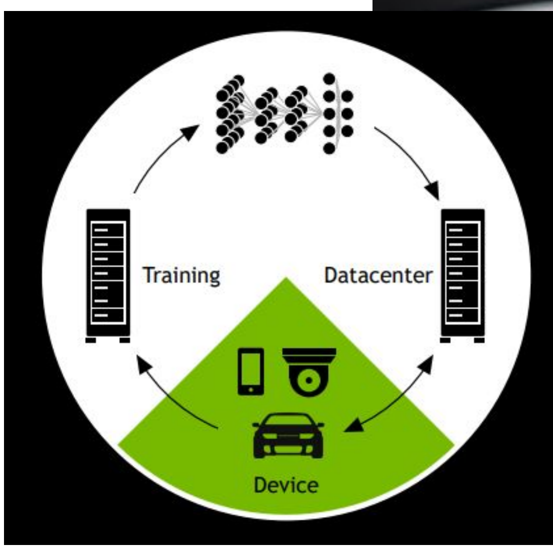
Universidade Federal do Paraná (UFPR)
Especialização em Engenharia Industrial 4.0

AI / ML / Deep Learning



David Menotti
www.inf.ufpr.br/menotti/am-192

Objetivos

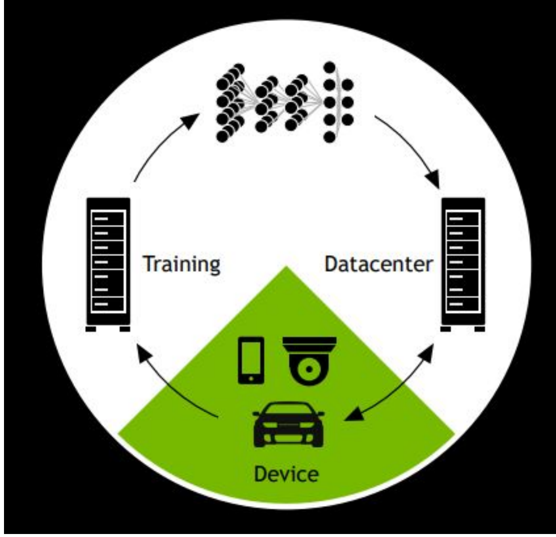
- *Deep Learning - Breakthroughs*
  - Histórico e eventos

Deep Learning

- ***Deep Learning*** é como sexo na adolescência.
  - Todo mundo fala sobre
  - Ninguém realmente sabe como fazer
  - Todo mundo pensa que todo mundo está fazendo
  - Então, todo mundo diz que está fazendo

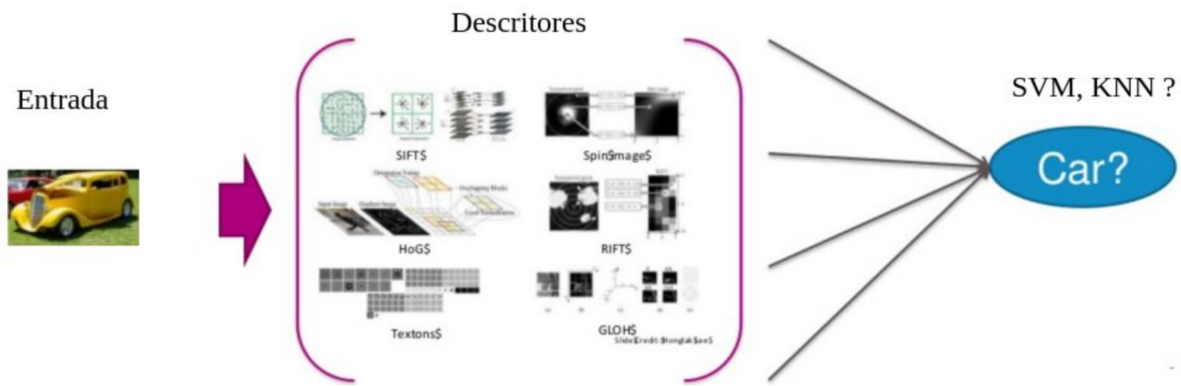
Adaptado de Dan Ariely (Big Data)

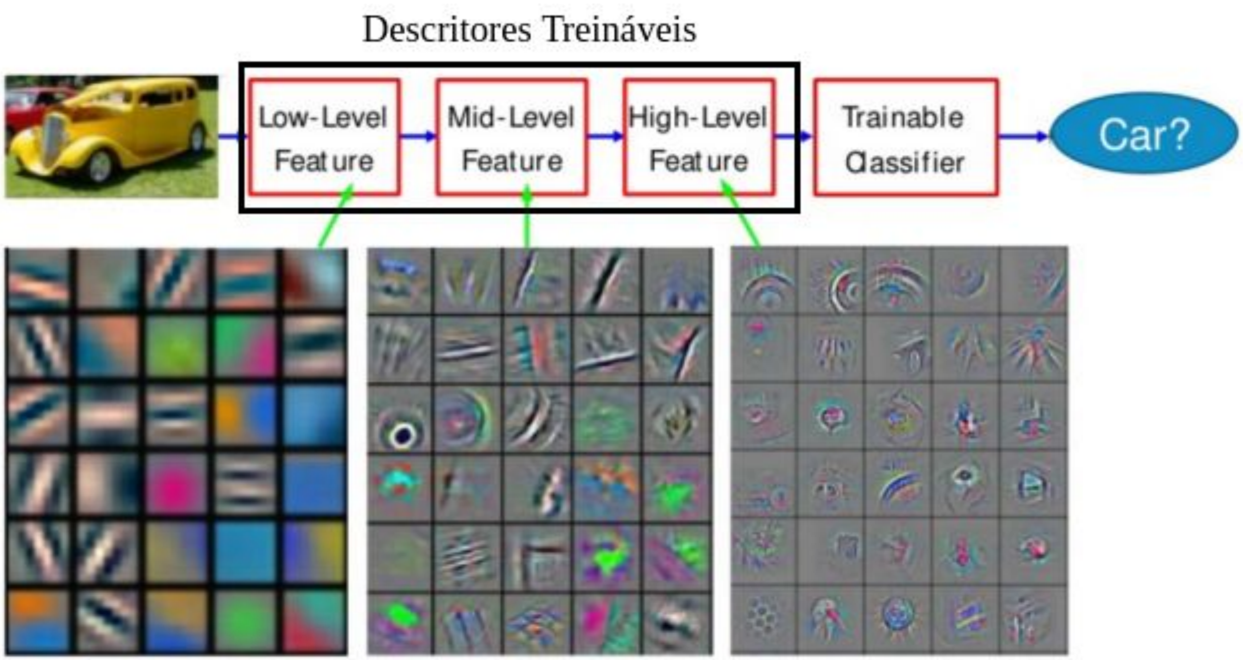
Deep Learning



- Fatores que influenciam
  - Desenvolvimento de Algoritmos
  - Hardware capaz de processar grandes massas de dados (GPU)
  - Disponibilidade de dados anotados
    - Imagens no Facebook
    - Voz e Texto no Google

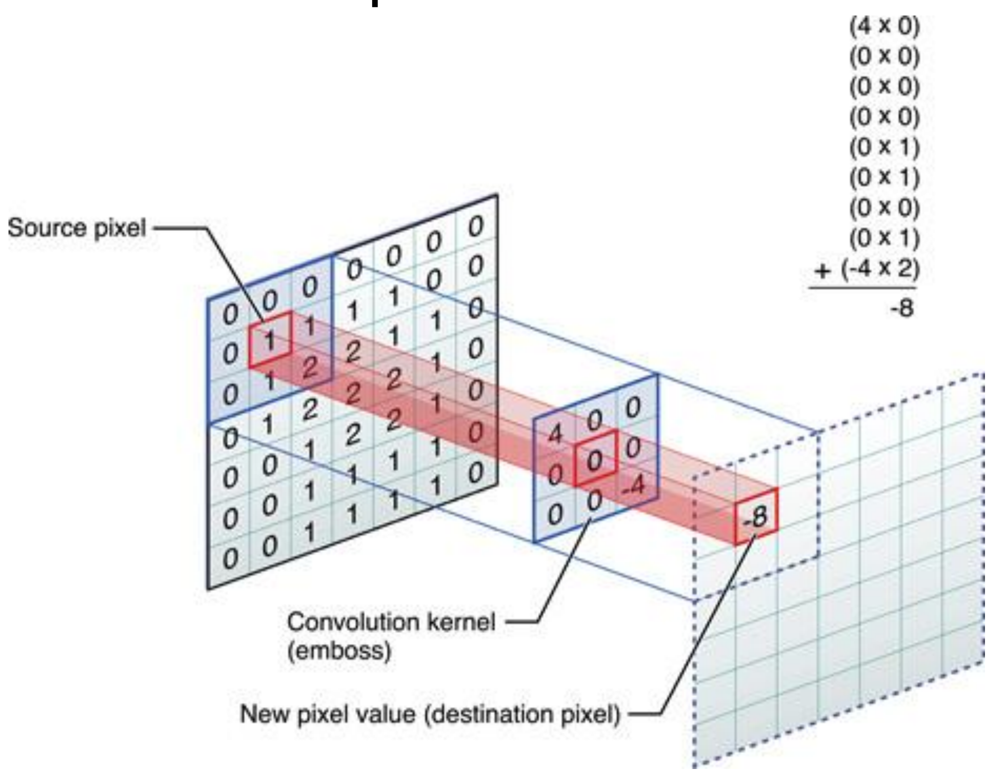
Tradicional x Deep Learning

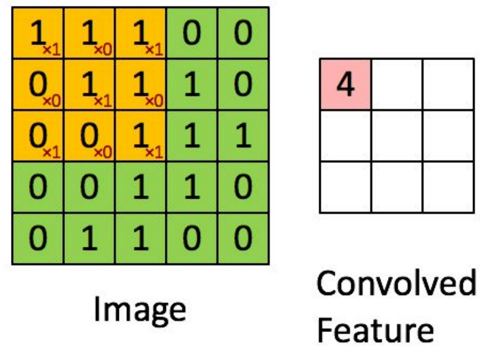




Convolução

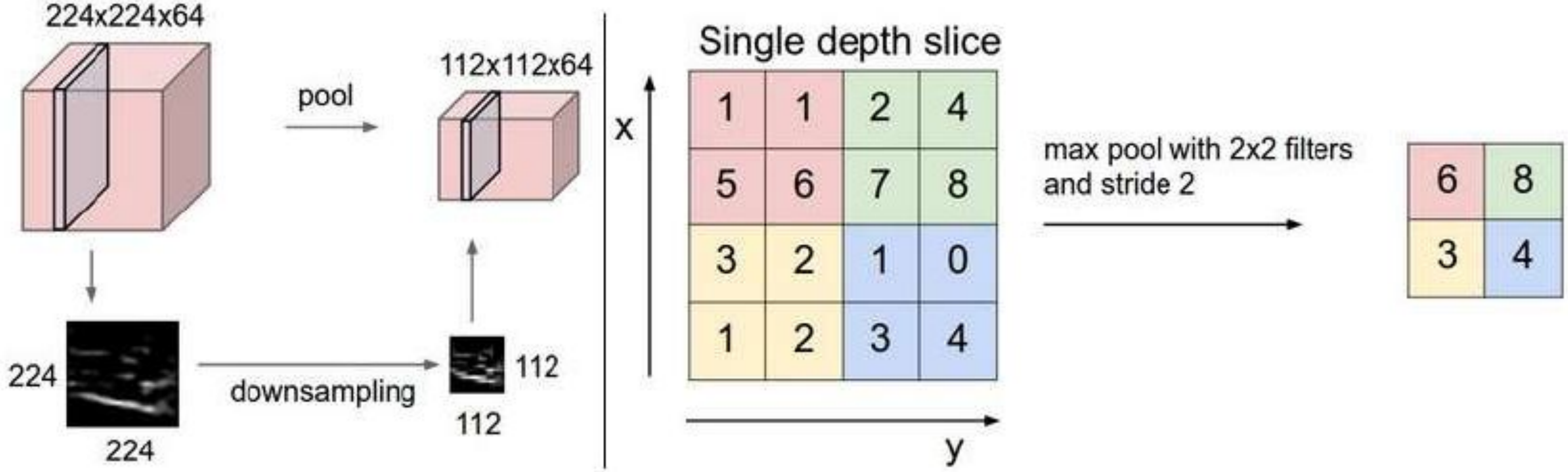
- Chave em *Deep Learning*
  - Pesos compartilhados





