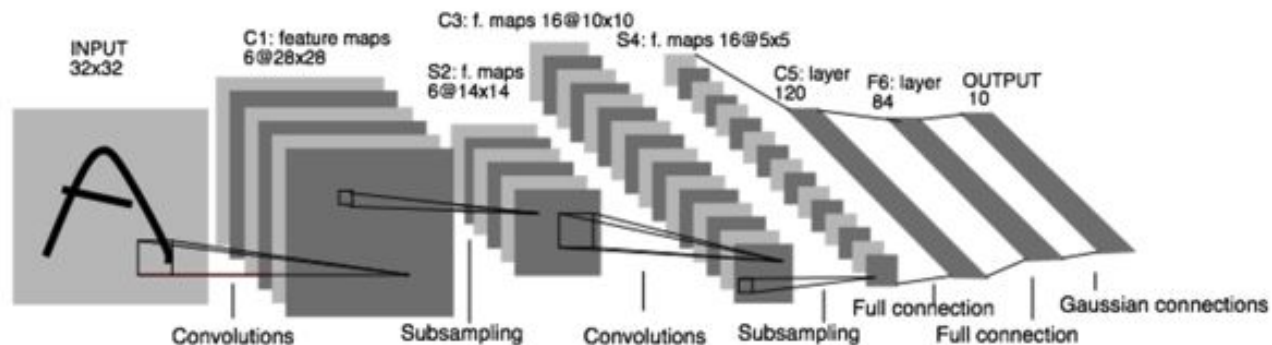

Deep Learning



— Redes Neurais Convolucionais —
(CNN)

Roberto Matheus Pinheiro Pereira

LeNet - Yan LeCun em 1998 (MNIST)



LeNet: a layered model composed of convolution and subsampling operations followed by a holistic representation and ultimately a classifier for handwritten digits. [LeNet]

não é um conceito tão recente

Redes Neurais Convolucionais

“Transforma um volume de entrada
em um volume de saída”

- Neurônios tridimensional
- Cada camada tem uma funcionalidade

- Camada de Convolução
- Camada de Pooling
- MLP
- Camadas de ativação
- Aprendizado supervisionado

Camada de Convolução

- Principal camada
- Dá nome ao tipo de rede
- Conceito matemático
 - Operador linear
 - Realiza uma operação sobre duas funções de entrada

