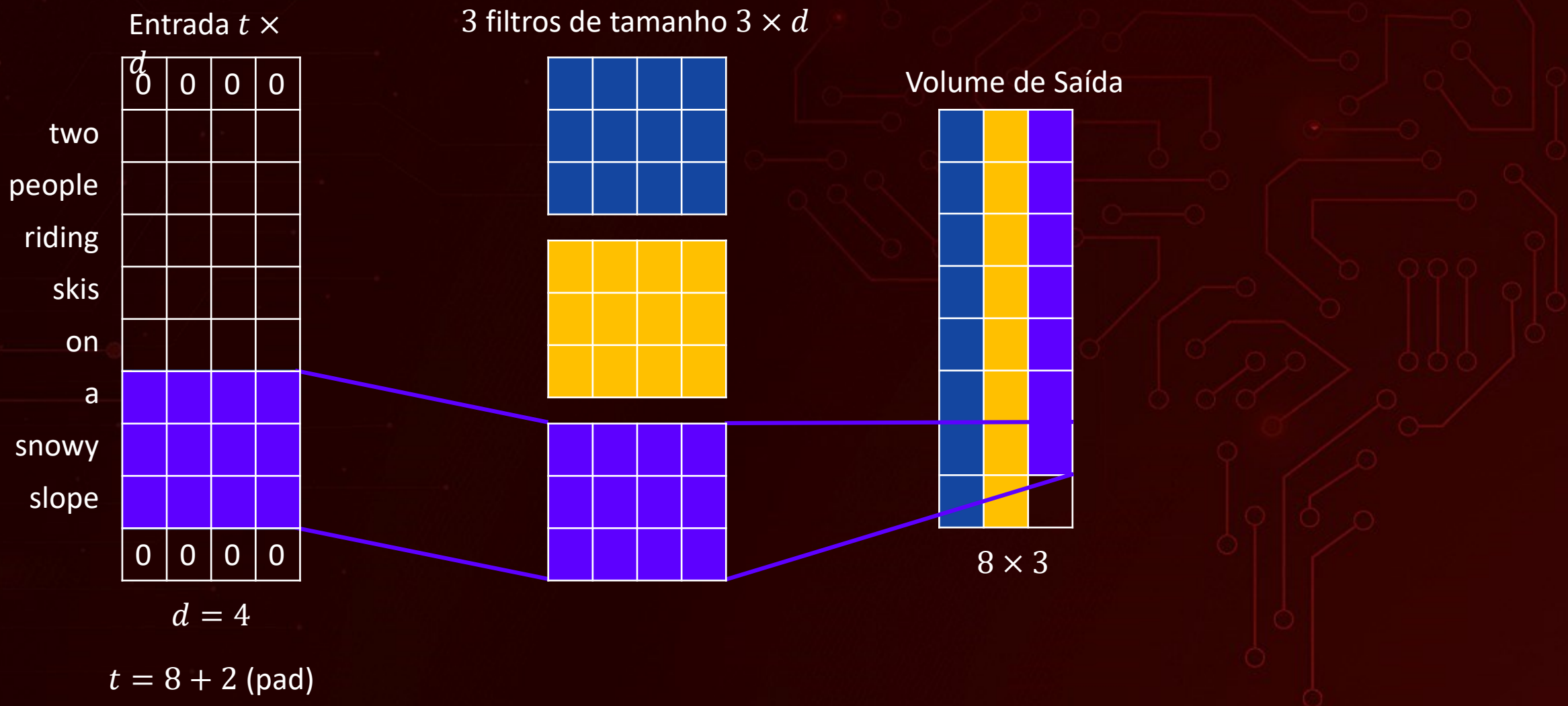
PLN COM CNN


Entrada $t \times d$
0 0 0 0
two
people
riding
skis
on
a
snowy
slope
0 0 0 0
$d = 4$
$t = 8 + 2$ (pad)
3 filtros de tamanho $3 \times d$
Volume de Saída
8×3