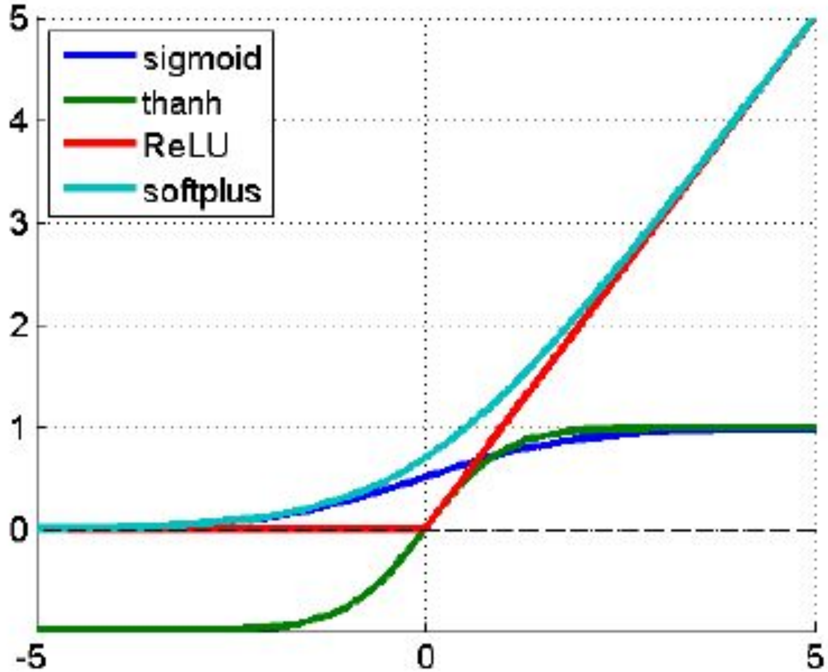
Pooling

- Largamente utilizado
  - Reduz **dim**. representação



Funções de ativação

- ReLU elimina o problema do **vanish gradient**



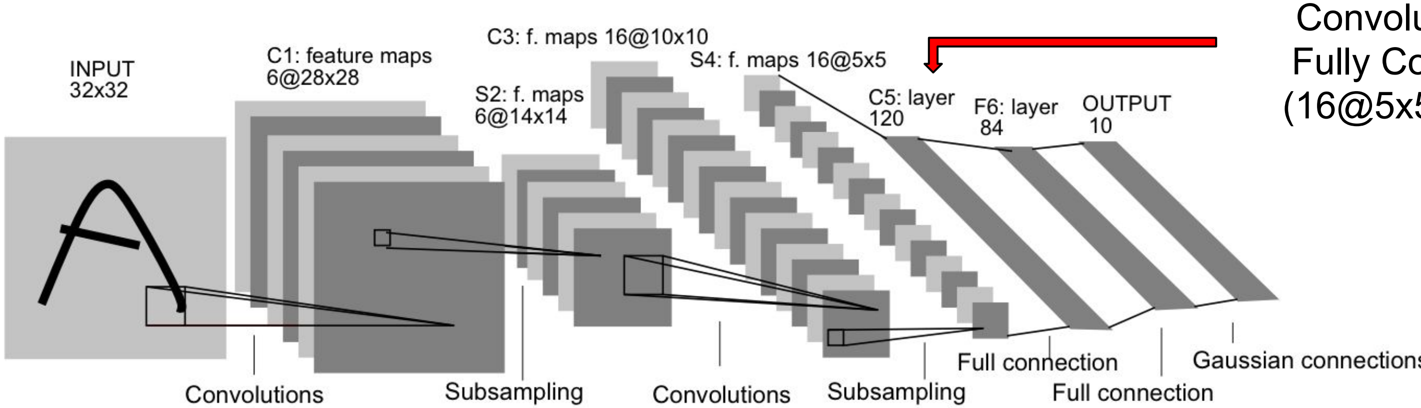
$$f(x) = \max(0, x)$$

$$f'(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases}$$

- Yoshua **Bengio**, Ian Goodfellow and Aaron Courville. Deep Learning. MIT Press. http://www-labs.iro.umontreal.ca/~bengioy/DLbook (citado por 3467 jul/2018)

LeNet (Lecun)

- 1998 - 1a. CNN implementada com sucesso (Bell Labs)
  - Reconhecimento de Dígitos Manuscritos
    - Imagem de entrada de **32x32 pixels**
    - MNIST Dataset (10 classes [0-9])
    - #60k *train* / #10k *test*
    - 0,95% / ~345 k conexões ~60k parâmetros



Convolutional &
Fully Connected
(16@5x5+1)x120

- Y. LeCun, L. Bottou, Y. Bengio & Patrick Haner, Gradient Based Learning Applied to Document Recognition, Proc. of IEEE, 1998

LeNet (Lecun)

- Arquitetura
  - Pesos & Bias (340.908 / 60.000)
  -
  - C1: conv: 6@5x5 => 122.304 (6x(5x5+1)=156) => 6@28x28
  - S2: pool: 2x2 => 5.880 (6x2=12) => 6@14x14
  - C3: conv: *@5x5 => 151.600 ((*60x5x5)+16=1.516) => 16@10x10
  - S3: pool: 2x2 => 2.000 (16x2=32) => 16@5x5
  - C5: fully: 16@5x5 => 48.120 (16x5x5+1)x120 => 120
  - F6: fully: 120x84 => 10.164 (84x(120+1)) => 84**

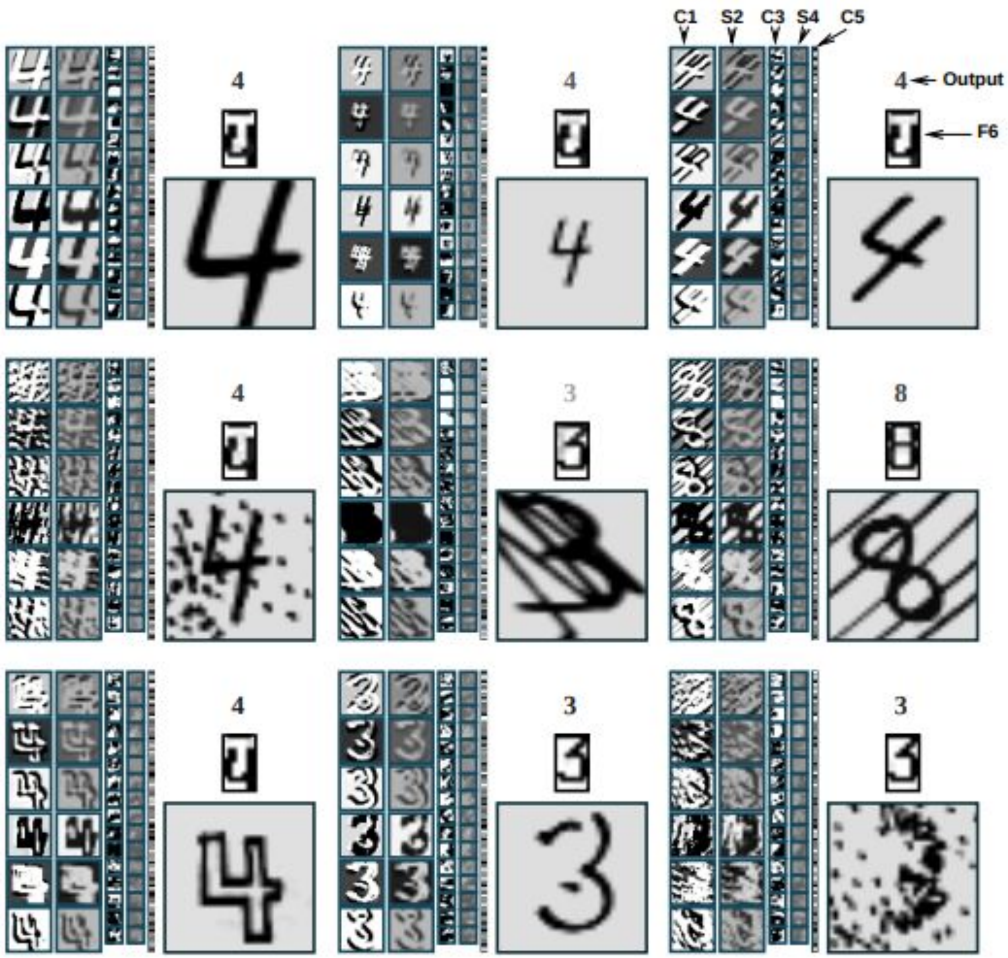
    => **60.000 total**
  - Out: fully: 84x10

Camada C3

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	X				X	X	X			X	X	X	X		X	X
1	X	X				X	X	X			X	X	X	X		X
2	X	X	X				X	X	X			X		X	X	X
3		X	X	X			X	X	X	X			X		X	X
4			X	X	X			X	X	X	X		X	X		X
5				X	X	X			X	X	X	X		X	X	X

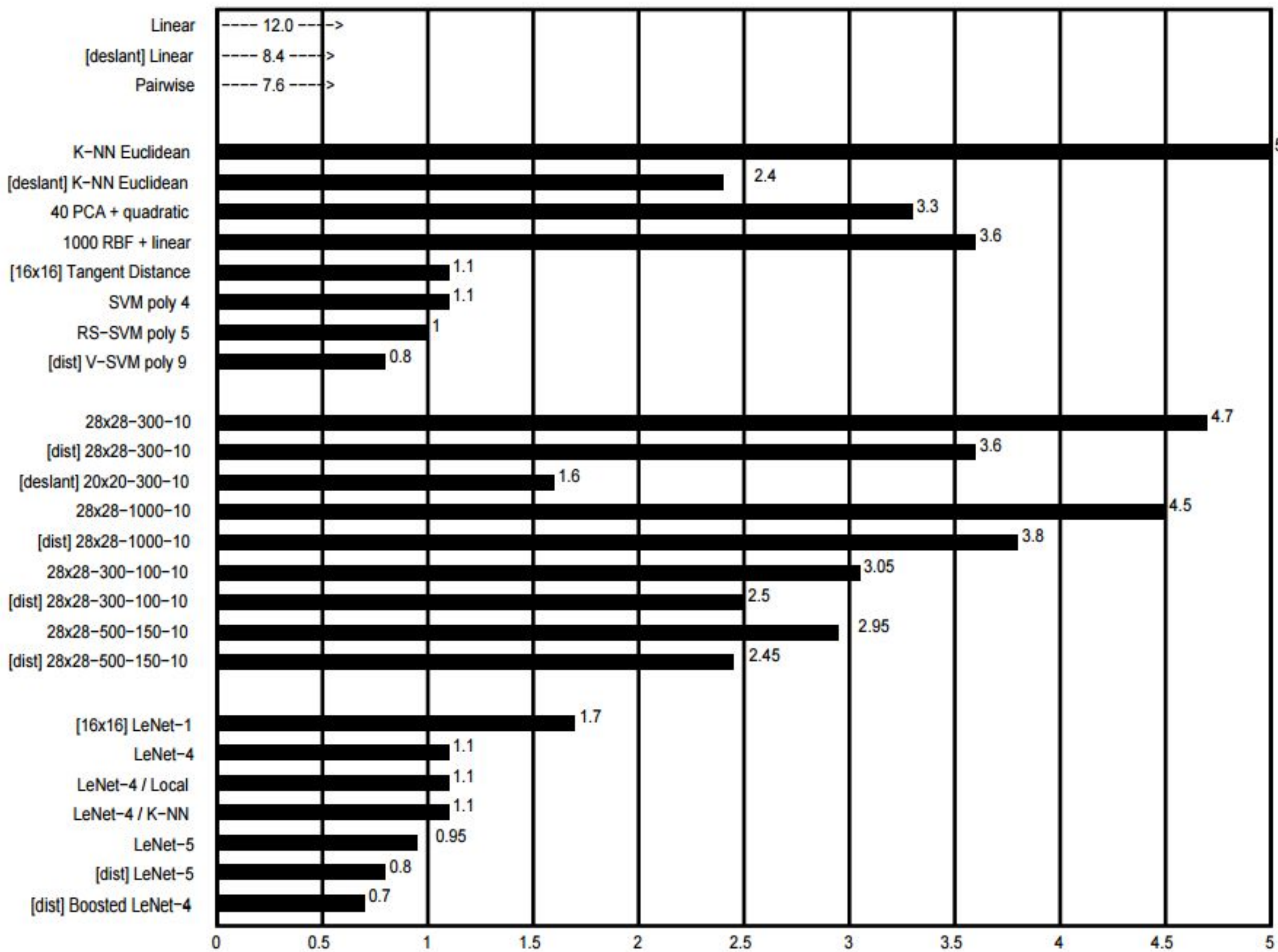
Camada F6

LeNet - Funcionamento

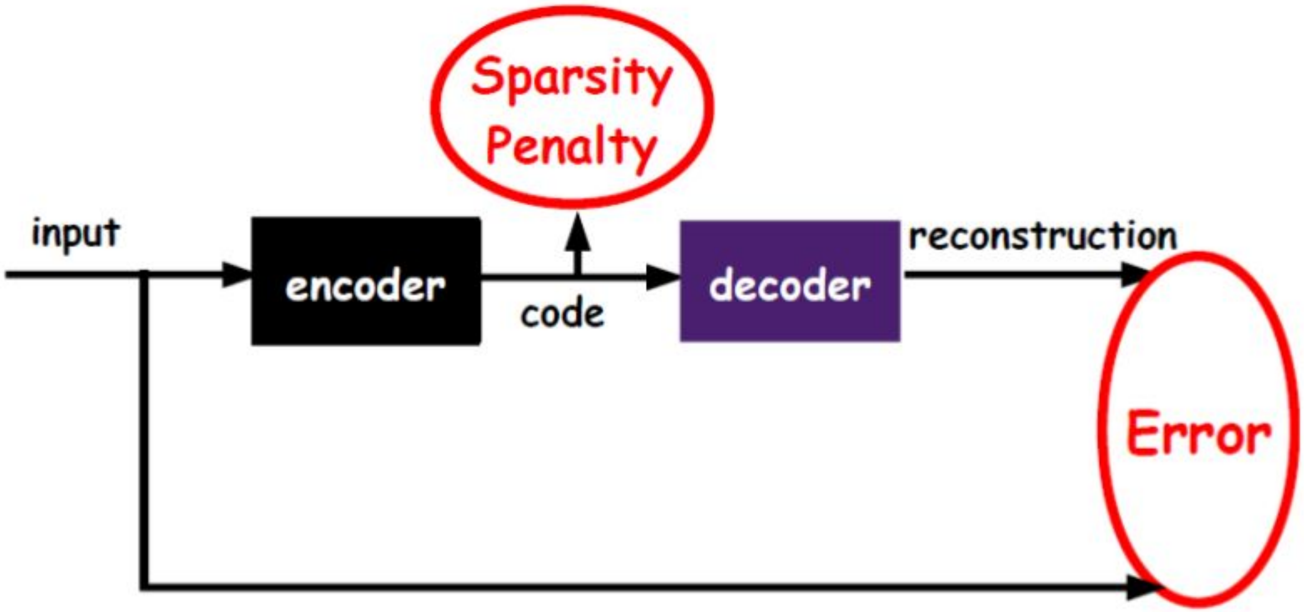


Robusto a ruídos, inclinação e forma

LeNet vs *State of the Art*



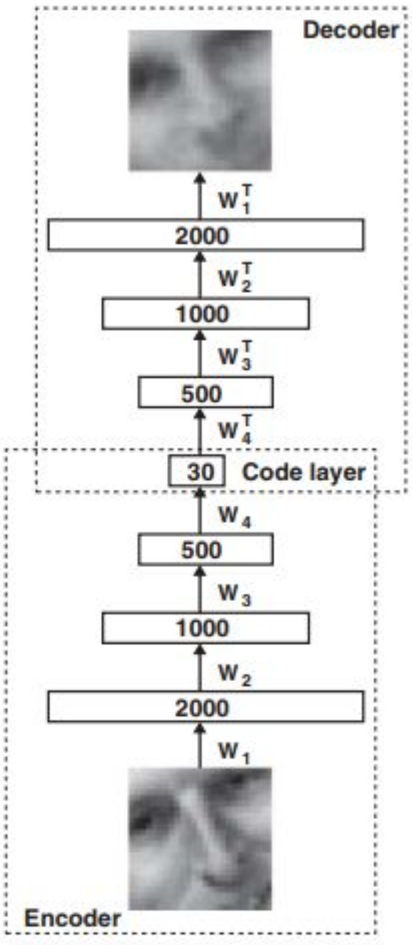
Sparse Autoencoders



- input: X code: $h = W^T X$
- loss: $L(X; W) = \|W h - X\|^2 + \lambda \sum_j |h_j|$

