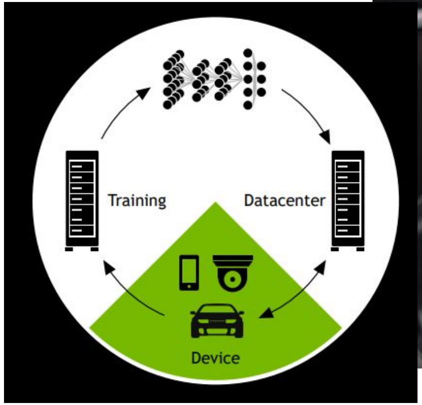
Universidade Federal do Paraná (UFPR)
Especialização em Engenharia Industrial 4.0

AI / ML / Deep Learning



David Menotti

www.inf.ufpr.br/menotti/am-212

www.inf.ufpr.br/menotti/am-202

Objetivos

- Inteligência Artificial
  - Consciência
- Aprendizado de Máquinas
- *Deep Learning - Breakthroughs*
  - Histórico e eventos
- Minhas Pesquisas
- *Bigshots*

Inteligência Artificial

- Questões filosóficas:
  - Uma máquina pode realmente pensar como um humano?
  - Uma máquina pode ser um humano?

- 1637 - **René Descartes** - *“Je pense, donc je suis!”*
  - ***Discours de la méthode: Pour bien conduire sa raison, et chercher la vérité dans les sciences***