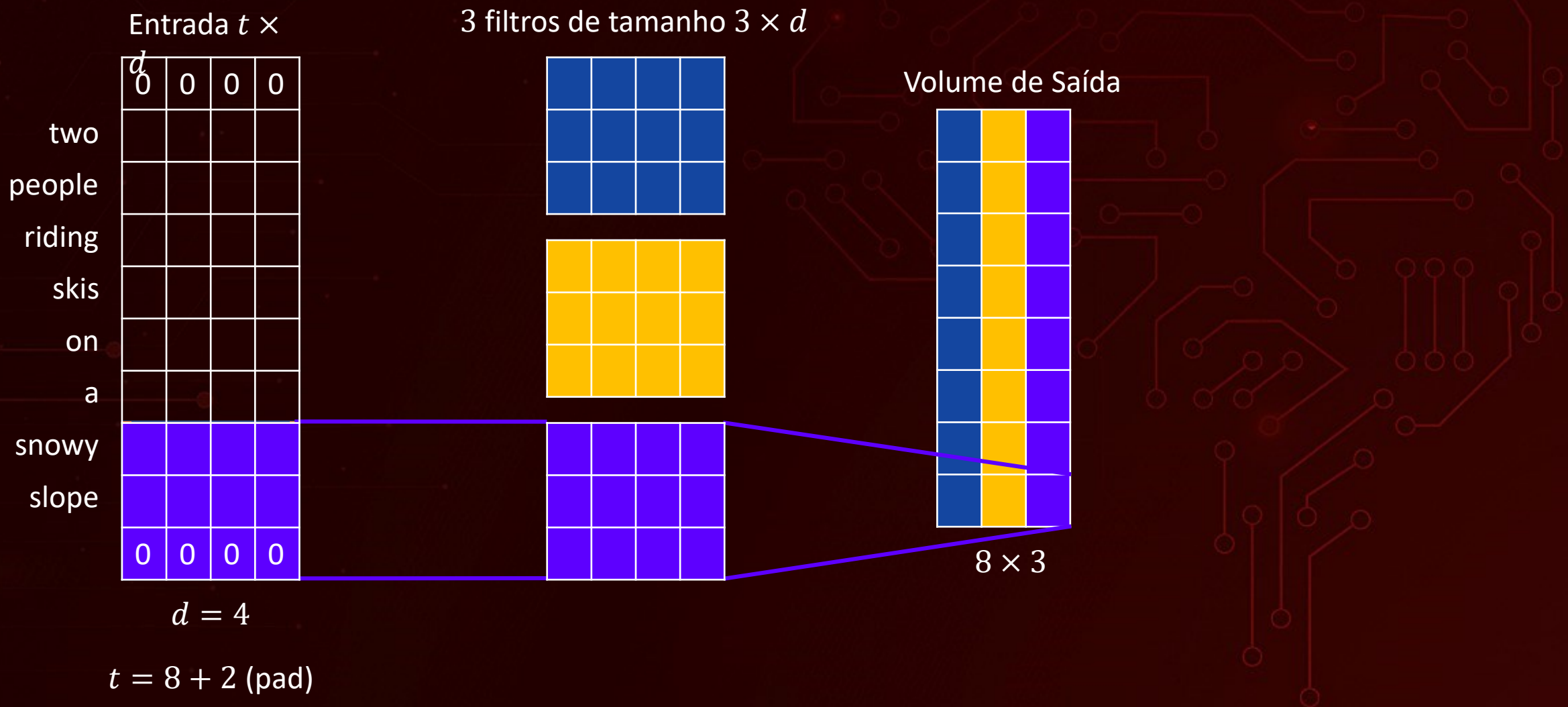
PLN COM CNN


Entrada $t \times d$
0 0 0 0
two
people
riding
skis
on
a
snowy
slope
0 0 0 0
$d = 4$
$t = 8 + 2$ (pad)
3 filtros de tamanho $3 \times d$
Volume de Saída
8×3

PLN COM CNN


Entrada $t \times d$
0 0 0 0
two
people
riding
skis
on
a
snowy
slope
0 0 0 0
$d = 4$
$t = 8 + 2$ (pad)
3 filtros de tamanho $3 \times d$
Volume de Saída
8×3

PLN COM CNN


Entrada $t \times d$
3 filtros de tamanho $3 \times d$
Volume de Saída
0 0 0 0
two
people
riding
skis
on
a
snowy
slope
0 0 0 0
8×3
$d = 4$
$t = 8 + 2$ (pad)