Eventos

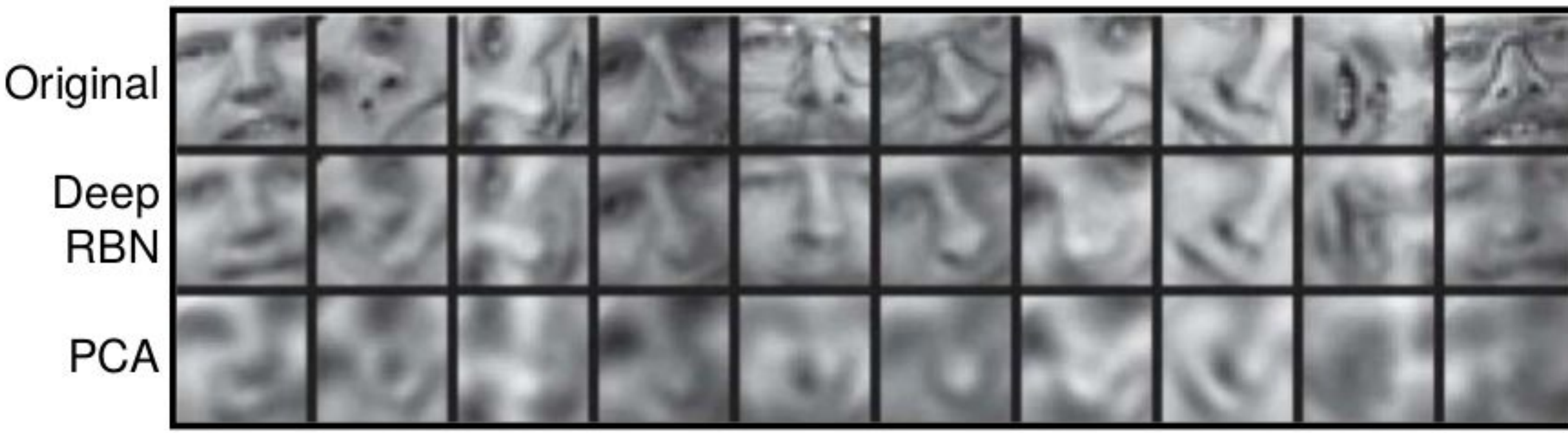
- **2006 - Autoencoders**
  - Redução de Dimensional.
  - Uso de RBM empilhados como um deep auto-encoder
    i. Treinar RBM usando imagens como saída e entrada



- G. E. Hinton & R. R. Salakhutdinov. *Reducing the Dimensionality of Data with Neural Networks*, Science 313 (2006), p. 504–507.

Eventos

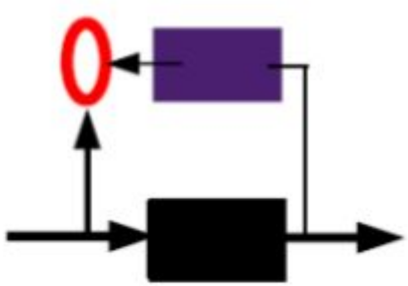
Olivetti face data, 25x25 pixel images reconstructed from 30 dimensions (625 → 30)



- G. E. Hinton & R. R. Salakhutdinov. *Reducing the Dimensionality of Data with Neural Networks*, Science 313 (2006), p. 504–507.

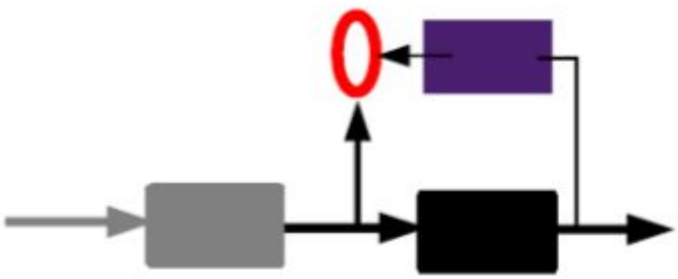
Unsupervised Pre-training

1) Dado uma imagem (não rotulada), aprenda características



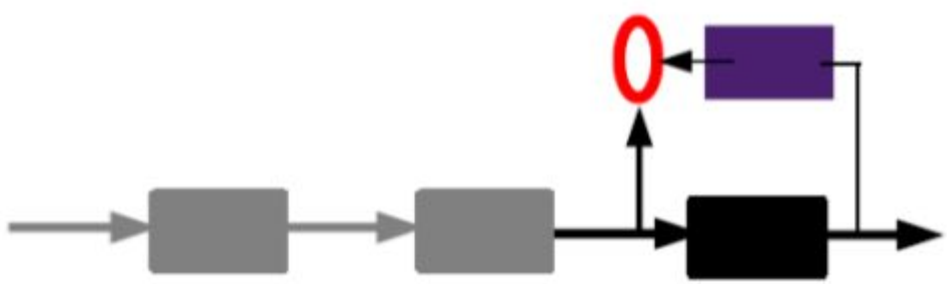
Unsupervised Pre-training

1) Dado uma imagem (não rotulada), aprenda características
2) Use o encoder para produzir novas features e treinar outra camada



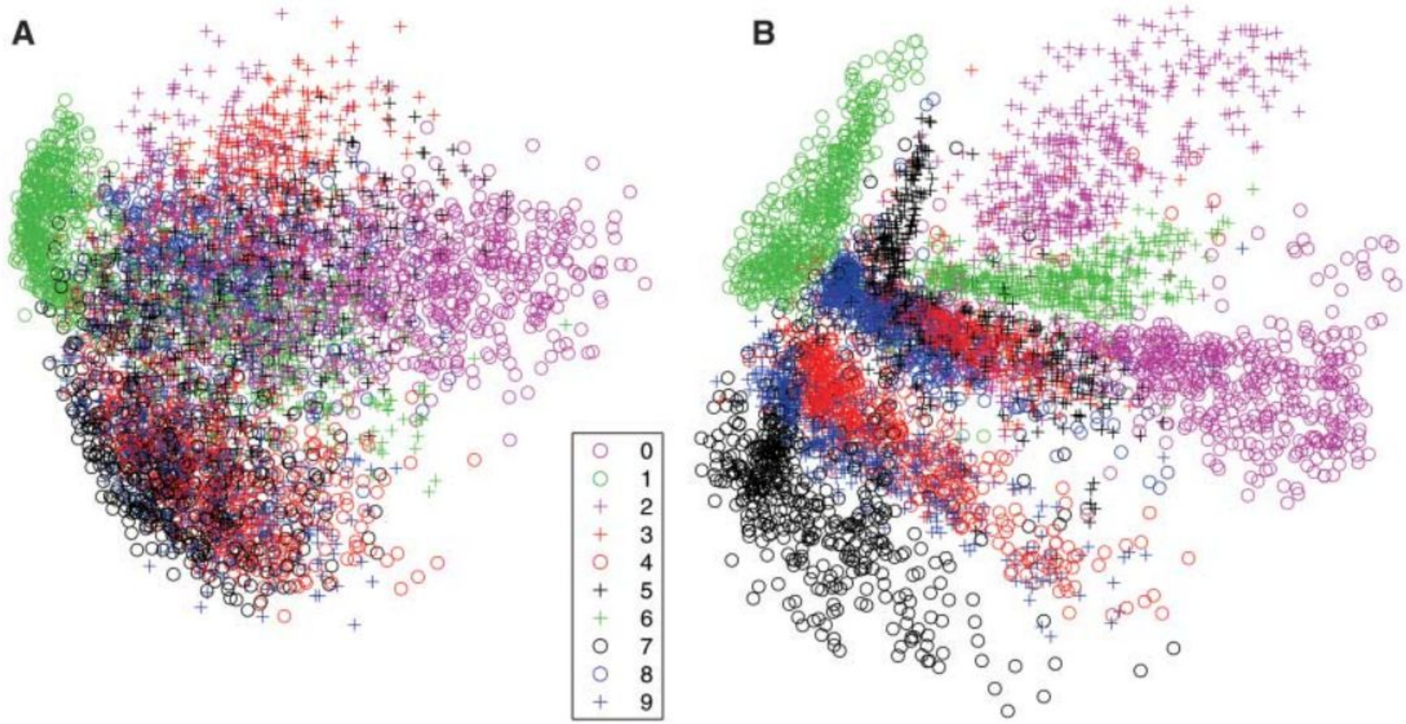
Unsupervised Pre-training

1) Dado uma imagem (não rotulada), aprenda características
2) Use o *encoder* para produzir novas *features* e treinar outra camada



Eventos

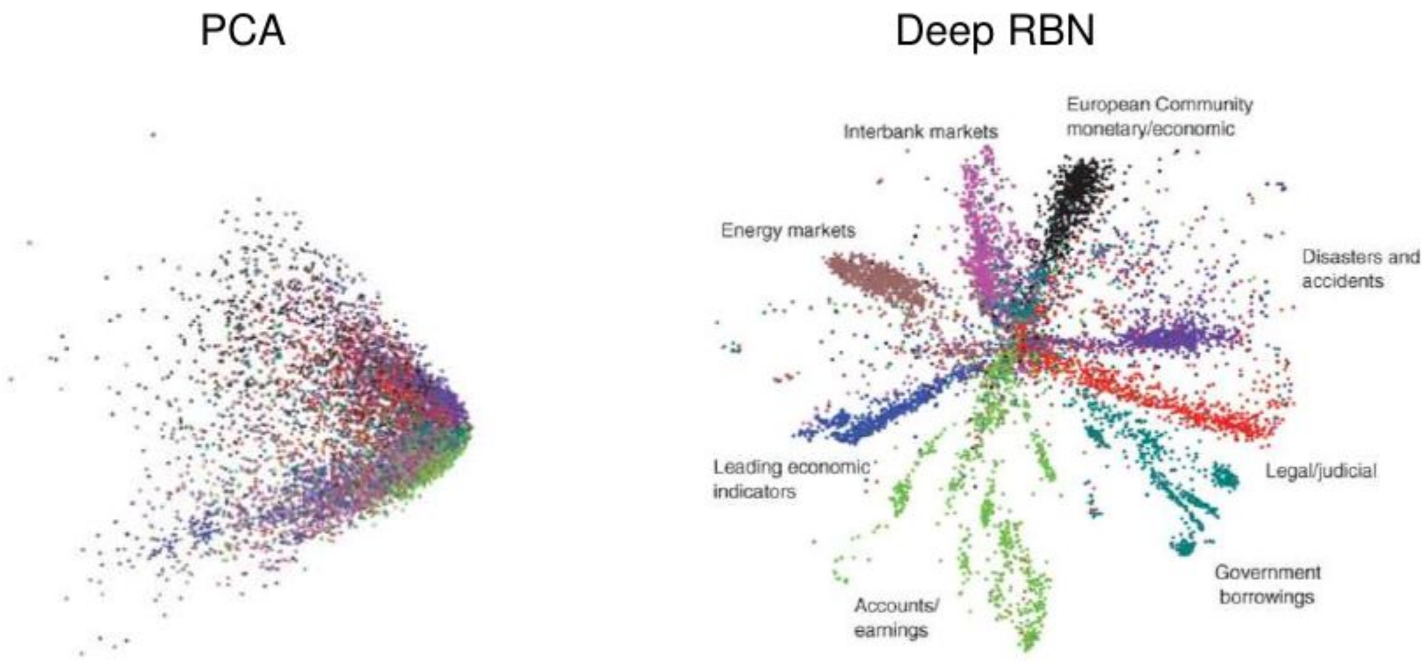
- **2006 - Autoencoders**
  - PCA vs (784-1000-500-250-2) autoencoder * $28^2 = 784$



- G. E. Hinton & R. R. Salakhutdinov. *Reducing the Dimensionality of Data with Neural Networks*, Science 313 (2006), p. 504–507.