- No nosso caso, as funções de entrada são entendidas como matrizes imagens

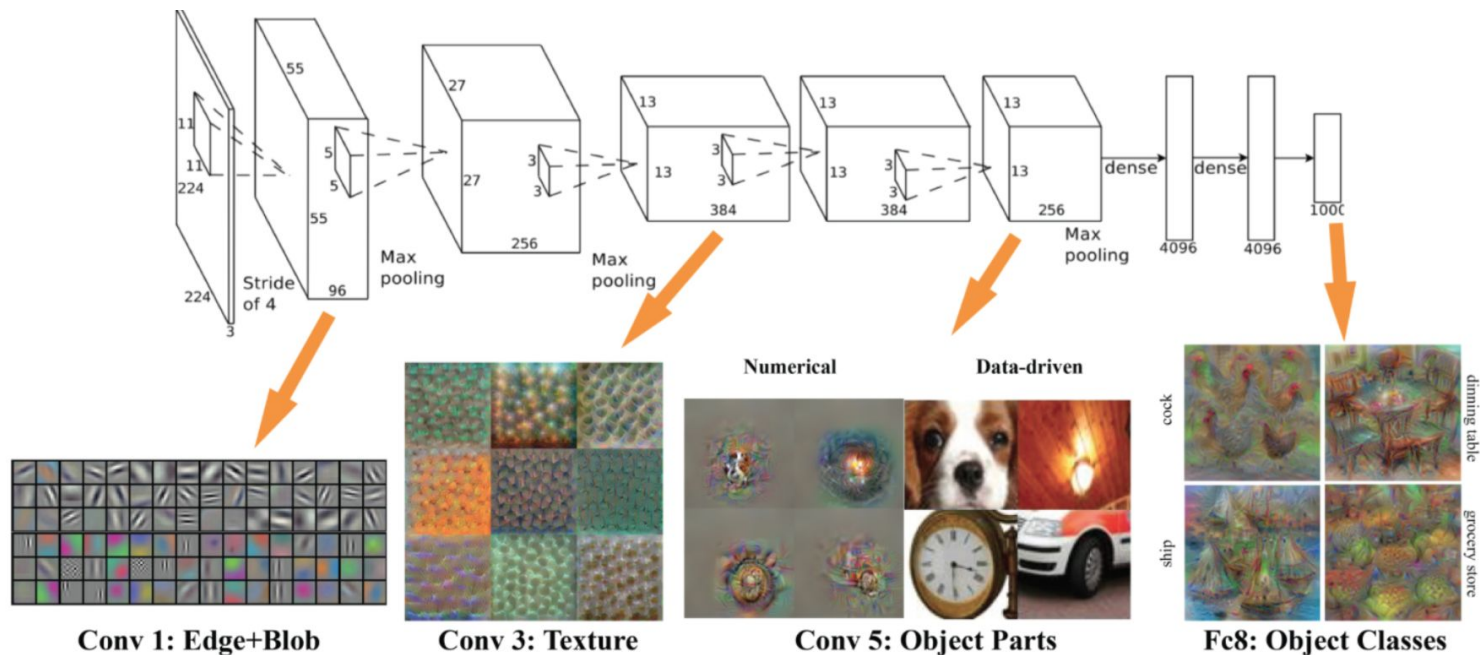
Camada de Convolução - Objetivo

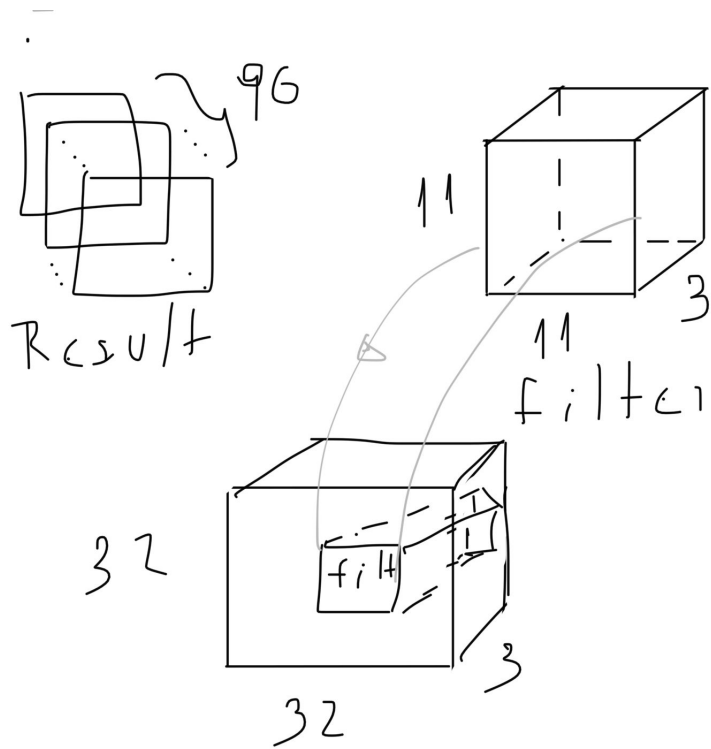
- Aprender filtros 2D sobre uma entrada.
- Cada filtro representa alguma característica importante.