PLN COM CNN

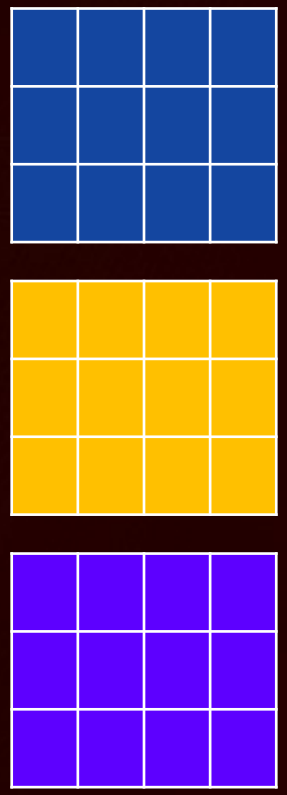
Entrada $t \times d$

	0 0 0 0
two	
people	
riding	
skis	
on	
a	
snowy	
slope	
	0 0 0 0

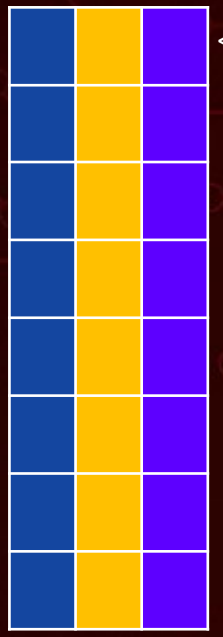
$d = 4$

$t = 8 + 2$ (pad)

3 filtros de tamanho $3 \times d$



Volume de Saída

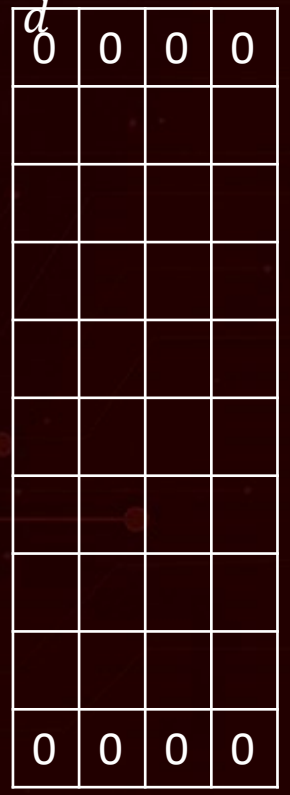


<pad>, two, people

8×3

PLN COM CNN

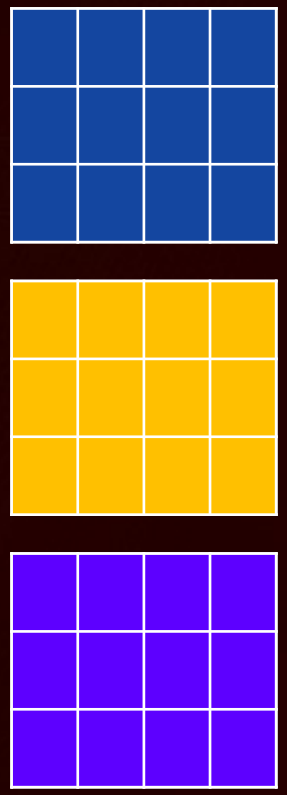
Entrada $t \times d$



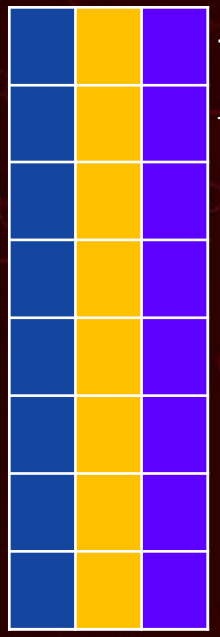
$d = 4$

$t = 8 + 2$ (pad)