Eventos

- **2006 - Autoencoders**
  - 804’414 news da Reuters, redução para 2D
  - PCA **vs** 2000-500-250-125-2 autoencoder



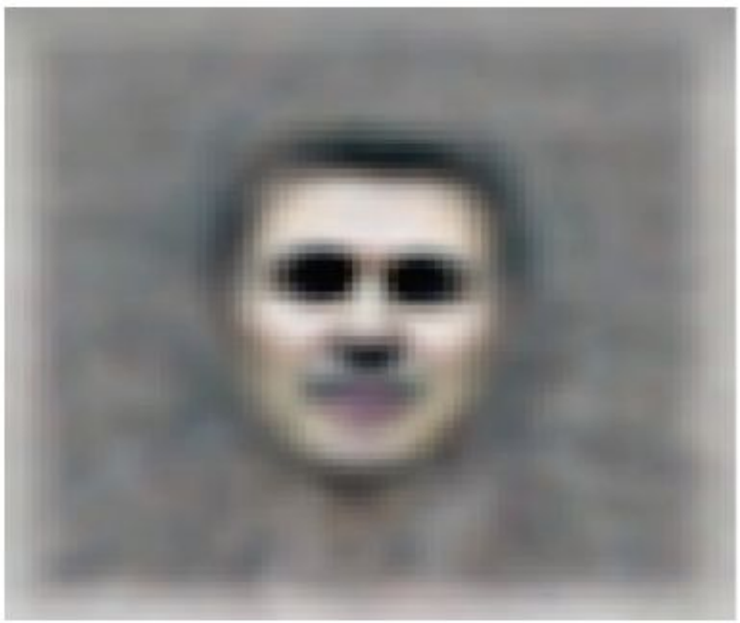
Eventos

- Dados não rotulados estão disponíveis
- Exemplos: Imagens da web
  a. Baixar 10.000.000 imagens
  b. Treinar uma DNN 9-camadas
  c. Conceitos são formados pela DNN

- Q. V. Le, M. Ranzato, R. Monga, M. Devin, K. Chen, G. S. Corrado, J. Dean, A. Ng. Building High-Level Features using Large Scale Unsupervised Learning in ICML 2012 (citado por 1610 jul/2018)

Eventos

- **Deep Auto-Encoder (Unsupervised) - ImageNet**
  - 10M Images 200x200 pixels / 20 Classes
  - Cluster (1k CPUs / 16k cores)

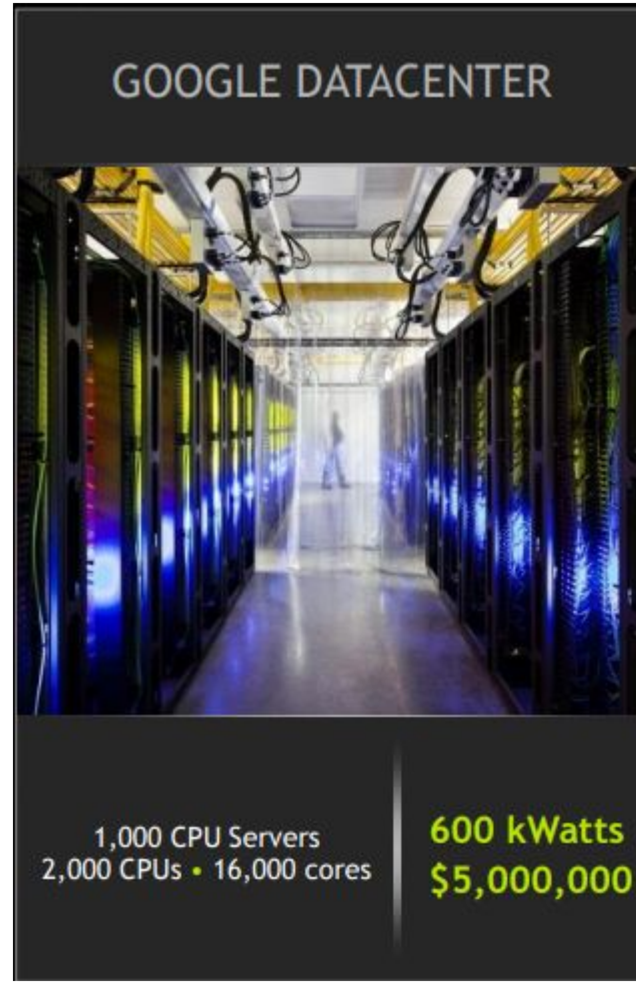


- Q. V. Le, M. Ranzato, R. Monga, M. Devin, K. Chen, G. S. Corrado, J. Dean, A. Ng. Building High-Level Features using Large Scale Unsupervised Learning in ICML 2012 (citado por 1610 jul/2018)

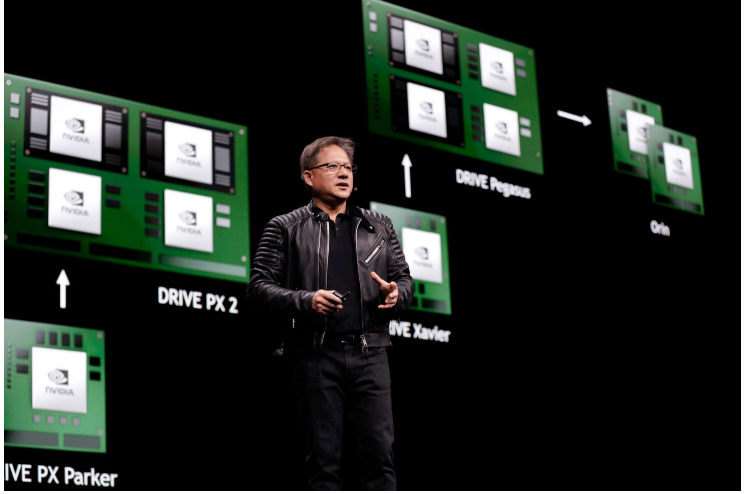
Eventos

- **2010 - Google Brain**
  - 1,000 CPUs (16k cores)
  - 600 kWatts
  - U$ 5.000.000
  - Prédio / Refrigeração

  - *Self-driving cars*
  - *Voice Recognition*



Eventos



NVIDIA CEO - Jensen Huang

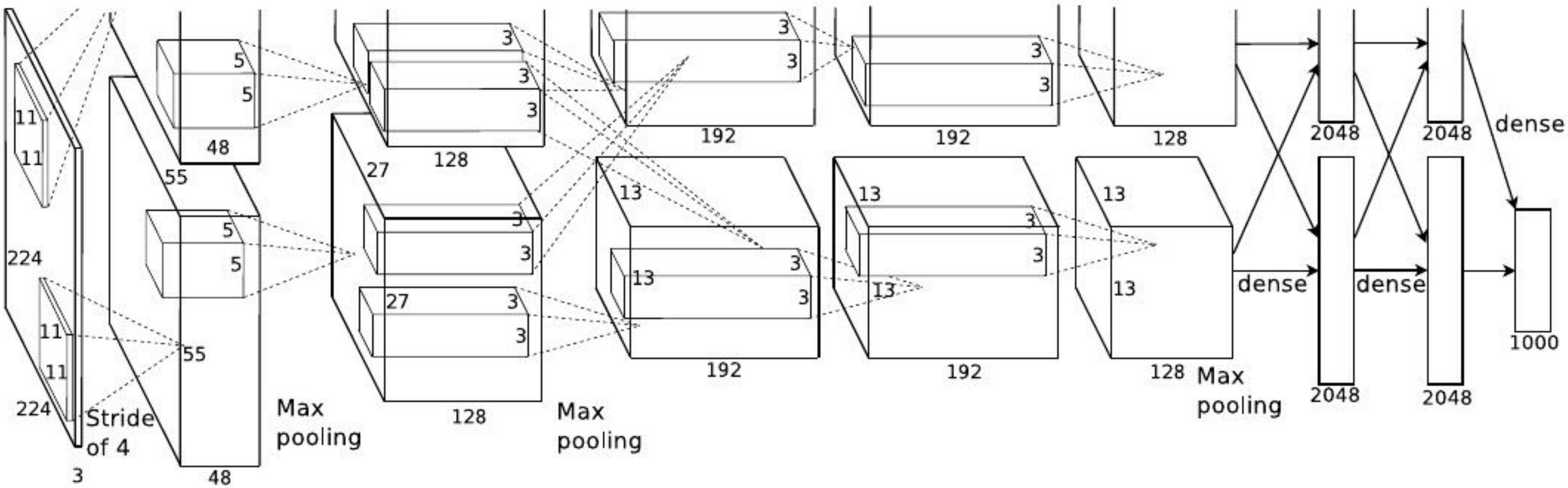
- **2013 - NVIDIA**
  - GTC (2014)
  - 3 Servidor (12 GPUs)
    - 18.432 cores
  - 4 KWatts
  - U$ 33.000

- GPUs 20-50x +CPUs
  - 2010?

- A. Coates, B. Huval, T. Wang, D. Wu, A. Ng, B. Catanzaro (421 citações - jul/2018) Deep learning with COTS HPC systems - Int. Conf Machine Learning (ICML) 2013

Eventos

- **AlexNet (ImageNet) /** Cuda-ConvNet



  - 60M parametrôs / 650K neurônios
  - 3 GPUs - limitado tamanho modelo pela RAM das GPUs
  - ImageNet (15M images / 22k classes) - erro caiu para 15%

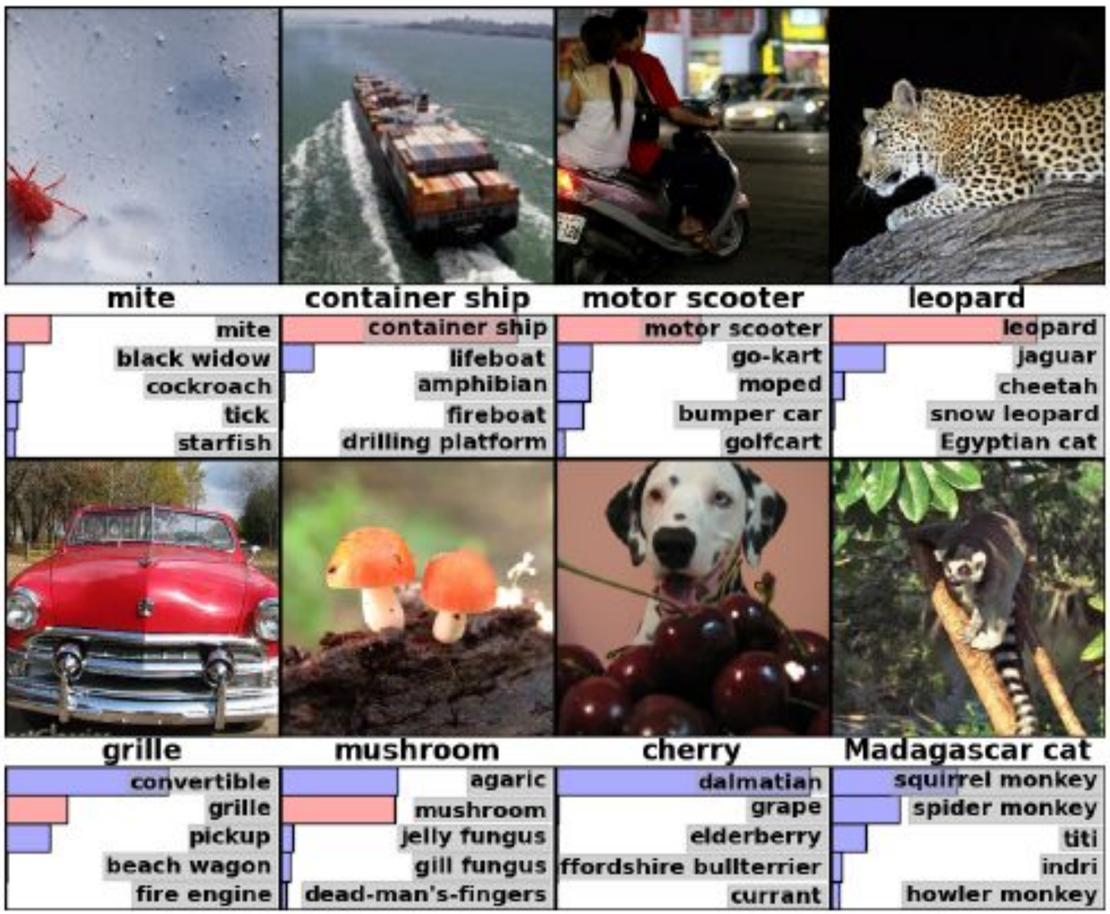
    - A. Kryzhevsky, I. Sutskever, G.E. Hinton. ImageNet Classification with Deep Convolutional Neural Networks. NIPS, 2012. (+26 mil citações jul/2018)

AlexNet (ImageNet)

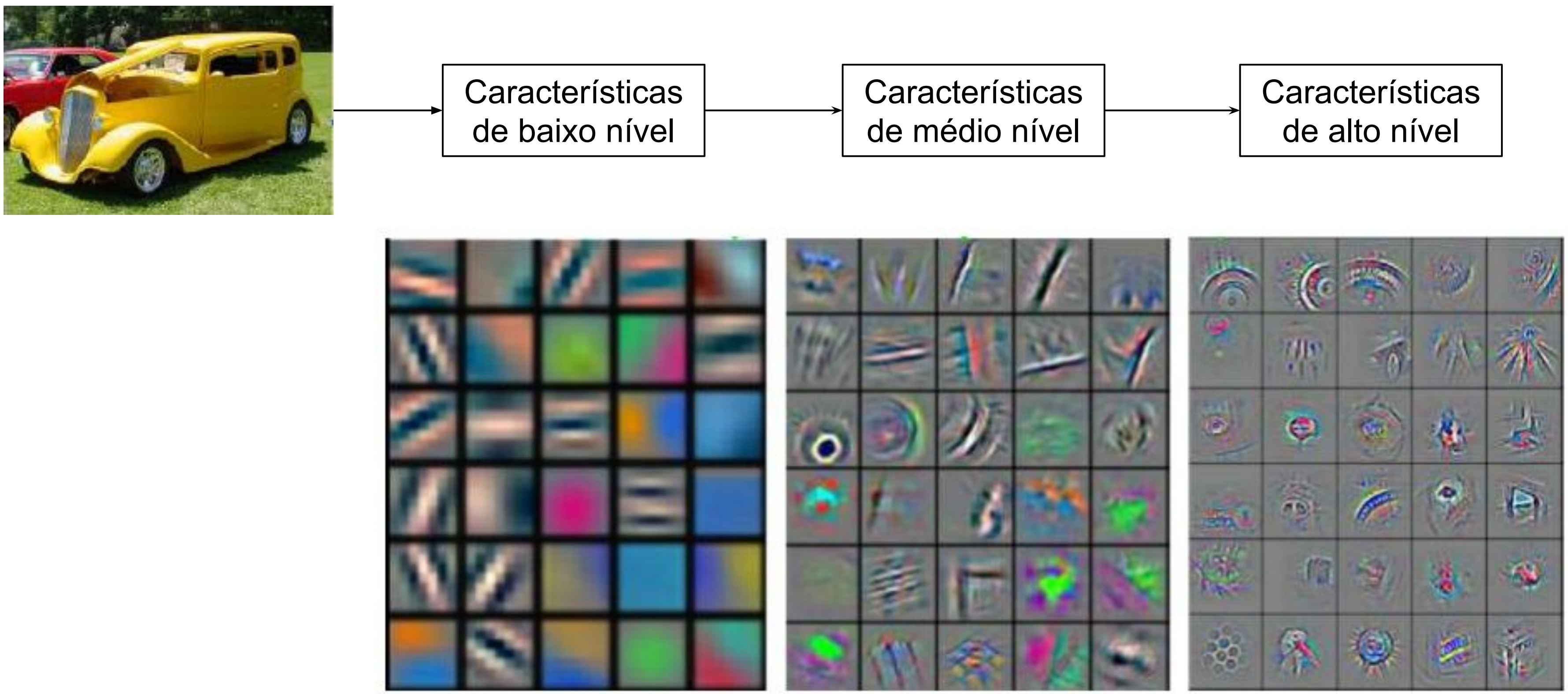
- **1 *Test* / 6 *Train* (menor distância)**