    “Se houvesse máquinas que tivessem uma semelhança com nossos corpos e imitassem nossas ações o mais próximo possível para todos os propósitos práticos, nós ainda deveríamos ter dois meios muito certos de reconhecer que eles não eram homens de verdade.”

    - Máquina não poderia “falar“ ou “juntar sinais”

Inteligência Artificial

- **1950 -** O teste de Turing / “*The Imitation Game”* (1950)

  “ Se um juiz não pode diferenciar entre um humano e uma máquina (interface texto), a máquina pode enganar o juiz a pensar que ele é o humano? “

  – Previsão em 50 anos - Computadores com 1GB de RAM
    - 70% de certeza / 5 minutos de prova

  – **Alan Turing** (Pai da Ciência da Computação)
    - Nasceu em 23.06.1912 / Morreu 07.02.1954
    - *The Imitation Game* (2014)

Inteligência Artificial

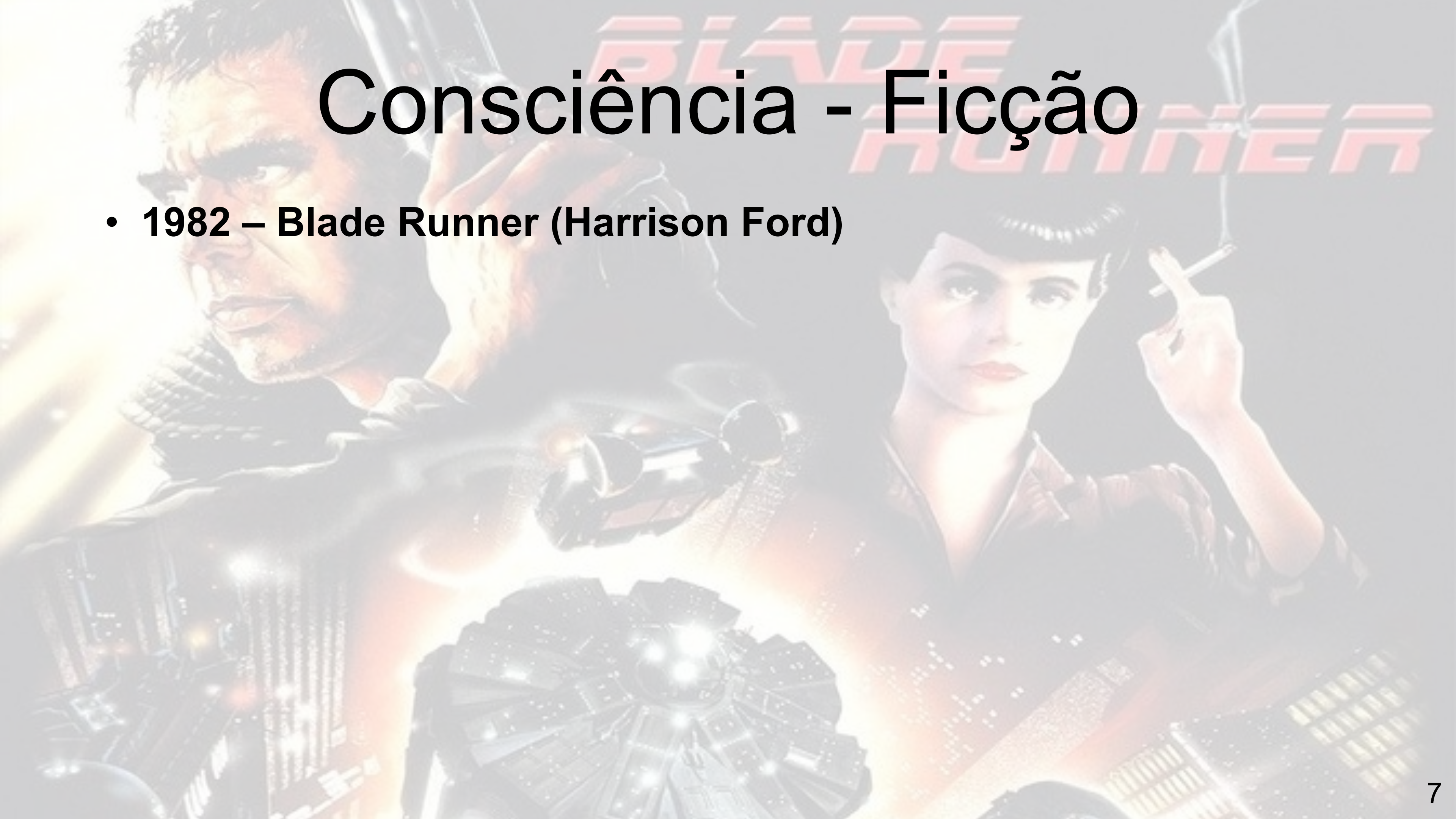
- Computadores / Máquinas tem consciência?

