

# Deep Residual Learning for Image Recognition

Kamila de Almeida Benevides  
kab@ic.ufal.br

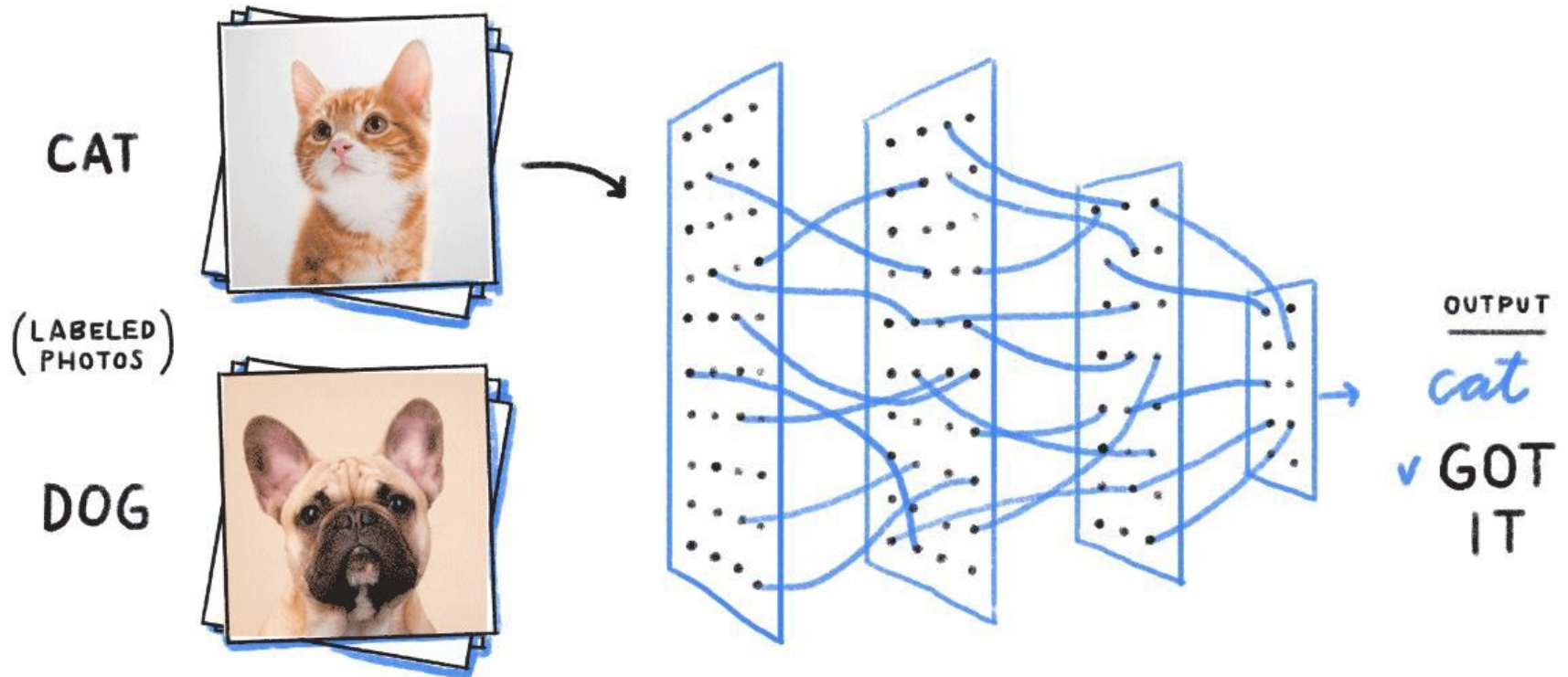
# Introdução

O estudo apresenta uma estrutura de aprendizado de uma rede neural residual.

É apresentado experimentos que dizem que essa rede residual é extremamente profunda e simples de otimizar, produzindo resultados substancialmente melhores e com ganho de precisão de profundidade maior.

Com o objetivo de classificar uma imagem pixel a pixel.

