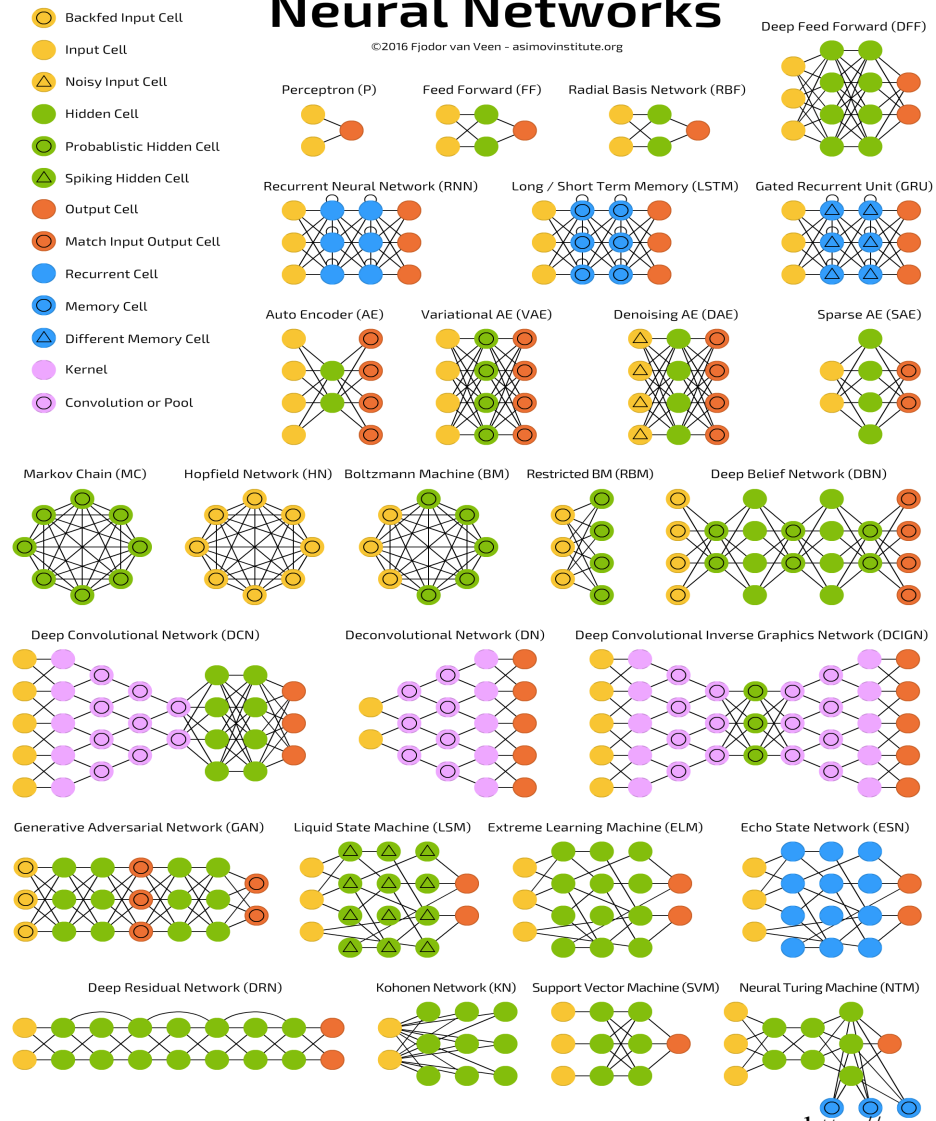


# Deep Neural Networks: Contextualização

Gisele L Pappa

# A mostly complete chart of Neural Networks

©2016 Fjodor van Veen - asimovinstitute.org



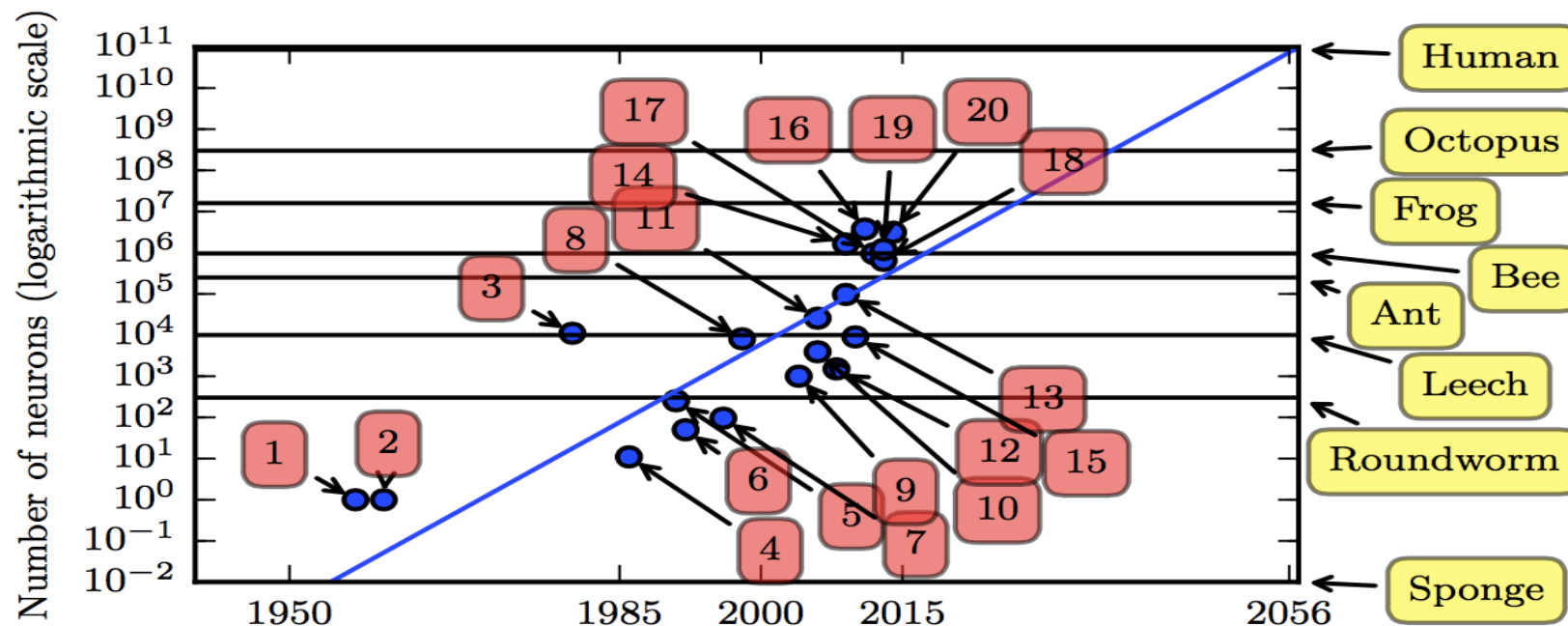
# Contexto Histórico

- Deep learning:
  1. Tem uma história longa, já teve vários nomes e ondas de popularidade
  2. Se tornou muito útil com o aumento do volume de dados disponível
  3. Passou a gerar modelos maiores e mais complexos com o aumento do poder computacional
  4. Tem resolvido problemas cada vez mais complexos com taxas de acerto cada vez maiores

# As diferentes faces de Deep Learning

- 1940-1960: Sistemas cibernéticos
  - McCulloch and Pitts, 1943
  - Perceptron, 1958
- 1960-1990: Conexionismo
  - Back-propagation, 1986
  - Long short-term memory (LSTM), 1997
- 2006: Deep Learning
  - Deep belief networks treinadas eficientemente com *greedy layer-wise pre-training*, 2006