AlexNet (ImageNet)

- **Classes *Top 5***

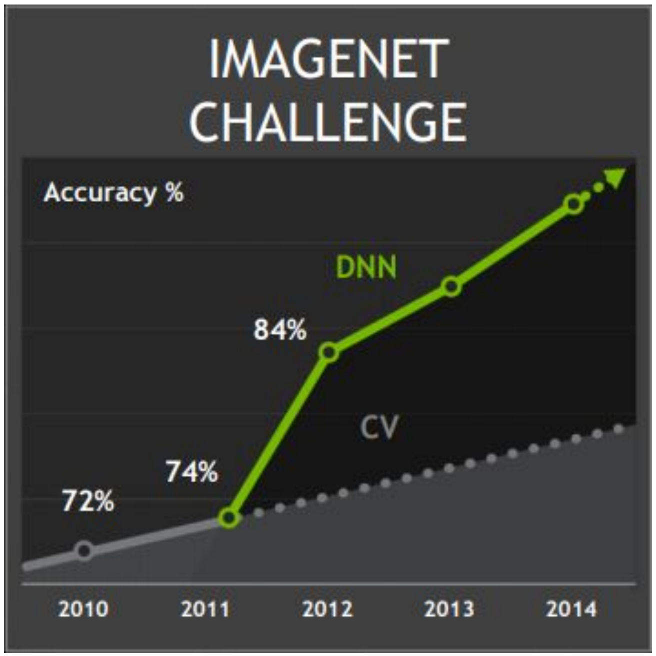


Hierarquia das Representações Treinadas



- Zeiler & Fergus. *Visualizing and Understanding Convolutional Networks*. European Conference on Computer Vision, 2014

ImageNet



"Deep Image: Scaling up Image Recognition"

– Baidu: 5.98%, Jan. 13, 2015

"Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification"

– Microsoft: 4.94%, Feb. 6, 2015

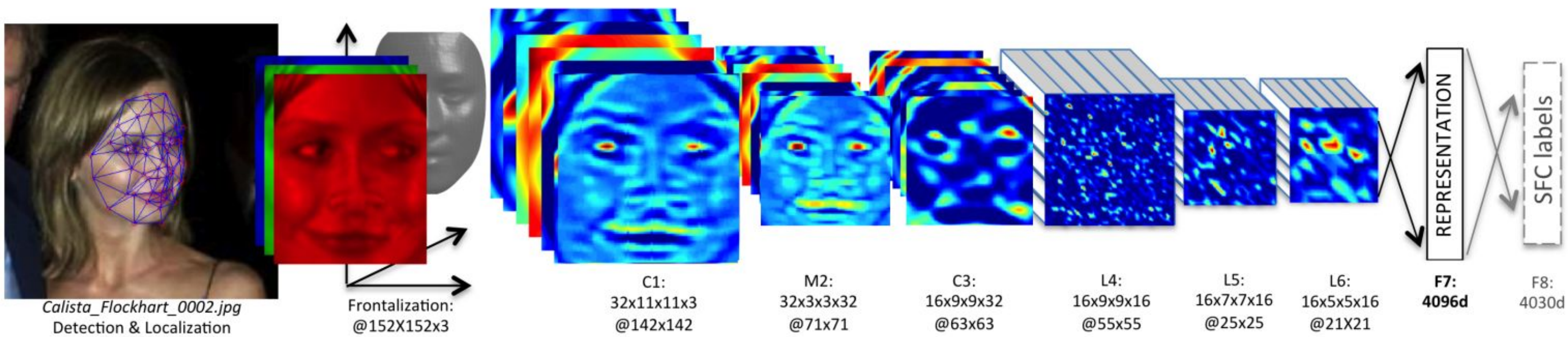
"Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariant Shift"

– Google: 4.82%, Feb. 11, 2015

- Larry Brown (Ph.D), Deep Learning with GPUs,
  GEOINT 2015 - http://www.nvidia.com/content/events/geoInt2015/LBrown_DL.pdf

Eventos

- **2014 - Deep Face (Facebook) / LFW database**



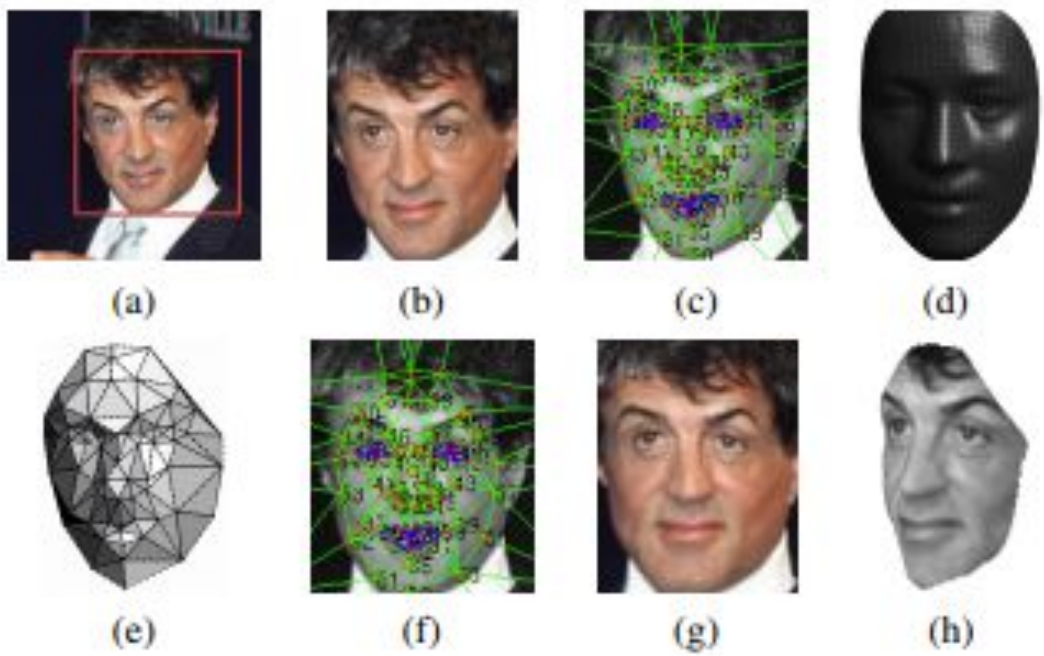
  – Treinada com 4M imagens faciais

Modelo	# de parâmetros	Acurácia (%)
Deep Face	128 M	97,35
State-of-the-art [2013]	~	79,70
Nível Humano	N / A	97,50

- Y. Taigman, M. Yang, M.A. Ranzato, L. Wolf. *DeepFace: Closing the Gap to Human-Level Performance in Face Verification*. CVPR 2014.

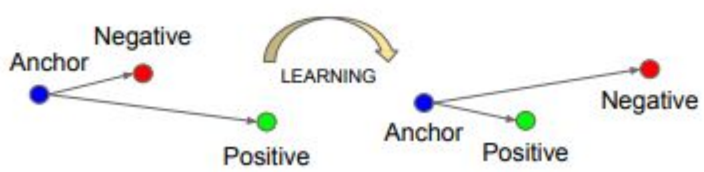
Eventos

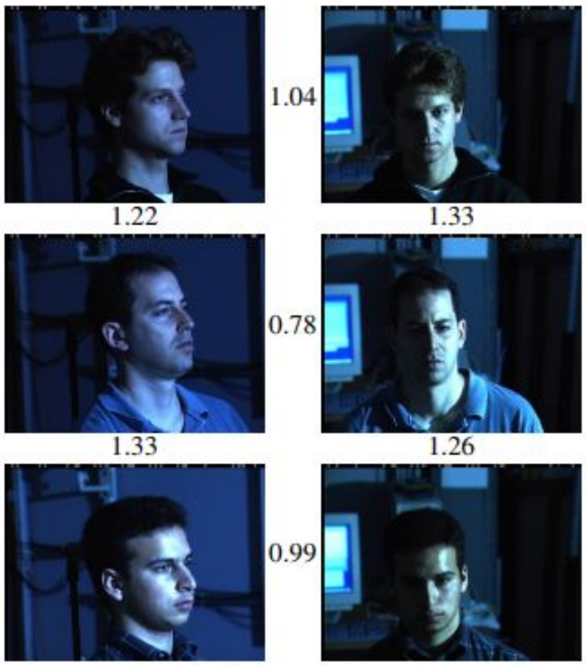
- **2014 - Deep Face (Facebook)**
  - **Frontalization**



Eventos

- **2015 - FaceNet (Google-DeepMind2014)**
  – Acurácia em 99,63% (LFW)
  – Representação Compacta
    - (128 bytes por Face)
  – Usaram 100M-200M
    - De 8M de entidades





Invariância a iluminação e pose

- F. Schroff, D. Kalenichenko, J. Philbin. FaceNet: A Unified Embedding for Face Recognition and Clustering. CVPR 2015.

2015 - FaceNet - Erros

- F. Schroff, D. Kalenichenko, J. Philbin. FaceNet: A Unified Embedding for Face Recognition and Clustering. CVPR 2015.

2015 - FaceNet - Erros

- F. Schroff, D. Kalenichenko, J. Philbin. FaceNet: A Unified Embedding for Face Recognition and Clustering. CVPR 2015.

Eventos

- **2010/2013 - TIMIT Phoneme Recognition**

Modelo	# de parâmetros	Erro (%)
Hidden Markov Model (HMM)	N/A	27,3
Deep Belif Network DBN	~4M	26,7
Deep RNN	4,3M	17,7

– Dados:
  - Pessoas (#): 462 *train* / 24 *test*
  - Tempo (h): 3,16 *train* / 0,14 *test*

- Mohamed, A. and Hinton, G. E. (2010). Phone recognition using restricted Boltzmann machines // ICASSP pages 4354–4357.
- Graves, A., Mohamed, A.-R., and Hinton, G. E. (2013). Speech recognition with deep recurrent neural networks ICASSP, pages 6645–6649.

Eventos

- **2014 - Google Large Vocabulary Speech Recognition**
  - Data de treinamento: 3M expressões (1900 hrs)

Modelo	# de parâmetros	Cross-entropy (%)
ReLU DNN	128 M85M	11,3
Deep Projection LSTM RNN	13M	10,7

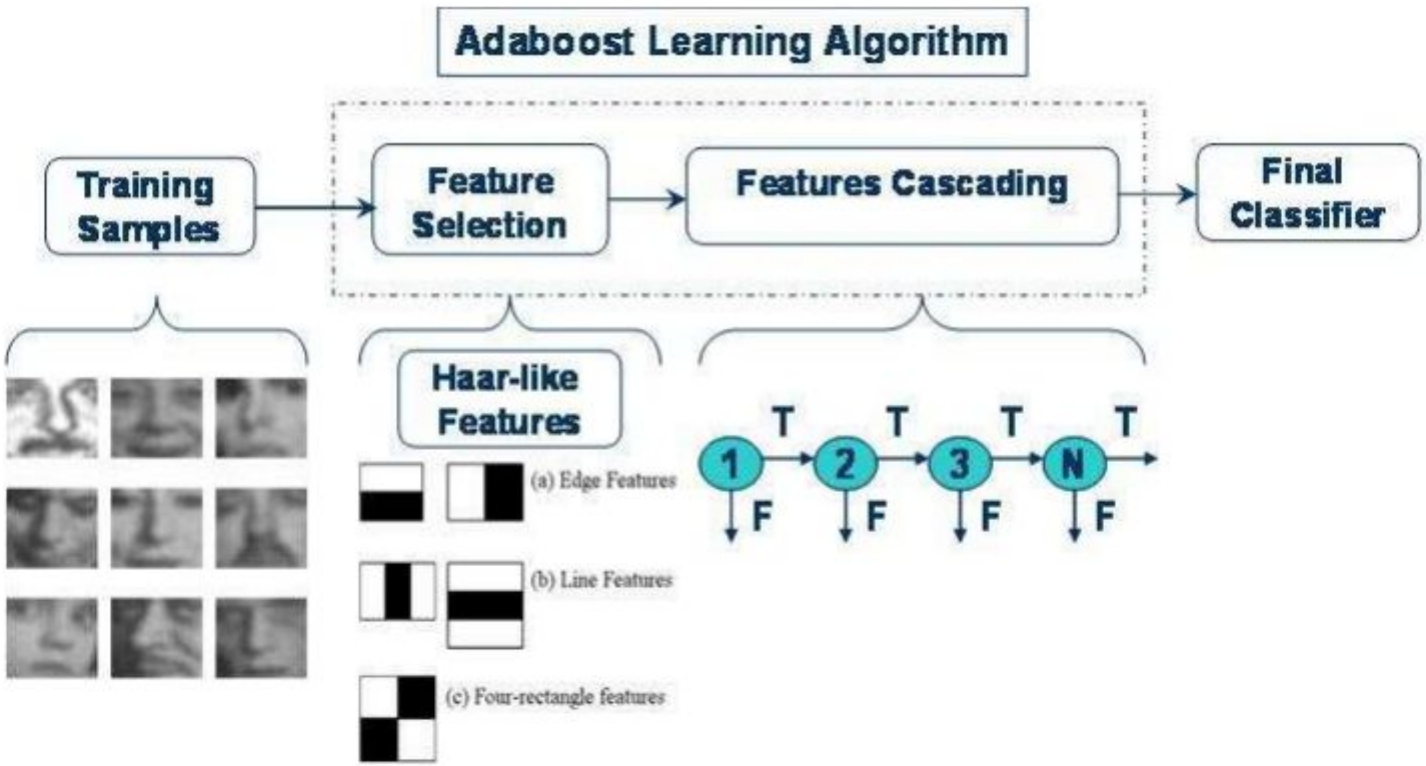
- H. Sak, A. Senior, F. Beaufays. Long Short-Term Memory Recurrent Neural Network Architectures for Large Scale Acoustic Modeling // INTERSPEECH'2014.
- K. Vesely, A. Ghoshal, L. Burget, D. Povey. Sequence-discriminative training of deep neural networks // INTERSPEECH'2014.