Consciência - Ficção

- **1968 – *2001: A Space Odyssey* (Stanley Kubrick)**
  - **Hal**
  - **Monolito**

Consciência - Ficção

- **1982 – Blade Runner (Harrison Ford)**



Consciência - Ficção

- **2017 – Blade Runner (Ryan Gosling / Harrison Ford)**

Consciência - Ficção

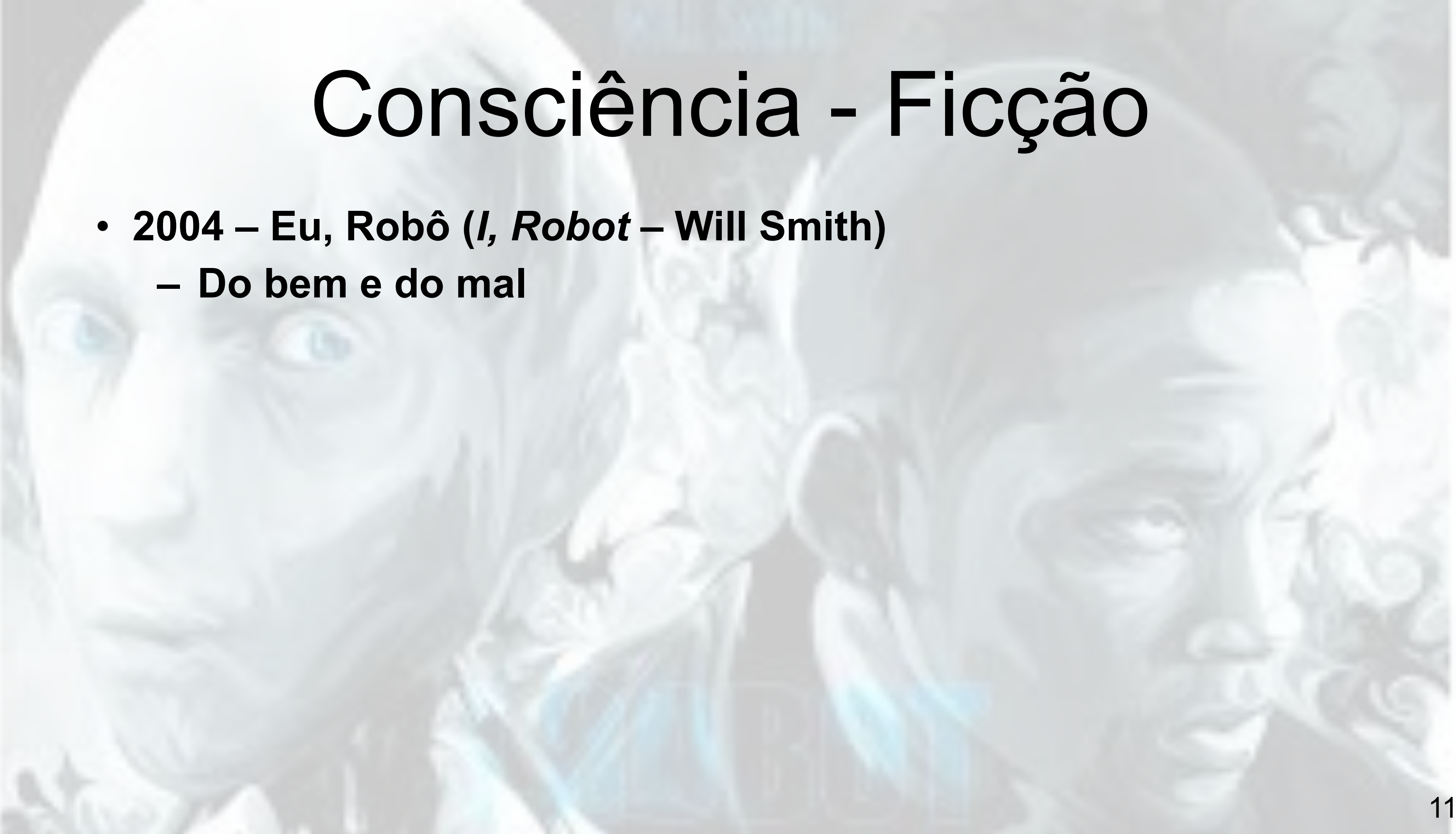
- **1984 – *The Terminator* (Arnold Schwarzenegger)**
  - **Sara Connor**
- **1991 – *The Terminator 2* (Arnold Schwarzenegger)**
  - **Skynet - 29 de Agosto de 1997**
- **???**

Consciência - Ficção

- **1999 – *Bicentennial Man* (Robin Williams)**
  - **Quis se tornar uma máquina**

Consciência - Ficção

- **2004 – Eu, Robô (*I, Robot* – Will Smith)**
  - **Do bem e do mal**



Consciência - Ficção

- **2008 – Controle Absoluto (Eagle Eye – Michelle Monaghan)**
  - **Conspiração vs estado**