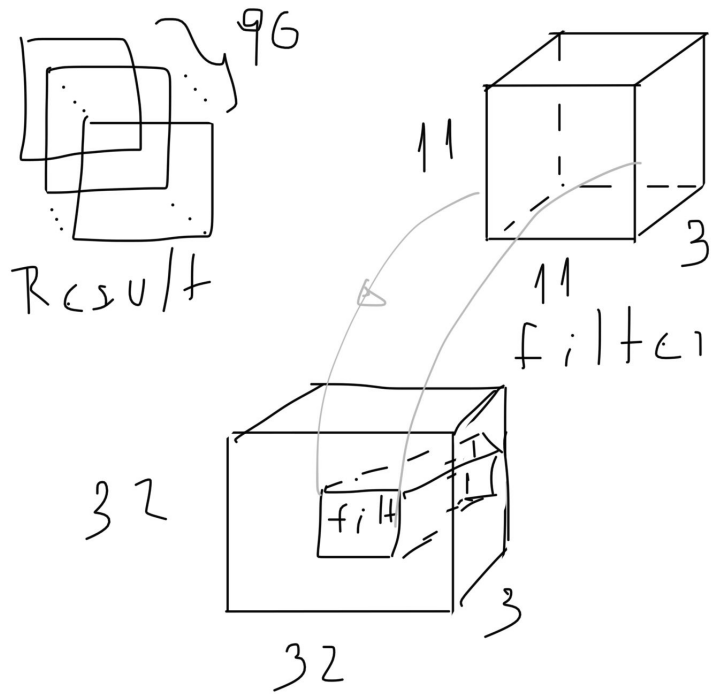
Filtros da primeira camada da rede neural AlexNet utilizando a base CIFAR-10.

Camada de Convolução - Objetivo





- **MLP**
- Cada camada possui 55x55 neurônios
- cada neurônio está ligado a 11x1x3 pesos +1 bias
- Temos 96 lasers ou filtros
- $55 \cdot 55 \cdot 96 \cdot (11 \cdot 11 \cdot 3 + 1) = 105'705'600$ pesos para serem aprendidos
- 10! > 3,1 milhões

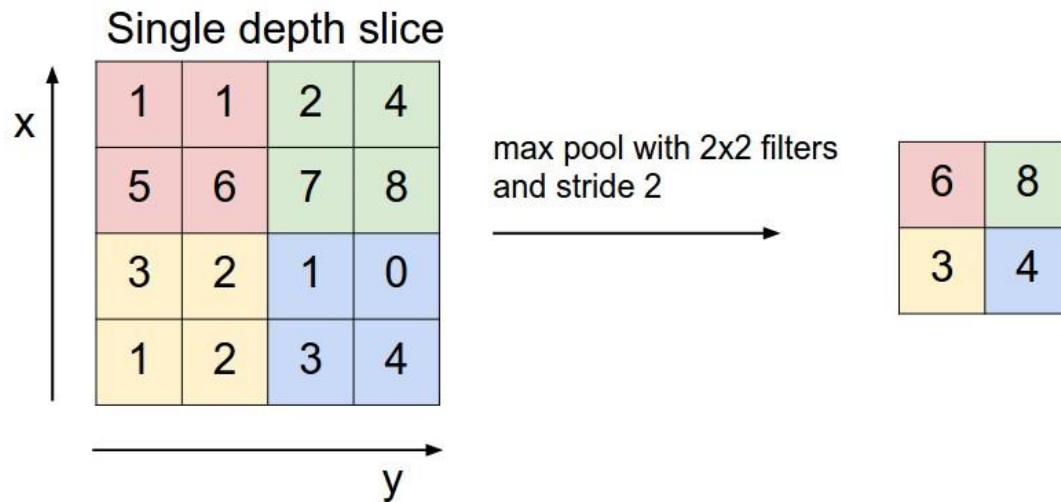


- **Convolucao**
- Cada Layer representa uma mesma característica, então faz sentido termos que aprender somente um peso é um bias.
- Não precisamos de um neurônio por que a camada por si só já representa uma característica.
- Aprender somente um peso reflete em todo mundo
- Previne overfitting
- 96*(11*11*3) 34 mil !!