Inteligência Artificial

- **Processamento de Linguagem Natural (2010)**
  - Siri (Apple)
  - The Assistant (Google)
  - Microsoft (Cortana)
  - Alexa (Amazon

Eventos

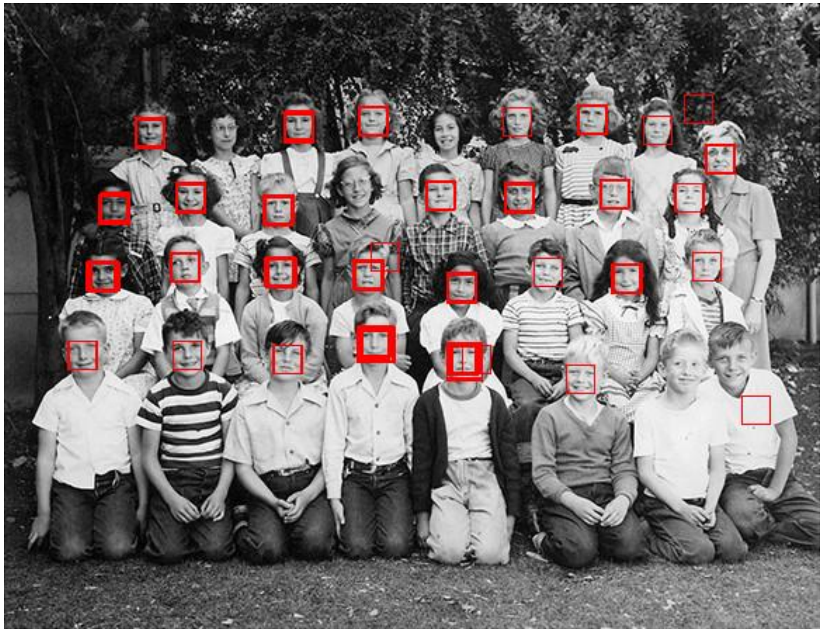
- **2001 - Face Detection - Viola & Jones**



- P. Vila & Jones, Rapid object detection using a boosted cascade of simple features IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2001

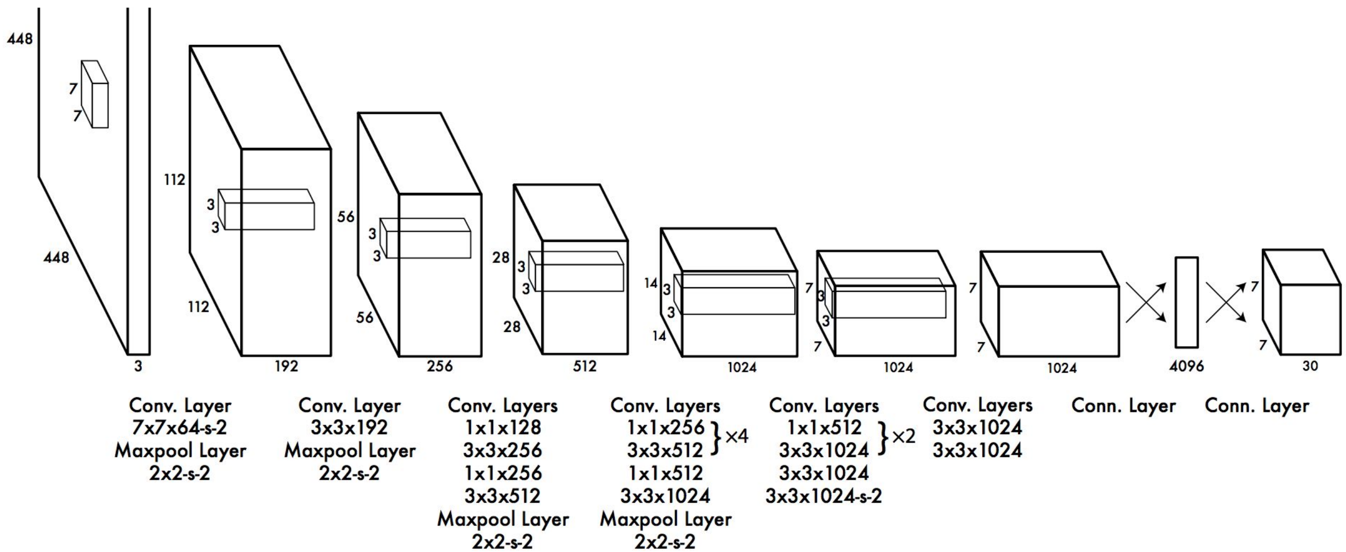
Eventos

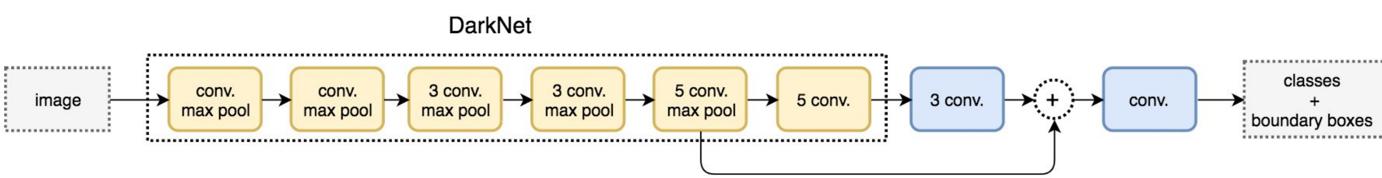
- **2001 - Object Detection - Viola & Jones**
  - Semanas para treinar
    - Características simples
  - Imagens de 384x288 pixels
  - 15 fps 700 Mhz
    - Pentium III
  - ~94 de acurácia
  - Máquinas Fotogr.



Eventos

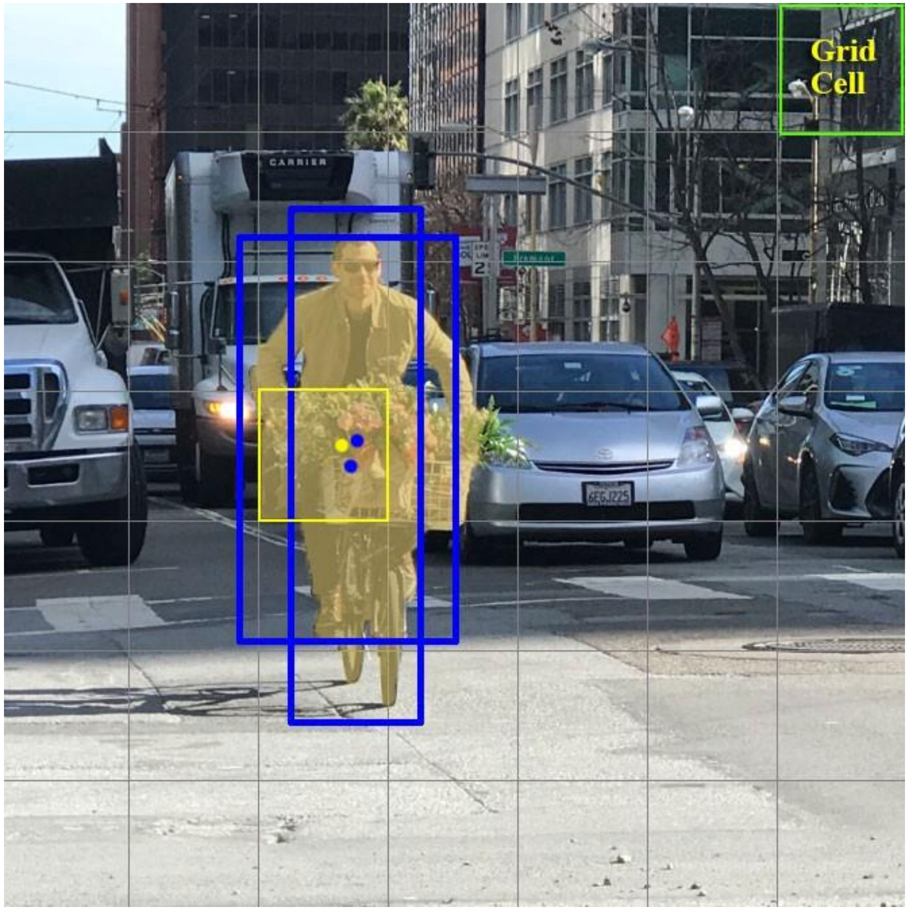
- **2015-2017 Yolo (Object Detection)**





Eventos

- **2015-2017 Yolo (Object Detection)**



Object Detection

- **Demonstração YoloV2 & YoloV3**
  - **Vídeos ~ 2 minutos**

Eventos

- **2015-2018 Identificação de Veículos**
  - Câmera e Veículo em movimento

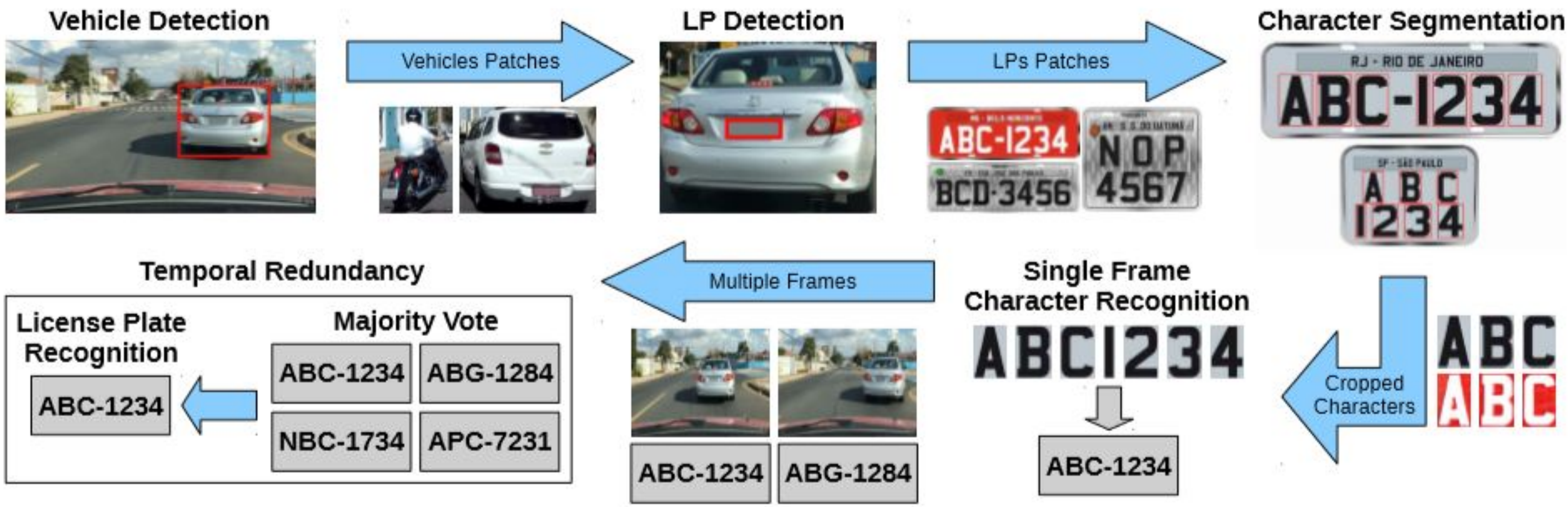
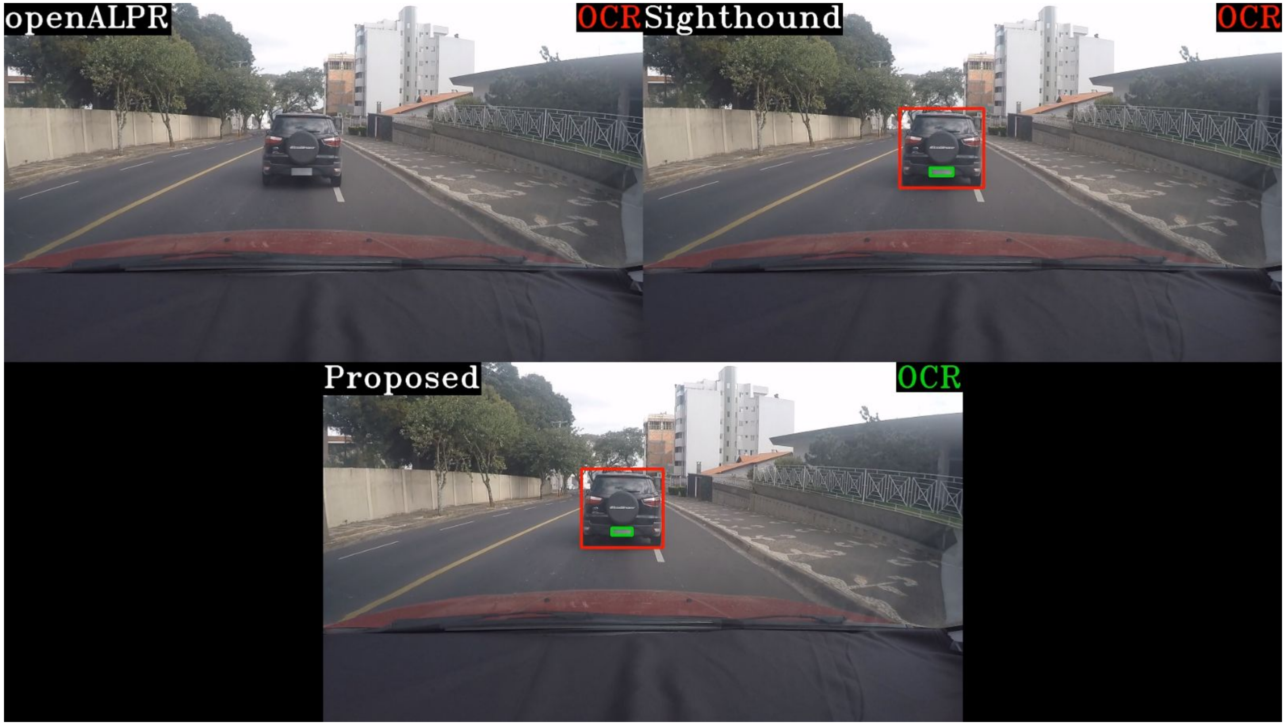


Fig. 5. An usual ALPR pipeline having temporal redundancy at the end.