Consciência - Ficção

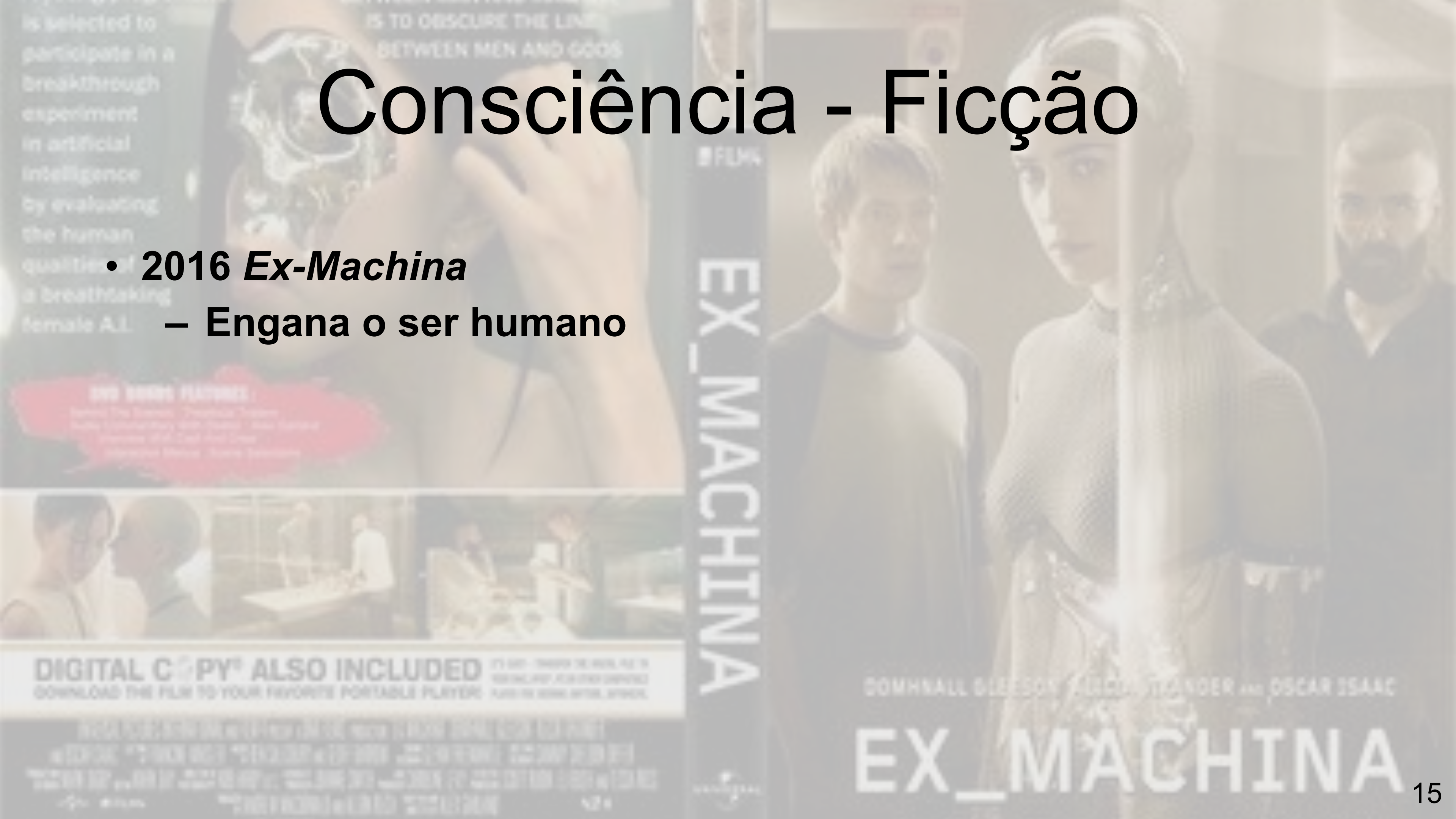
- **2011 - *Artificial Intelligence* (Jude Law)**
  - **Não sabe que é máquina**

Consciência - Ficção

- **2014 - *Transcendence* (Johnny Depp)**
  - **Consciência transplantada**

Consciência - Ficção

- **2016 *Ex-Machina***
  - **Engana o ser humano**



Consciência - Ficção

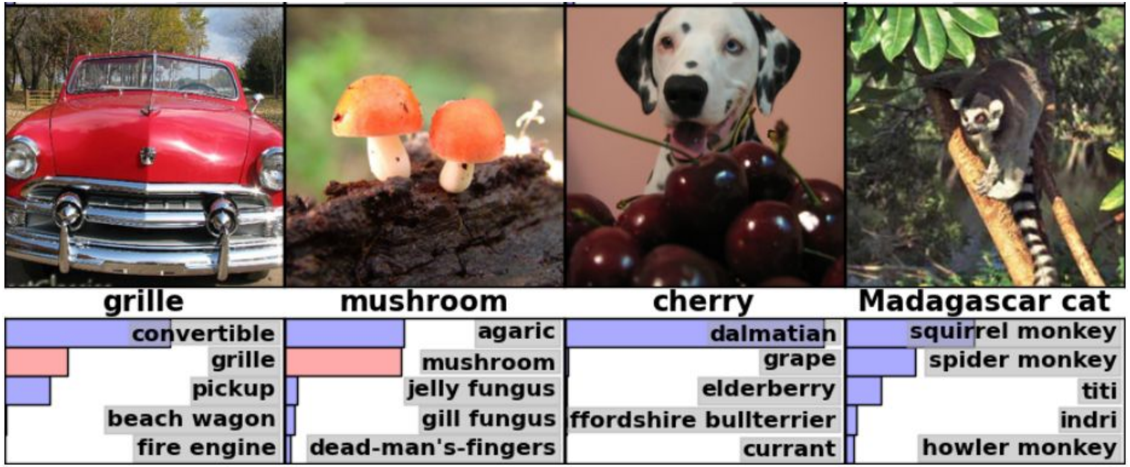
- **2017 *Westworld***
  - **Humano vivendo no mundo artificial**

Consciência - Realidade

- **As máquinas têm consciência?**

Aprendizado de Máquinas

- Mas as máquinas já “aprendem” !!!