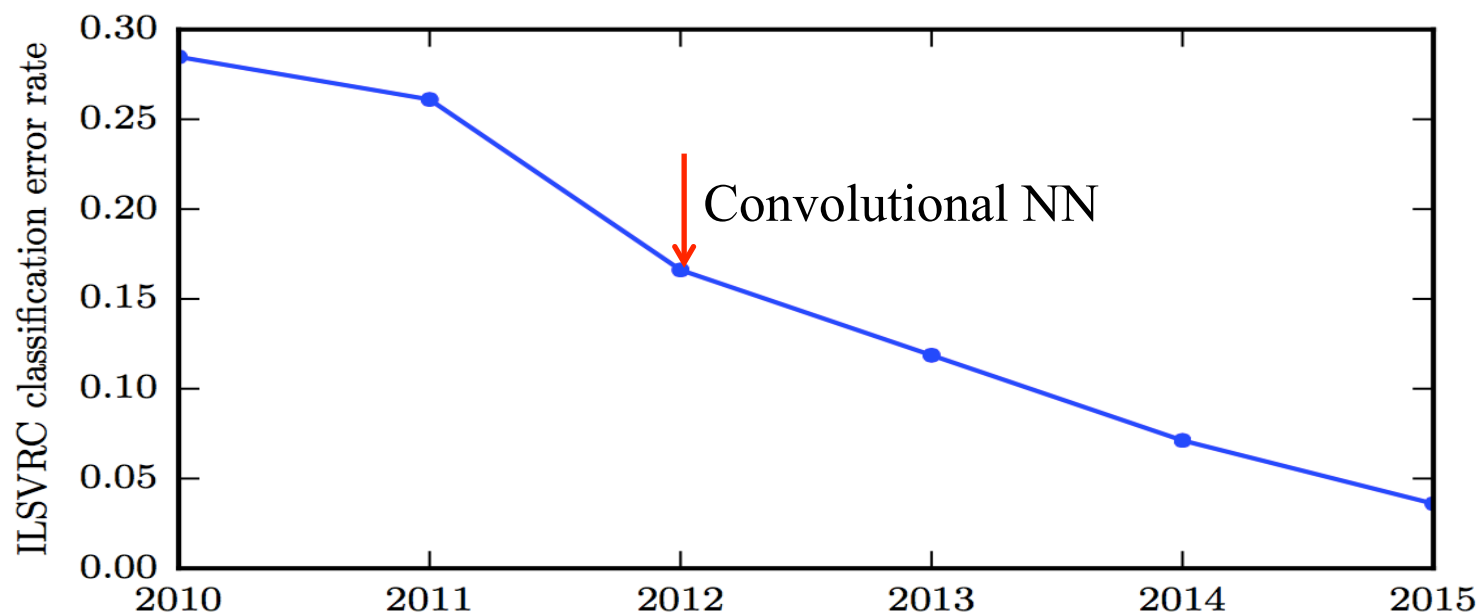
# Número de neurônios



- Redes neurais dobram de tamanho a cada 2.4 anos

# Qualidade de predição dos modelos

- Decaimento do erro do ImageNet Visual Recognition Challenge



# O que caracteriza uma rede profunda?

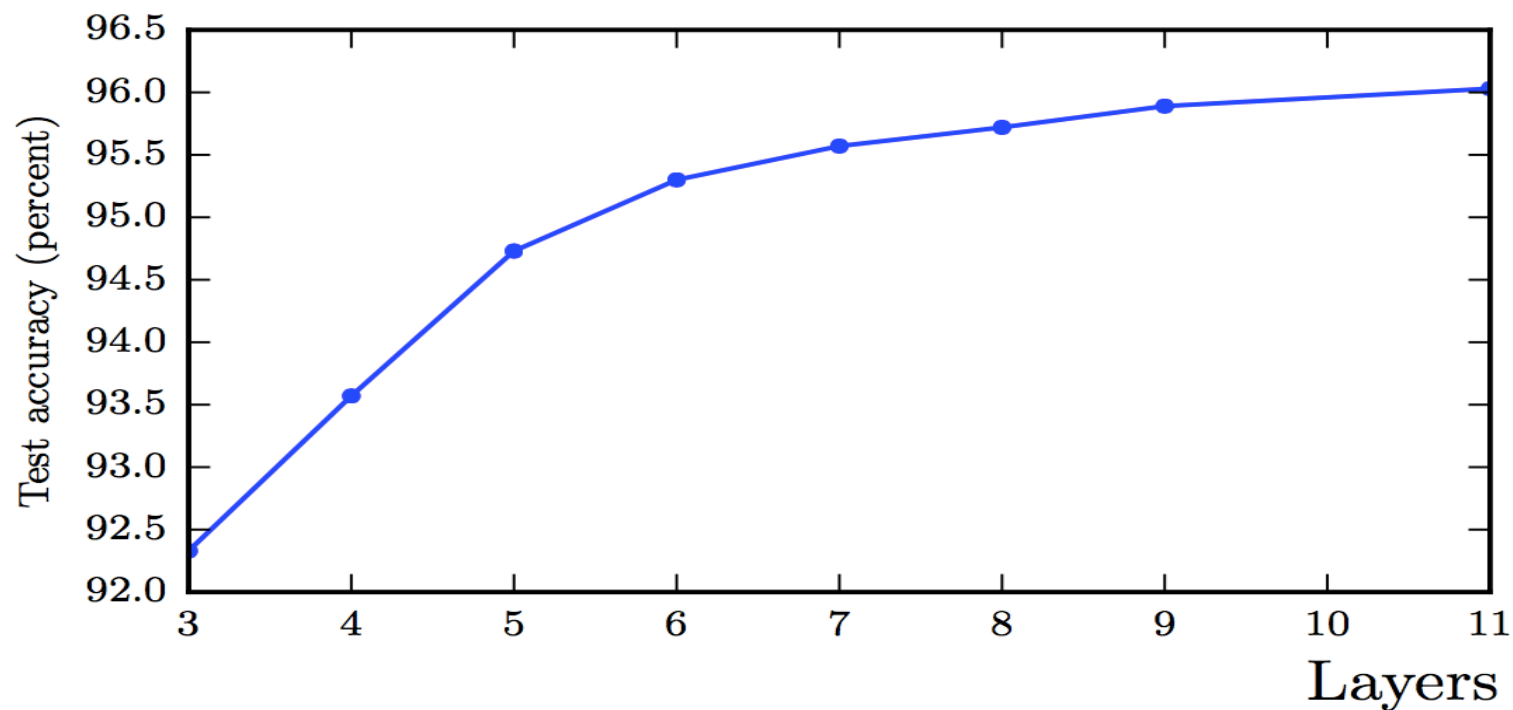
- Quantas camadas escondidas são necessárias?
  - Trabalho clássico que ajudou a popularizar o termo *deep learning* usa 3 camadas escondidas (Hinton et al 2006)
- Trabalhos que se auto denominam como “very deep networks” usam mais de 10 camadas escondidas
- Trabalham que se auto denominam “extremely deep” usam de 50 até 1000 camadas escondidas