- R. Laroca, others, D. Menotti, A robust real-time automatic license plate recognition based on the YOLO detector. IEEE International Joint Conference on Neural Networks (IJCNN) 2018.

Eventos

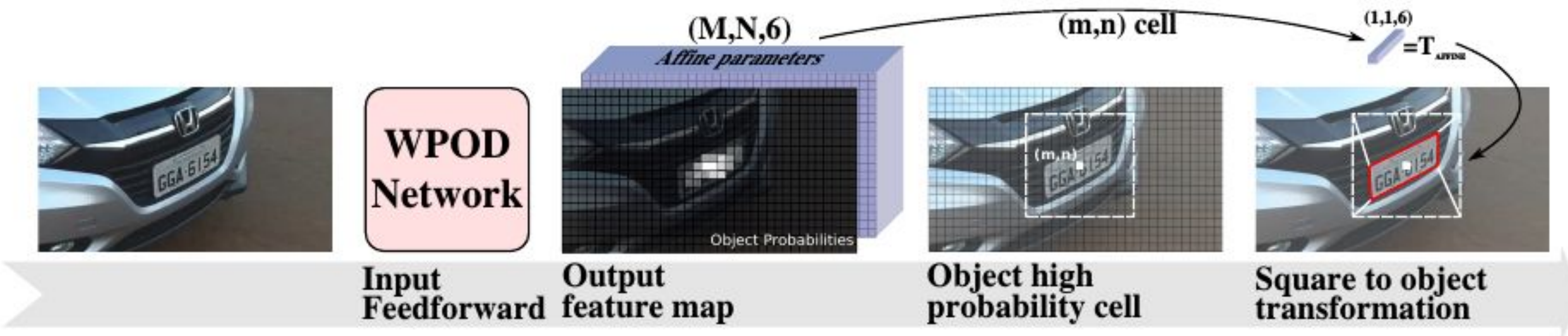
- **2015-2018 Identificação de Veículos**
  - Vídeo

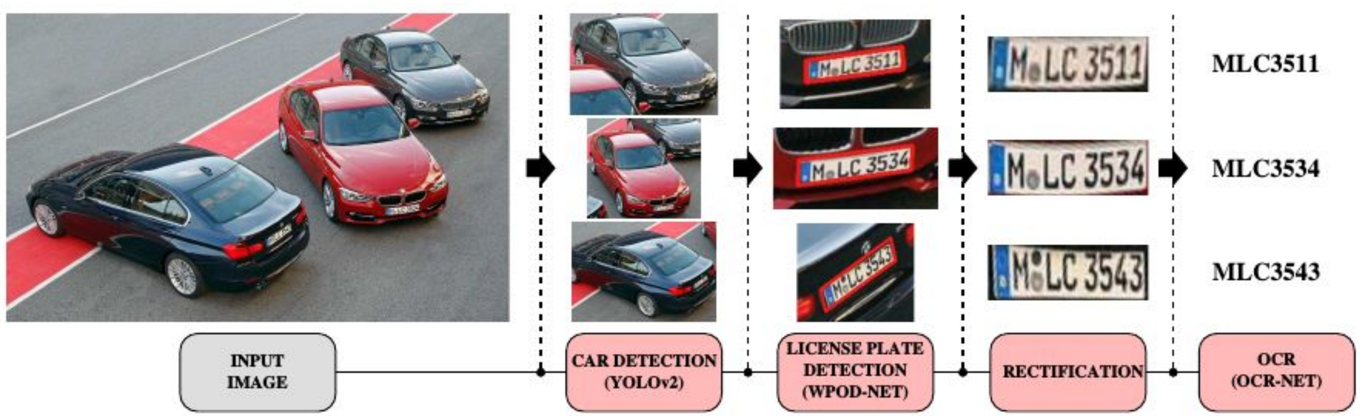


- R. Laroca, others, D. Menotti, A robust real-time automatic license plate recognition based on the YOLO detector. IEEE International Joint Conference on Neural Networks (IJCNN) 2018.

Eventos

- **2018** Retificação da Placa





- S. Montazolli & C. R. Jung, License Plate Detection and Recognition in Unconstrained Scenarios. European Conference on Computer Vision (ECCV) 2018.

Eventos

- **2018 Identificação de Veículos**
  - Vídeo
    - http://www.inf.ufrgs.br/~crjung/alpr-datasets/

- S. Montazolli & C. R. Jung, A robust real-time automatic license plate recognition based on the YOLO detector. IEEE International Joint Conference on Neural Networks (IJCNN) 2018.

Eventos

- **2016 - AlphaGo (Image)**
  - Usa conhecimento humano & experiência
  - DL vence por 4 a 1 **Lee Sedol** (echo de **Garry Kasparov**)
  - Jogo muito mais complexo que xadrez
  - Google usou 1920 CPUs & 280 GPUs

Eventos

- **2017 - AlphaGo Zero (*Reinforcement learning*)**
  - Conhece regras do jogo / sem usar dados de treinamento
  - Aprende jogando contra si mesmo,
  - Venceu AlphaGo Lee (3 dias), competitivo com AlphaGo Master (21 dias) e superou todos com 40 dias
  - 19 servidores CPU e 64 estações GPU

  - Melhor que ser humano sem conhecimento
    - **Implicações?**

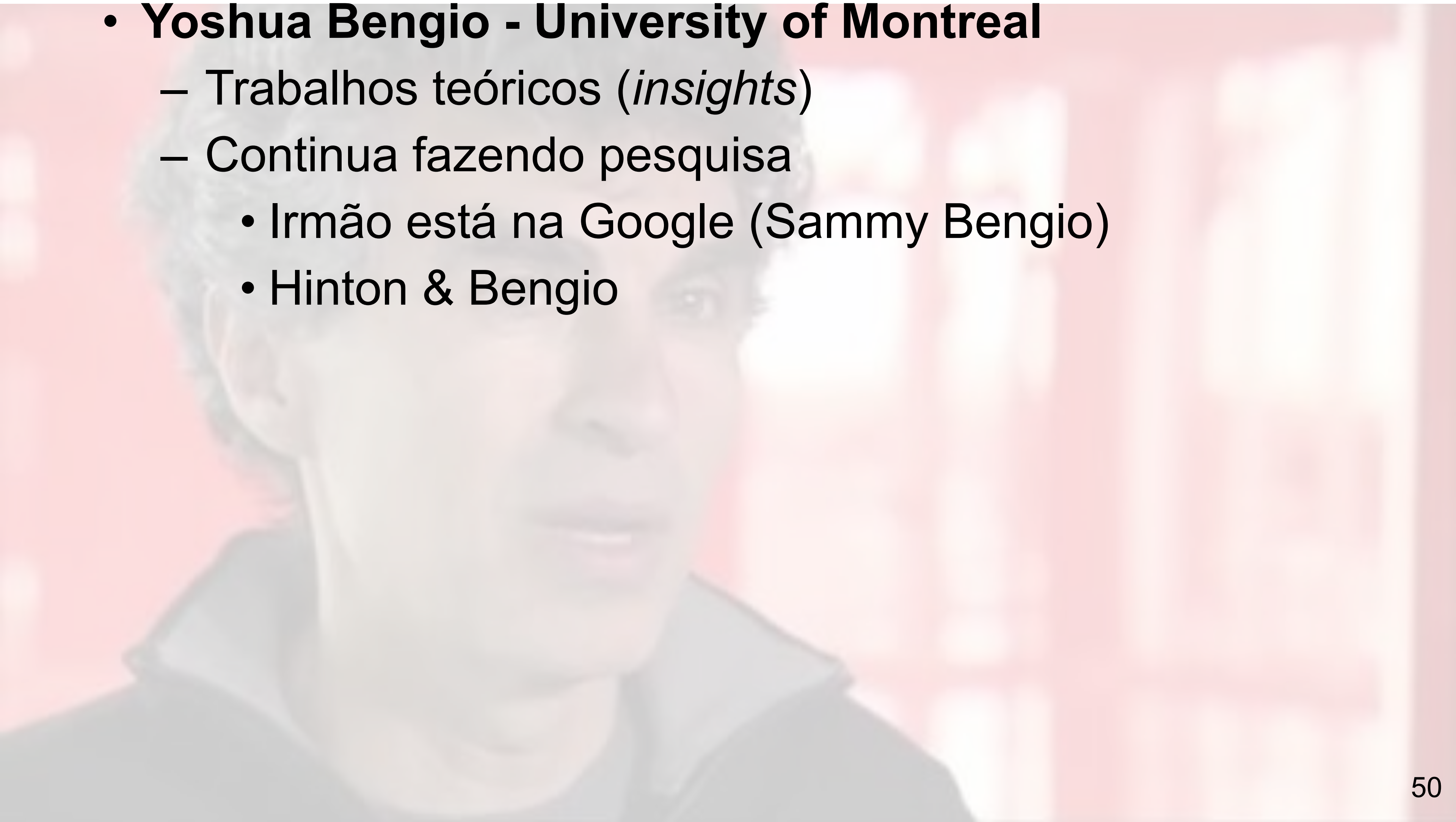
- Silver, David; Schrittwieser, Julian; Simonyan, Karen; Antonoglou, Ioannis; Huang, Aja; Guez, Arthur; Hubert, Thomas; Baker, Lucas; Lai, Matthew; Bolton, Adrian; Chen, Yutian; Lillicrap, Timothy; Fan, Hui; Sifre, Laurent; Driessche, George van den; Graepel, Thore; Hassabis, Demis (19 October 2017). "Mastering the game of Go without human knowledge". *Nature*. **550** (7676): 354–359

Bigshots - Machine Learning

- **Geoffrey Hinton (University of Montreal)**
  - Nature em 1986
    - Backpropagation
  - Science em 2006
    - Stacker Auto-encoder
      - redução de dimensionalidade
  - Explorou ImageNet 2012
    - Lecun & Bengio

Bigshots - Machine Learning

- **Yoshua Bengio - University of Montreal**
  - Trabalhos teóricos (*insights*)
  - Continua fazendo pesquisa
    - Irmão está na Google (Sammy Bengio)
    - Hinton & Bengio



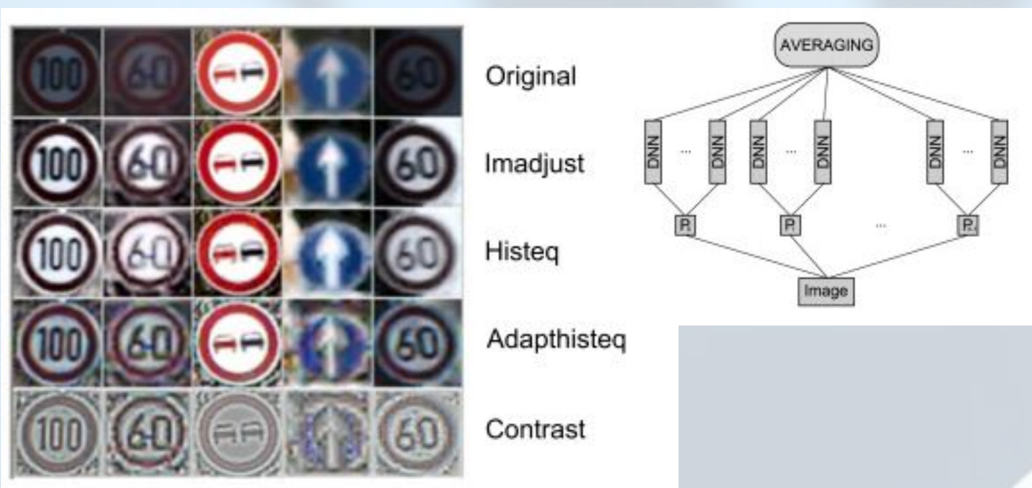
Bigshots - Machine Learning

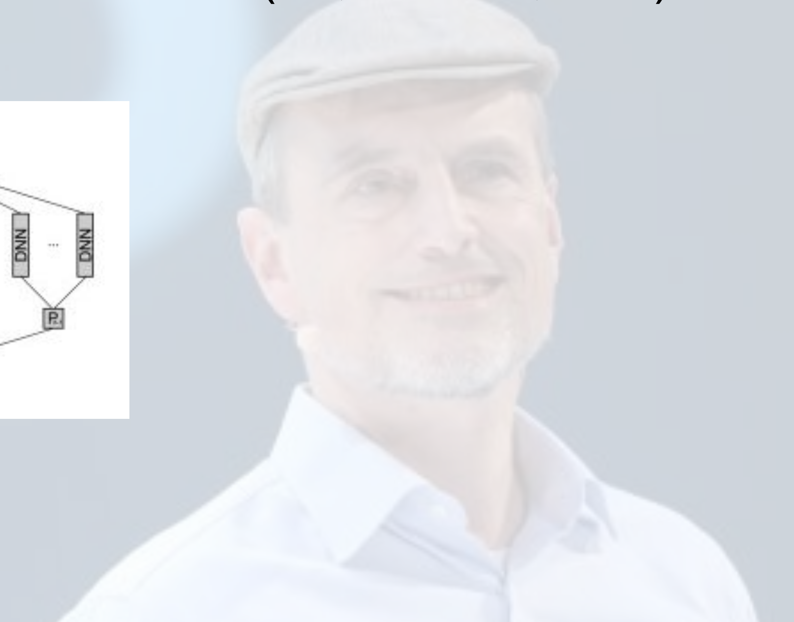
- **Yann Lecun - Facebbok / NYU**
  - “Pai” da Rede de Convolução
  - 1989 Trabalho de graduação (ESIEE-Paris)
  - 1998 - Convolução para Back-propagation
  - 1991 PostDoc (University of Toronto)
    - Hinton
  - 1996 AT&T Labs - Image Processing Head
  - 2003 NYU (*New York University*)
  - 2013 - Chief AI Scientist for Facebook AI Rs
    - Hinton & Bengio