Inteligência Artificial

- **1995 Veículos Autônomos (*Self-driven Car*)**
  - 1980s Vario Mercedes-Benz
  - 1995 S-Class W140 modificado - Mercedes-Benz (1986 → 2021)
    - trecho : Munique até Copenhagen (1678 Km) - **quase autonomo**
    - *Computer Vision* & Microprocessadores
    - 180 km/h, Ultrapassagem, Interpretava Placas de Trânsito
  - 2009 Waymo (Google)
    - Hoje (2017) empresa independente
      - Sensores & Câmeras
  - 2020 Uber & Waymo - U$ 100k
    - 19.03.2018 - Vítima Fatal (Arizona)

Inteligência Artificial

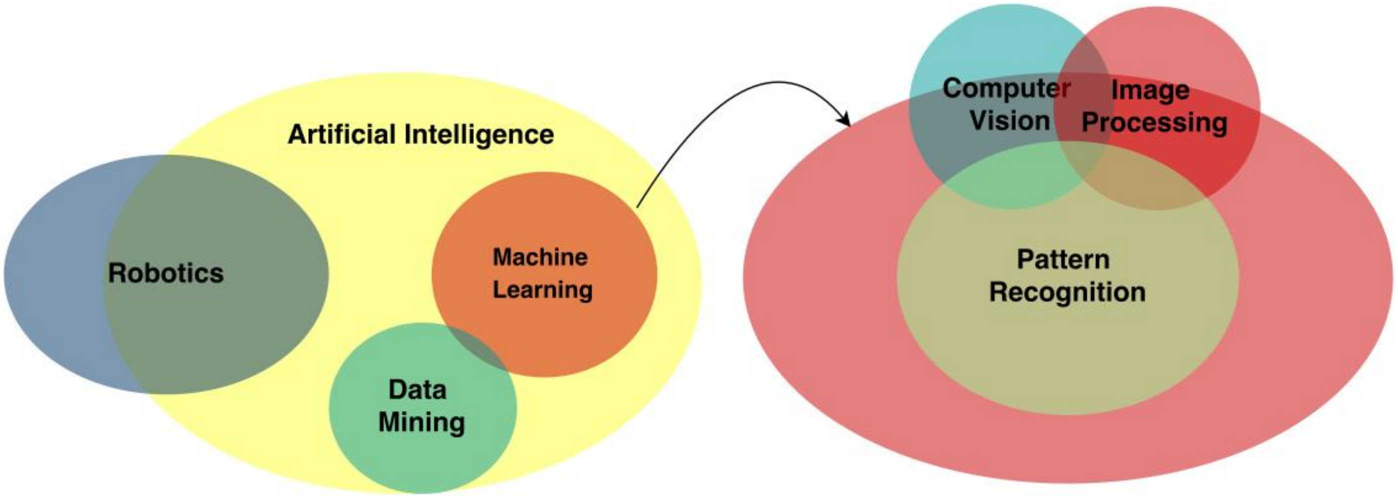
- **1996/1997 Garry Kasparovy vs *Deep Blue***

  – 1996 Kasparov vence *Deep Blue* (IBM) por 4-2
    - Ainda usando modelos baseados em regras (IA clássica)

  – 1997 *Deep Blue* 2 - 3 (empates) - 1 Kasparov
    - Inteligência Artificial ilusória
      – "Ele está jogando contra os fantasmas dos grandes mestres do passado"
        » IBM alimentou o sistema com dados de milhares de jogos anteriores
        » *Deep Blue* não “aprendia”, usava a experiência

Inteligência Artificial

- **Uma visão da “hierarquia” da área**



Aprendizado de Máquinas

- **Redes Neuronais Artificiais**
  - 1951 “SNARC”
    (*Stochastic Neural Analog Reinforcement Computer*)
    - Feita de tubos e válvulas
    - Ajudar um rato virtual a resolver um problema

  - 1960-1980 Pesquisa em Inteligência Artificial
    - Clássica (*ruled based*)

  - 1988 Estatística
    - IBM’s TJ Watson Research Center
      - *A statistical approach to language translation*

- Estatística (1988)

Como Aprende?