# Teorema da Aproximação Universal

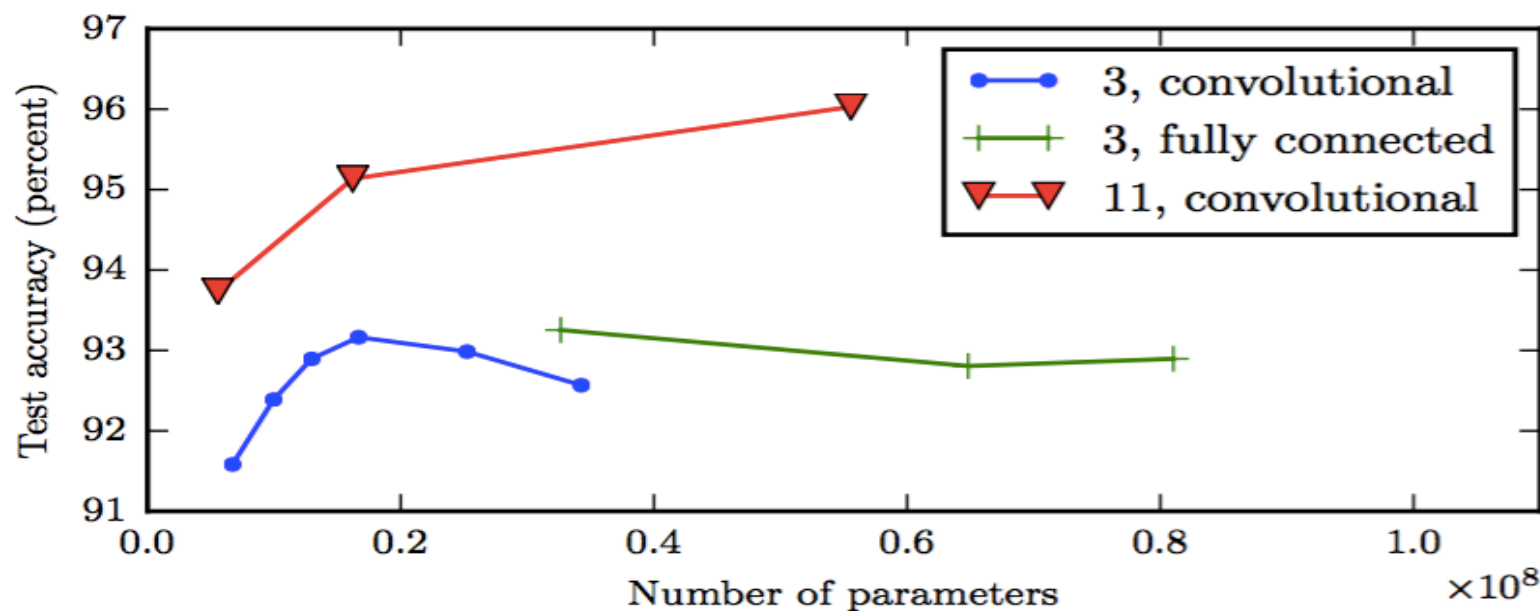
- Uma camada é suficiente para representar uma aproximação de qualquer função com um grau de acurácia arbitrário
- Então por que utilizar mais camadas?
  - Redes “rasas” podem precisar de um número exponencialmente maior de neurônios
  - Redes “rasas”, no geral, sofrem mais de overfitting