Camada de Pooling

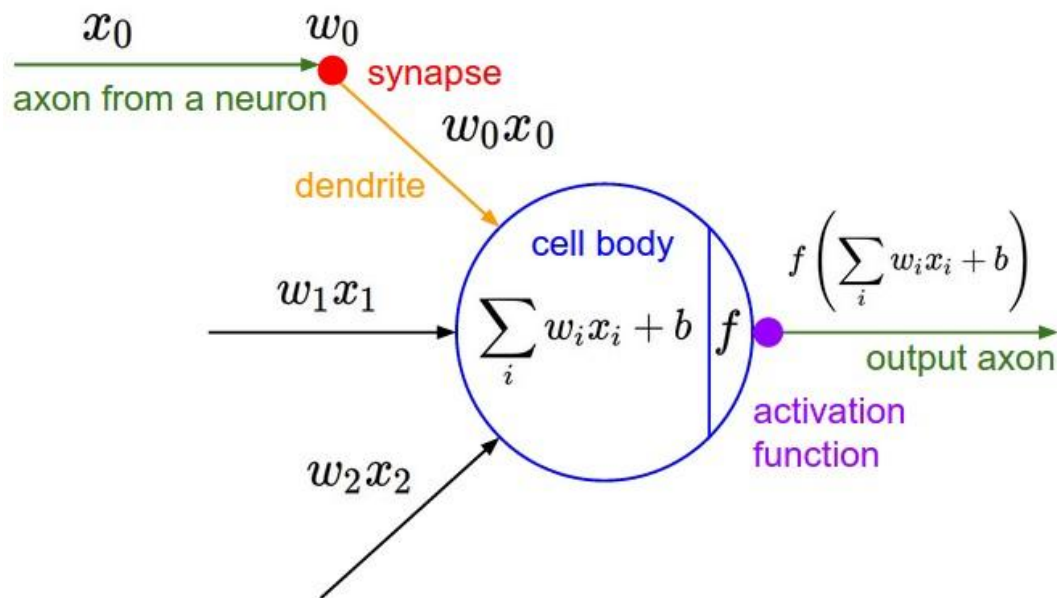
- Tamanho do kernel
- Tipo: Máximo ou Media

$$\text{Output} = (W-F)/S+1$$



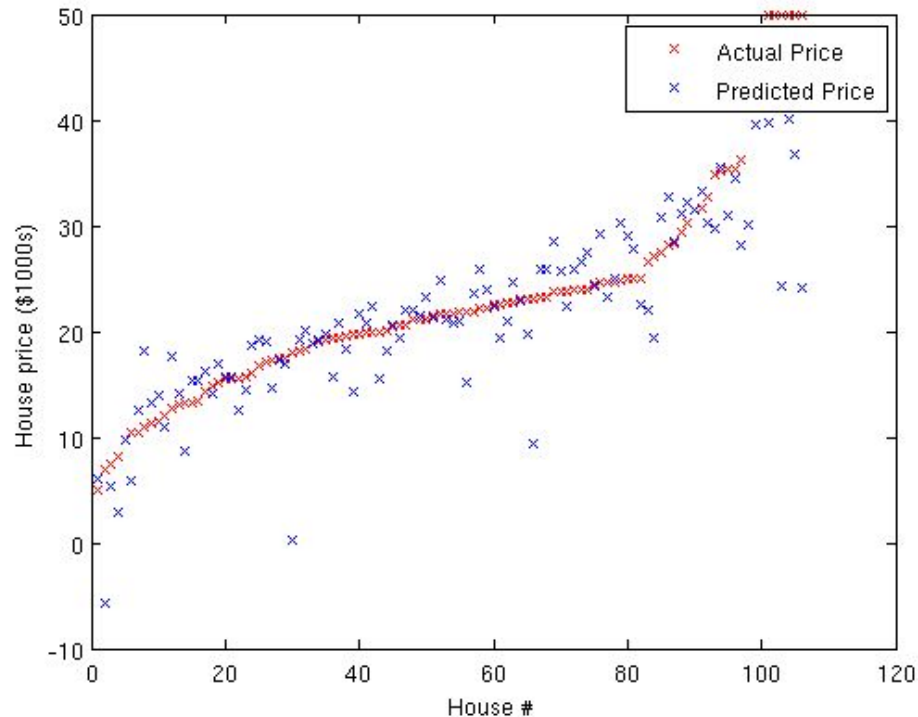
Camada de Ativação

- Também presente nas redes neurais MLP, mas de forma embutida.
- Embutida ou explícita em redes neurais convolucionais.
 - ReLu
 - Softplus
 - Softmax
 - tanh
 - sigmoid
 - linear



Cost function

- O quão errado foi o resultado.
- Regressão Linear ou Logística
- Geralmente o erro retropropagado é baseado na acurácia
- Gradiente descendente estocástico



Cost function - Least Square

$$J(\theta) = \frac{1}{2} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2.$$

Treino

- Mesmos passos de uma rede neural MLP.
 1. Carrega base
 2. Passa imagens pela rede e gera um resultado
 3. Calcula erro
 4. Retropropaga erro e ajusta pesos
 5. Refaz 2,3,4 por N épocas ou até chegar a um resultado esperado.
 6. Valida/Testa resultado

E se utilizássemos funções baseadas em outras métricas?