- Ser Humano:
  - Bons exemplos
  - Erros - correção

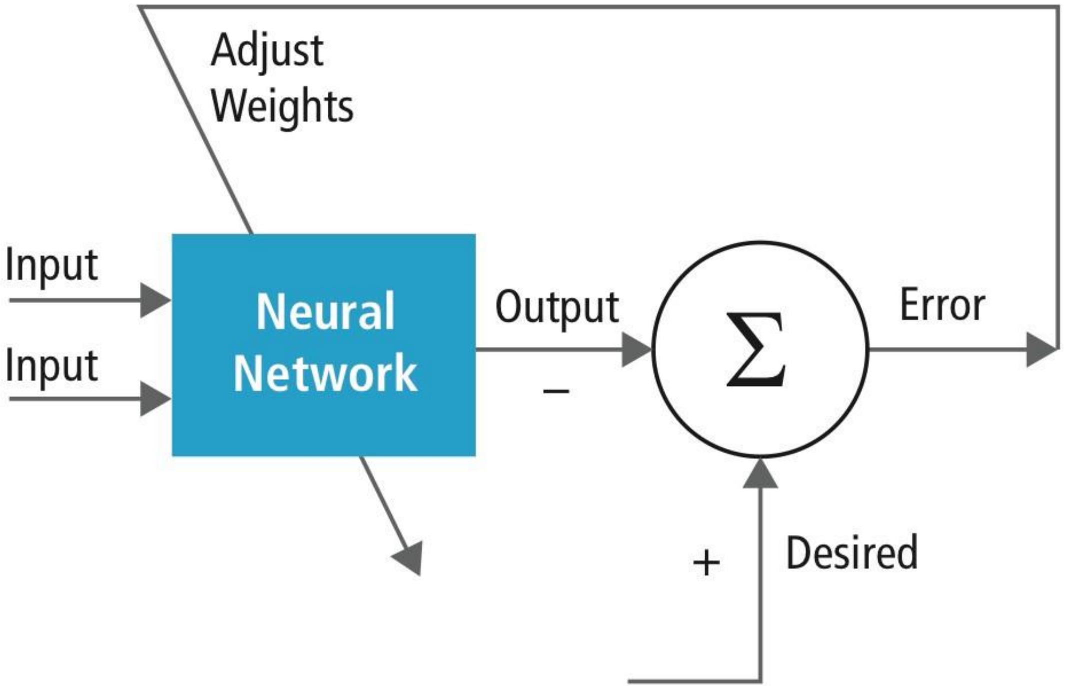
- Uma “máquina”?
  - Bons exemplos
  - Erros - correção

Humano vs "Máquina"

- Cérebro vs Modelo Neuronal (2000/2010)

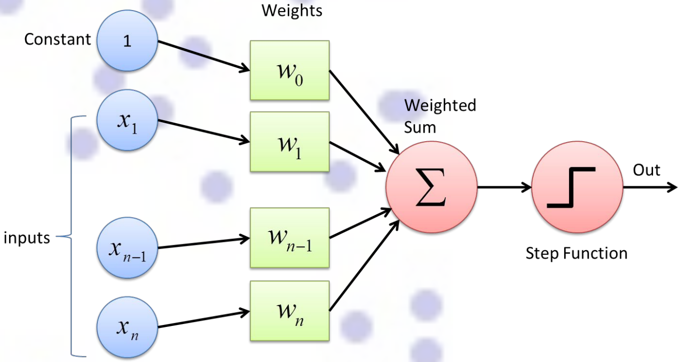
Parâmetro	Cérebro	Computador
Material	Orgânico	Metal e Plástico
Velocidade	Milisegundos	Nanosegundos
Tipo de Processamento	Paralelo	Sequêncial
Armazenamento	Adaptativo	Estático
Controle de Processos	Distribuído	Centralizado
Número de Elementos Processados	10^{11} a 10^{14}	10^{5} a 10^{6}
Ligações entre Elementos Processados	10.000	< 10

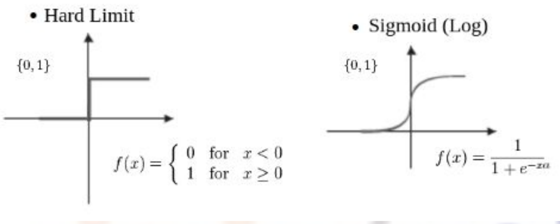
Rede Neuronal

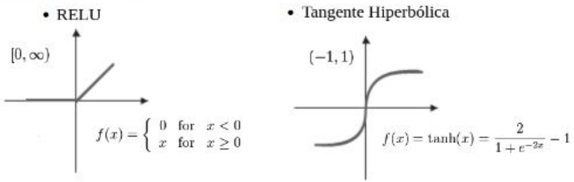


O Perceptron

- 1957 - Criação
  - Entradas
  - Pesos e *Bias*
  - Soma Ponderada
  - Função de Ativação
  - ***Classificador Binário