# Generalização vs número de camadas

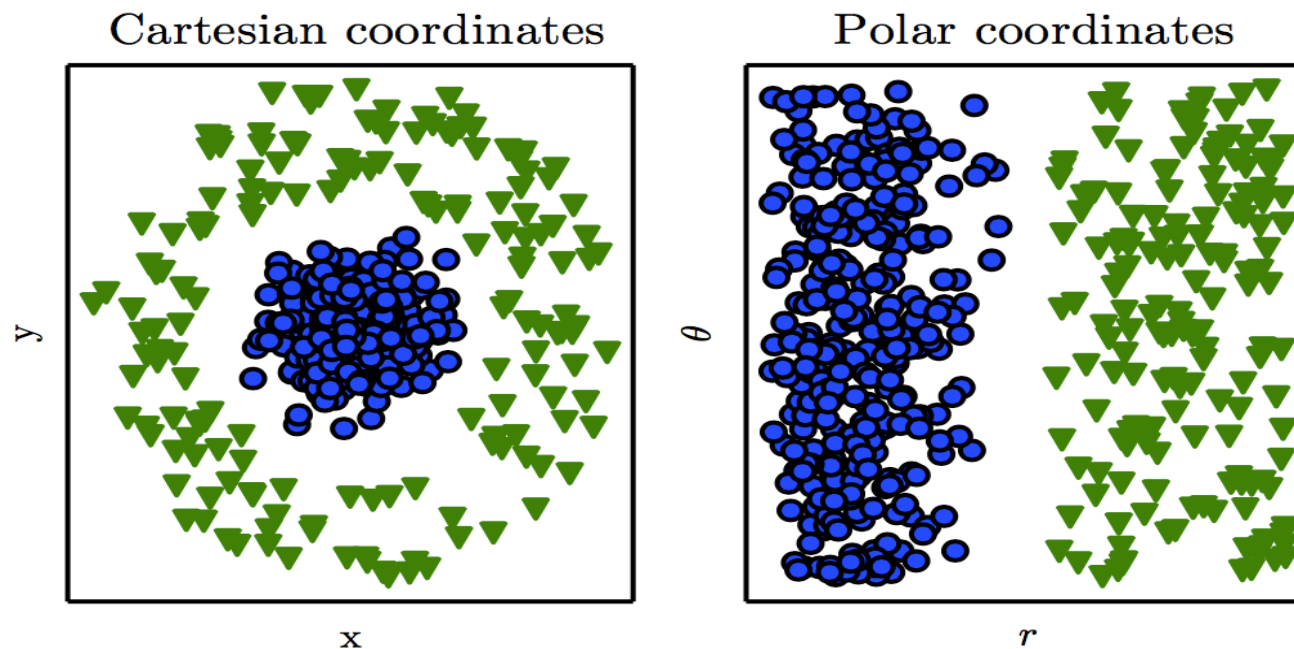


# Modelos mais “largos” e “rasos” generalizam menos



# Importância da Representação

- Aprendizado depende de uma representação apropriada da informação

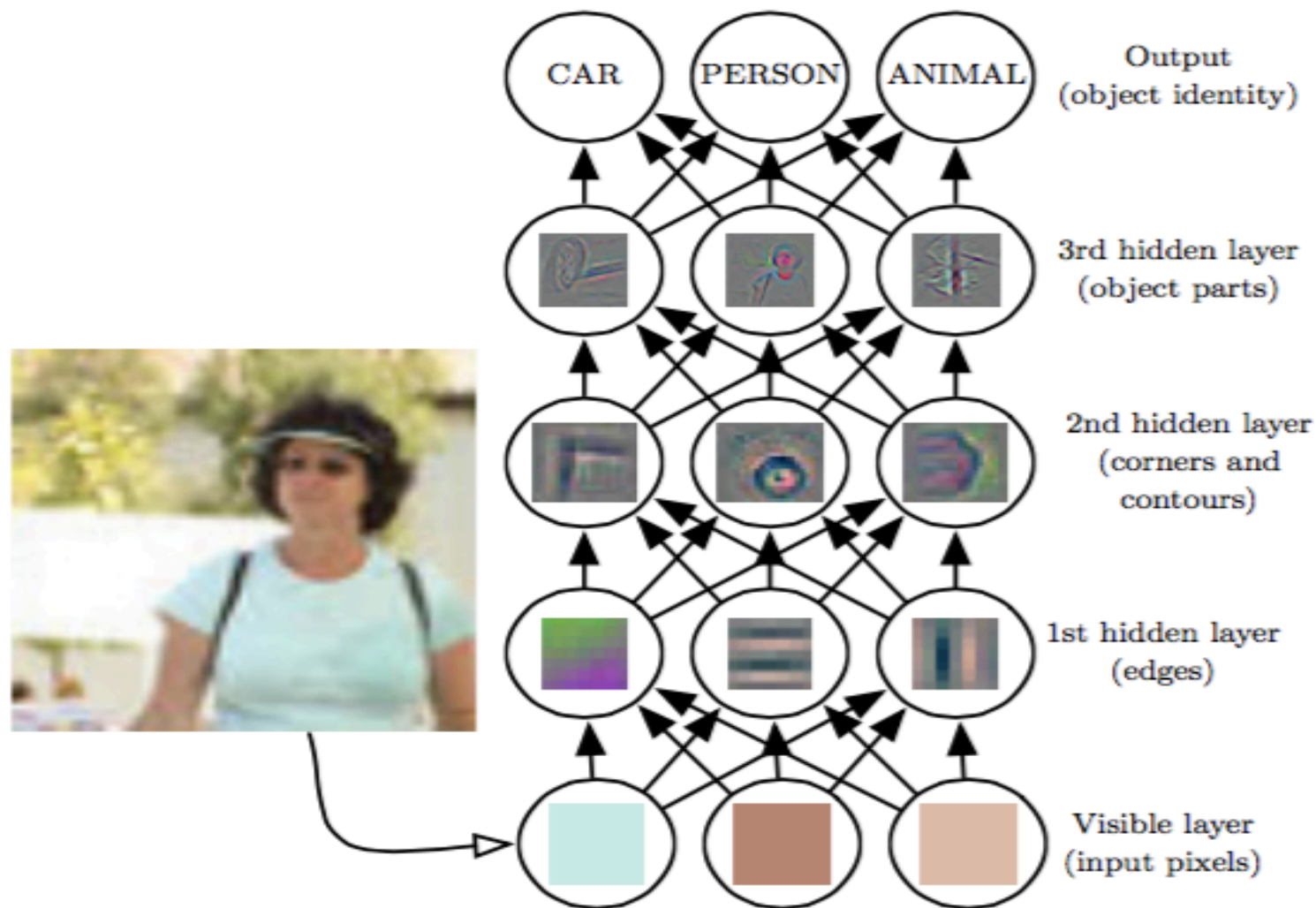


# Importância da Representação

- Para algumas tarefas, não é difícil escolher o conjunto apropriado de características para representá-las
  - Ex: previsão do tempo
- Para outras tarefas, é difícil saber que características utilizar
  - Aprendizado de representação (*Representation learning*) tem uma papel crucial
  - Ex: detecção de carros em fotos (rodas são formas geométricas simples mas podem ser afetadas por sombras, luz, etc)

# *Representation Learning*

- Deep Learning resolve o problema de aprendizado de representação utilizando representações complexas que são expressas em termos de outras representações mais simples
  - EX: como DL representa o conceito da imagem de uma pessoa combinando conceitos mais simples



# Feedforward Deep Network

- O exemplo mais popular de DL é uma feedforward deep network
- Em relação ao MLP clássico temos:
  - Redes maiores treinadas com mais exemplos
  - Função de ativação das camadas escondidas muda devido ao “*vanishing gradient problem*”
  - Pequenas modificações algorítmicas no BP

# Deep Neural Networks: Contextualização

Gisele L Pappa