- Especificidade (recall)
- Sensibilidade
- Precisão
- Outros tipos
 - Weighted
 - Mix

Matriz confusão

- Representa a quantidade de FP, FN, TP, TN.
- Basta inserir minha atual e quanto foi classificado pra cada classe

		Predicted		
		Cat	Dog	Rabbit
Actual class	Cat	5	3	0
	Dog	2	3	1
	Rabbit	0	2	11

Matriz confusão - TP

- Diagonal

		Predicted		
		Cat	Dog	Rabbit
Actual class	Cat	5	3	0
	Dog	2	3	1
	Rabbit	0	2	11

Matriz confusão - TN

- Todo mundo corretamente classificado como não eu

		Predicted		
		Cat	Dog	Rabbit
Actual class	Cat	5	3	0
	Dog	2	3	1
	Rabbit	0	2	11

		Predicted		
		Cat	Dog	Rabbit
Actual class	Cat	5	3	0
	Dog	2	3	1
	Rabbit	0	2	11

		Predicted		
		Cat	Dog	Rabbit
Actual class	Cat	5	3	0
	Dog	2	3	1
	Rabbit	0	2	11

Matriz confusão - FP

- Não eu classificado como eu

		Predicted		
		Cat	Dog	Rabbit
Actual class	Cat	5	3	0
	Dog	2	3	1
	Rabbit	0	2	11

		Predicted		
		Cat	Dog	Rabbit
Actual class	Cat	5	3	0
	Dog	2	3	1
	Rabbit	0	2	11

		Predicted		
		Cat	Dog	Rabbit
Actual class	Cat	5	3	0
	Dog	2	3	1
	Rabbit	0	2	11

Matriz confusão - FN

- Eu classificados como não eu.

		Predicted		
		Cat	Dog	Rabbit
Actual class	Cat	5	3	0
	Dog	2	3	1
	Rabbit	0	2	11

		Predicted		
		Cat	Dog	Rabbit
Actual class	Cat	5	3	0
	Dog	2	3	1
	Rabbit	0	2	11

		Predicted		
		Cat	Dog	Rabbit
Actual class	Cat	5	3	0
	Dog	2	3	1
	Rabbit	0	2	11

Matriz confusão - Por classe

	Classe Prevista		
		Sim	Não
Classe Original	Sim	VP	FN
	Não	FP	VN

Matriz confusão - Metricas

- Sensibilidade (recall) :

$$TPR = \frac{TP}{P} = \frac{TP}{TP + FN}$$

- Especificidade:

$$SPC = \frac{TN}{N} = \frac{TN}{FP + TN}$$

- Precisão:

$$PPV = \frac{TP}{TP + FP}$$

- Mix:

$$\frac{(acc + 2*re + sp)}{4}$$

Outras funções de custo

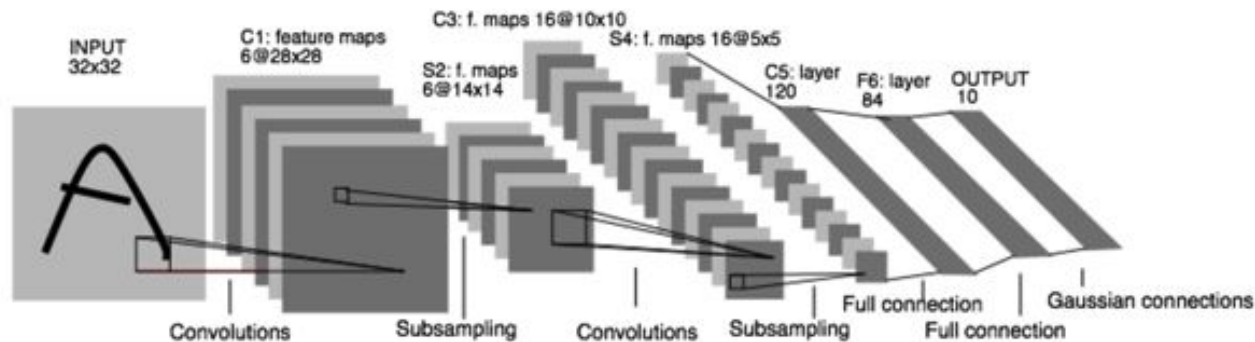
- Recall

$$\frac{m}{\sum_{i=1}^m recall(i)} \quad 1 - \frac{1}{m} \sum_{i=1}^m recall(i)$$

- Especificidade, precisão, mix ...

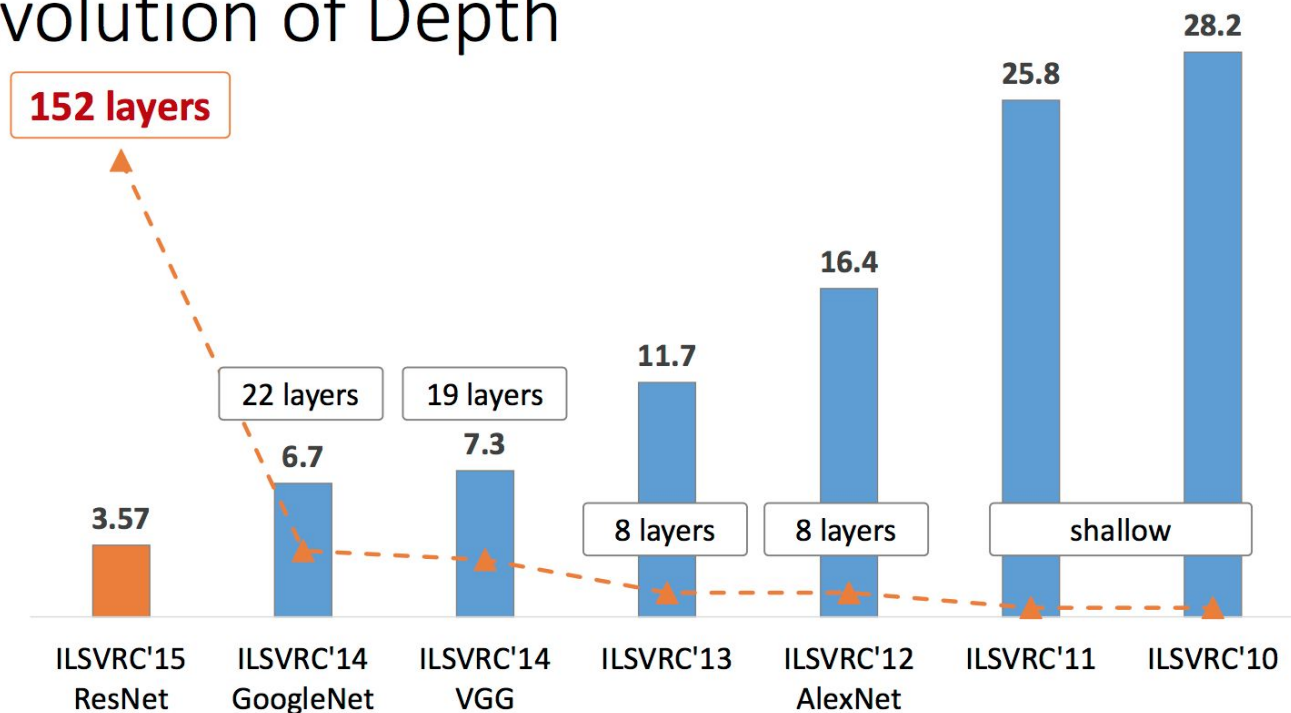
LeNet 1998

- Base de dados MNIST
- Yann LeCun
- Acima de 97%
- 5 camadas



Outras redes conhecidas - ResNet

Revolution of Depth



Problemas

- Definir uma função de custo boa.
- Um pouco mais complexa que o inverso de uma função f' .

Sugestões?

— Roberto Matheus Pinheiro Pereira —