3 filtros de tamanho $3 \times d$



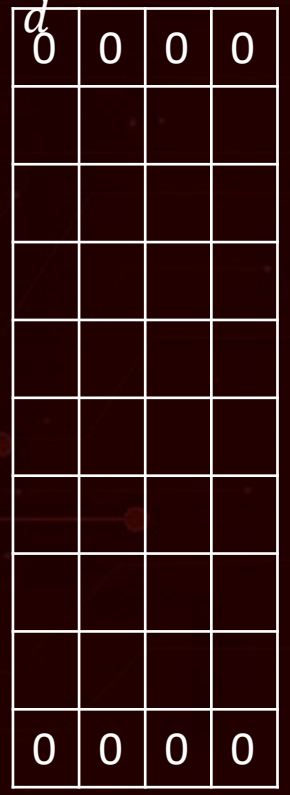
Volume de Saída



8×3

PLN COM CNN

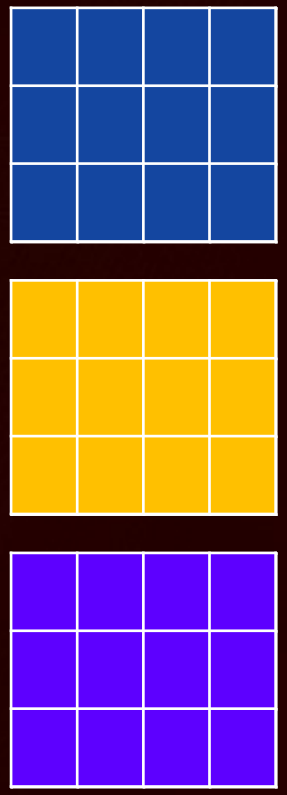
Entrada $t \times d$



$d = 4$

$t = 8 + 2$ (pad)

3 filtros de tamanho $3 \times d$



Volume de Saída

8×3

PLN COM CNN

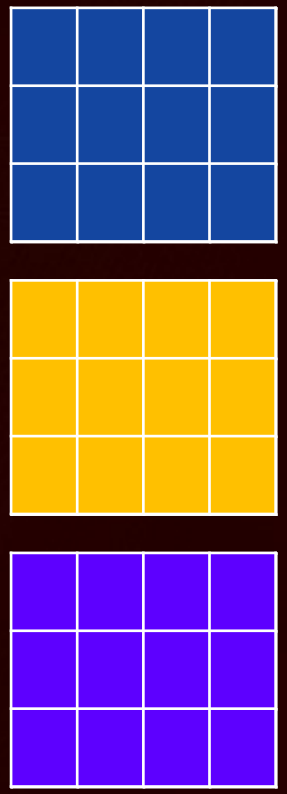
Entrada $t \times d$

	0	0	0	0
two				
people				
riding				
skis				
on				
a				
snowy				
slope				
	0	0	0	0

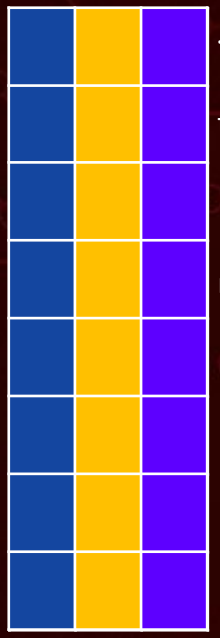
$d = 4$

$t = 8 + 2$ (pad)

3 filtros de tamanho $3 \times d$



Volume de Saída



<pad>, two, people

two, people, riding

people, riding , skis

riding, skis, on

skis, on, a

on, a, snowy

a, snowy, slope

snowy, slope, <pad>

8×3

PLN COM CNN

- **1 camada de Conv1D com filtro de tamanho 3 aprende relações entre 3 palavras vizinhas (trigramas)**
- **O que fazer para aprender relações entre mais palavras?**

PLN COM CNN

- **Aumentar o tamanho do filtro (features mais globais), porém, com aumento do número de parâmetros….**
- **Aumentar o campo receptor via mais camadas (mais flexibilidade com menos parâmetros)**

CNN: CLASSIFICAÇÃO DE TEXTOS

1. Word Embeddings + CNN
2. Usar uma CNN com uma única camada
3. Hiperparâmetros das CNNs
4. CNNs em nível de caractere
5. CNNs profundas

https://machinelearningmastery.com/best-practices-document-classification-deep-learning/

O QUE VIMOS?

- Introdução
- Camadas de convolução e *pooling*
- CNN para classificação de textos

PRÓXIMA VIDEOAULA

- **Prática: Classificação de Textos**