# Classificação tradicional de imagem



# Classificação pixel a pixel

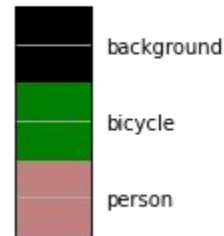
input image



segmentation map

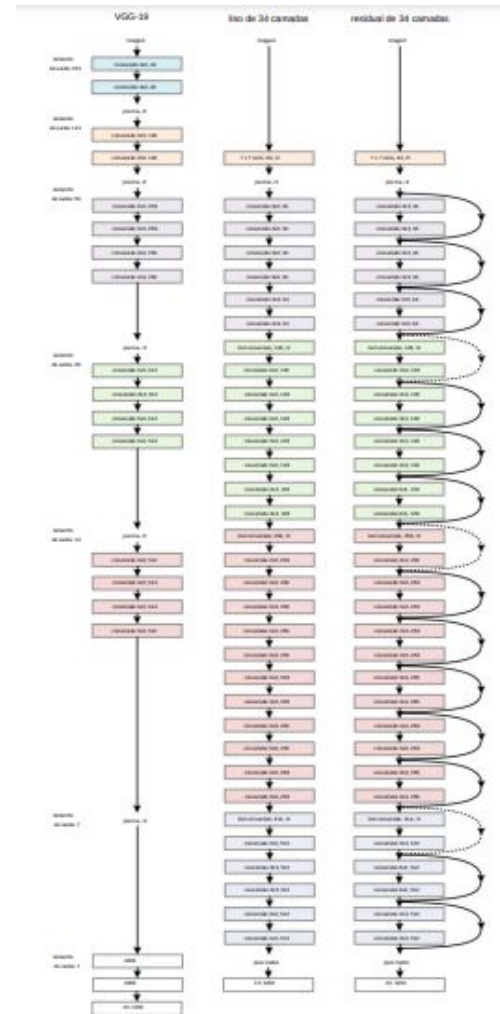


segmentation overlay



# Arquiteturas

- Arquitetura simples
  - Se o tamanho do mapa de características for reduzido pela metade, o número de filtros é duplicado para manter a complexidade de tempo por camadas.
- Arquitetura simples profunda
  - similar a arquitetura simples, onde possui mais camadas.
  - A informação que chega na última camada é completamente diferente da informação que chega no início.
- Arquitetura residual profunda
  - Com base na rede simples profunda, a rede residual é feita adicionando atalhos entre as camadas.
  - Existe uma preservação parcial das informações da primeira até a última camada



# Pré-processamento

- A imagem é redefinida para  $224 \times 224 \times 3$
- A média dos pixels é subtraída

# Base de dados

- ImageNet e CIFAR-10 (usada para treinamento inicial dos pesos)
  - ImageNet consiste em 1000 classes, onde os modelos são treinados nas 1,28 milhões de imagens de treinamento e avaliados nas 50k imagens de validação
  - CIFAR-10 consiste em 50k imagens de treinamento e 10k imagens de teste em 10 aulas.
- Pascal e MS COCO (usada com transferência de conhecimento)
  - MS COCO envolve 80 categorias de objetos. as métricas PASCAL VOC e a métrica COCO padrão, 80k de imagens no conjunto train para treinamento e 40k de imagens no conjunto de validação

# Treinamento

- É usado SGD com um tamanho de mini-lote de 256.
- A taxa de aprendizado começa em 0,1 e é dividida por 10 quando o erro se estabiliza, e os modelos são treinados para até  $60 \times 10^4$  iterações.
- É Usado um decaimento de peso de 0,0001 e um momento de 0,9.
- Nos testes, para estudos de comparação, adotamos o teste padrão de 10 colheitas.



# Resultados

method			error (%)
Maxout [10]			9.38
NIN [25]			8.81
DSN [24]			8.22
	# layers	# params	
FitNet [35]	19	2.5M	8.39
Highway [42, 43]	19	2.3M	7.54 (7.72±0.16)
Highway [42, 43]	32	1.25M	8.80
ResNet	20	0.27M	8.75
ResNet	32	0.46M	7.51
ResNet	44	0.66M	7.17
ResNet	56	0.85M	6.97
ResNet	110	1.7M	<b>6.43</b> (6.61±0.16)
ResNet	1202	19.4M	7.93