Bigshots - Machine Learning

- **Jurgen Schmidhuber (Switzerland)**
  - (Re)Clama / Questiona autorias
  - 2012 - Reconhecedor de Placas de Trânsito

    Acurácia superior a de seres humanos (99,36%/98,48%)

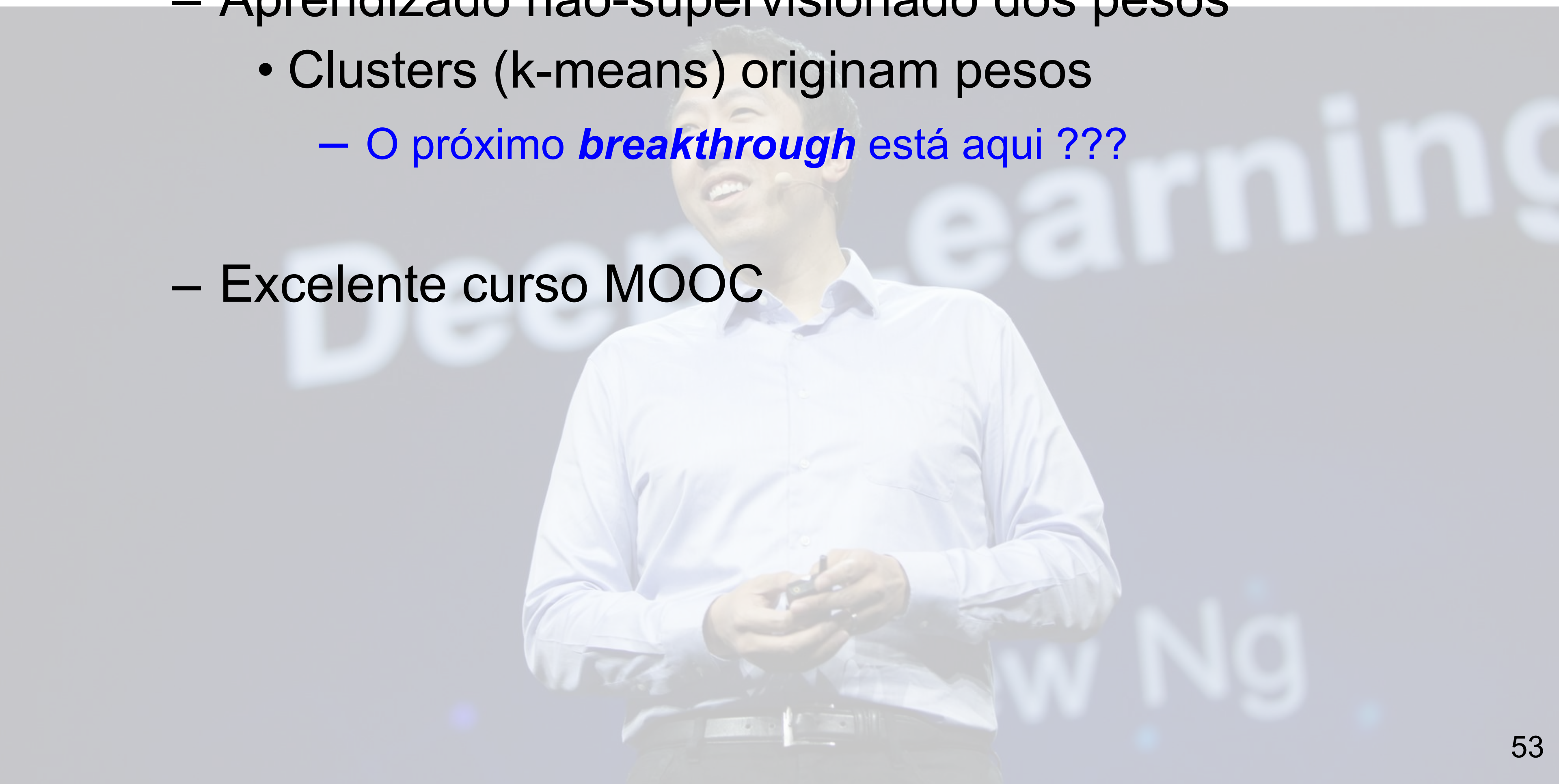




- D. Cireşan, U. Meier, J. Masci, Jürgen Schmidhuber. Multi-column deep neural network for traffic sign classification, Neural Networks, 2012 (citado por 468 jul/2018)

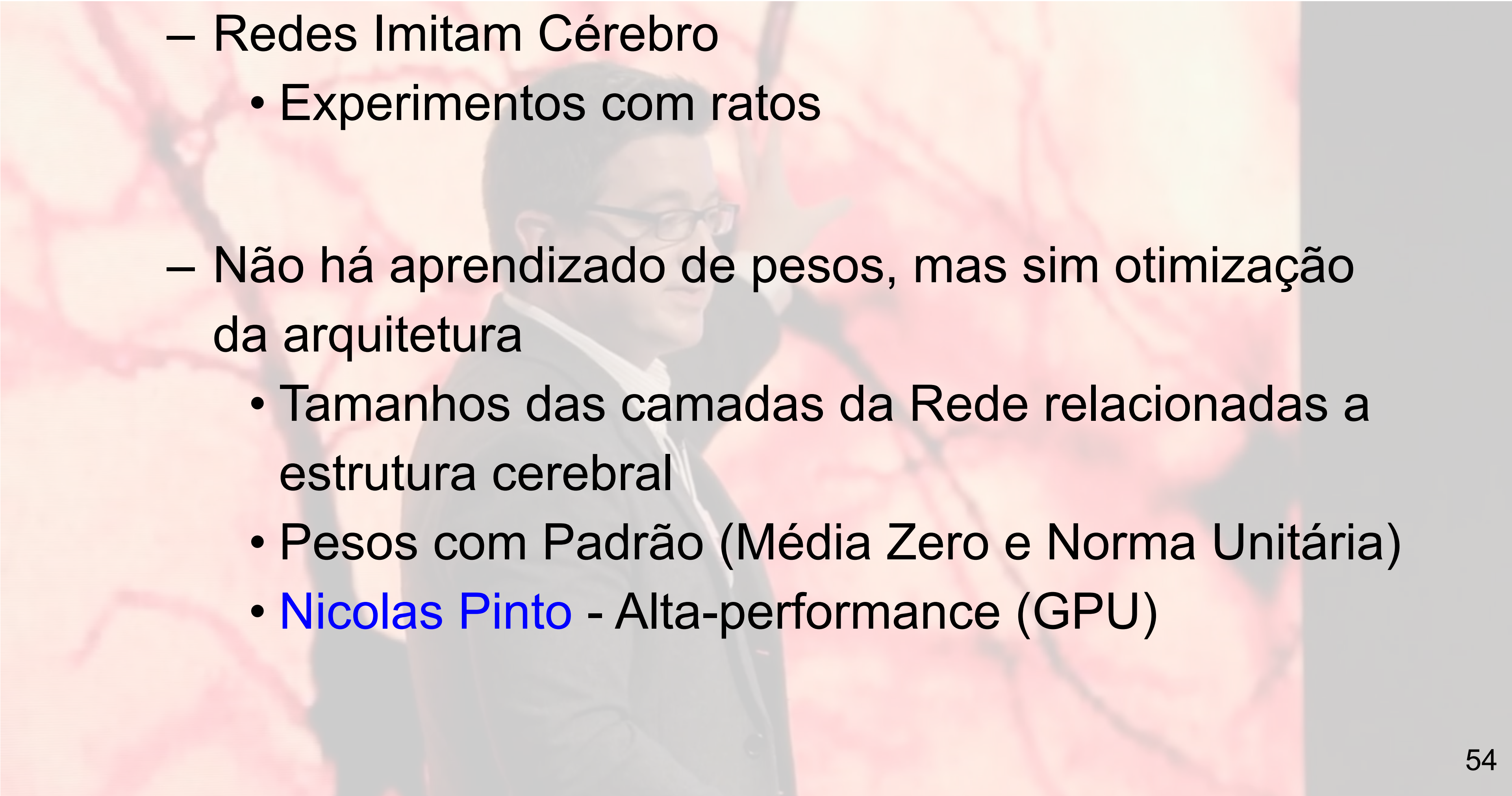
Bigshots - Machine Learning

- **Andrew NG - (Stanford University)**
  - Aprendizado não-supervisionado dos pesos
    - Clusters (k-means) originam pesos
      - O próximo ***breakthrough*** está aqui ???
  - Excelente curso MOOC



Bigshots - Machine Learning

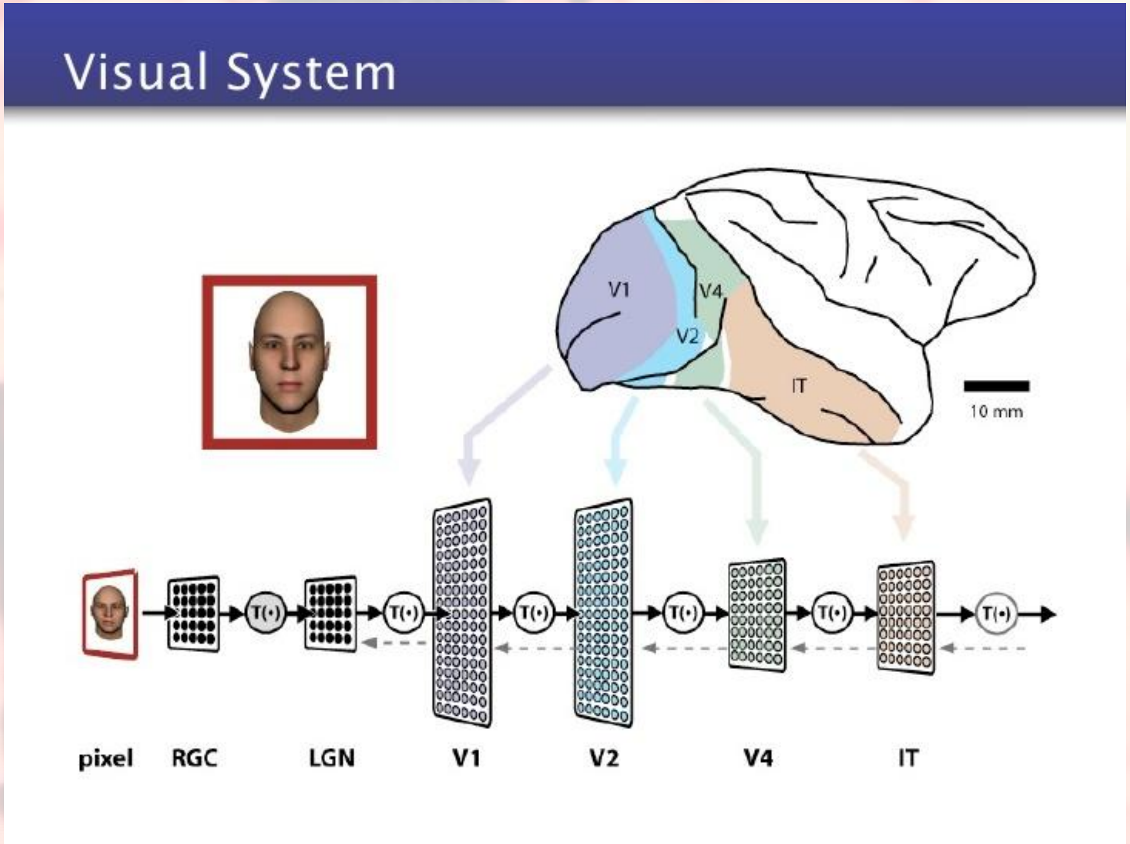
- **David Cox - Harvard / MIT**



  - Redes Imitam Cérebro
    - Experimentos com ratos
  - Não há aprendizado de pesos, mas sim otimização da arquitetura
    - Tamanhos das camadas da Rede relacionadas a estrutura cerebral
    - Pesos com Padrão (Média Zero e Norma Unitária)
    - Nicolas Pinto - Alta-performance (GPU)

Bigshots - Machine Learning

- **David Cox - Harvard / MIT**
  - Cérebro & Redes de Convolução




Questions

Referências

- James O'Malley **The 10 most important breakthroughs in Artificial Intelligence** https://www.techradar.com/news/the-10-most-important-breakthroughs-in-artificial-intelligence
- Seth Weidman **The 4 Deep Learning Breakthroughs You Should Know About** https://towardsdatascience.com/the-5-deep-learning-breakthroughs-you-should-know-about-df27674ccdf2
- Artem Chernodub, Geroge Pashchenko, **Details of Lazy Deep Learning for Images Recognition in ZZ Photo app**, Kharkob AI Club, 2015
- Alex Oagana **A Short History of Mercedes-Benz Autonomous Driving Technology** **https://www.autoevolution.com/news/a-short-history-of-mercedes-benz-autonomous-driving-technology-68148.html**