- 



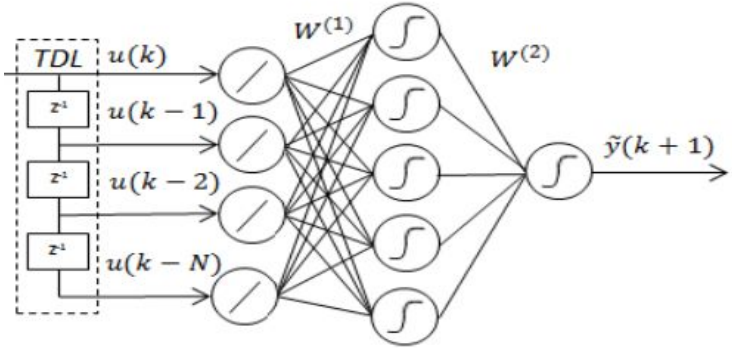
- Rosenblatt, Frank (1957), The Perceptron -- a perceiving and recognizing automaton. Report 85-460-1, Cornell Aeronautical Laboratory.

O Perceptron

- 1957 - Algoritmo de aprendizado???
- 1969 - Caiu em descrédito
  – Evolução da IA Clássica
- 1986 - O Back-propagation*
- 1989 - Teorema Universal de Aprixomação
  – Uma camada, qualquer função contínua
- 1991 - Multiple-Layer Perceptron
  – Função de ativação

- Rumelhart, David E.; Hinton, Geoffrey E.; Williams, Ronald J. (8 October 1986). "Learning representations by back-propagating errors". *Nature*. **323** (6088): 533–536.

Propagação (*forward*)



$$z_j = f(\sum_i w_{ji}^{(1)} x_i),$$

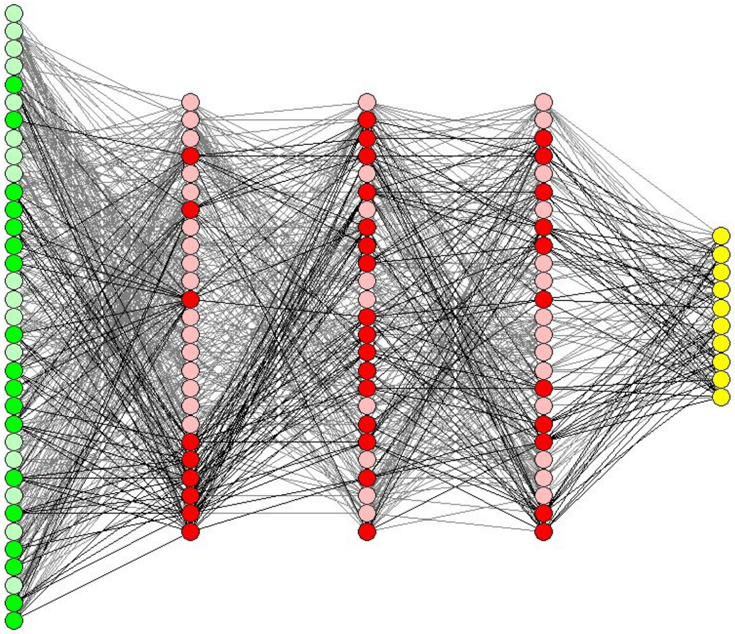
$$\tilde{y}(k+1) = g(\sum_j w_j^{(2)} z_j),$$

- onde z_j é o valor de saída do neurônio
  - ỹ(k+1) é o valor do último neurônio
  - $w^{(1)}$ e $w^{(2)}$ são os pesos da cam. escondida e saída
  - f() and g() são as saídas das funções de ativação da camada escondida e saída

  - Rumelhart, David E.; Hinton, Geoffrey E.; Williams, Ronald J. (8 October 1986). "Learning representations by back-propagating errors". *Nature*. **323** (6088): 533–536.

Propagação (*forward*)

- *Multi-Layer Perceptron* (MLP)

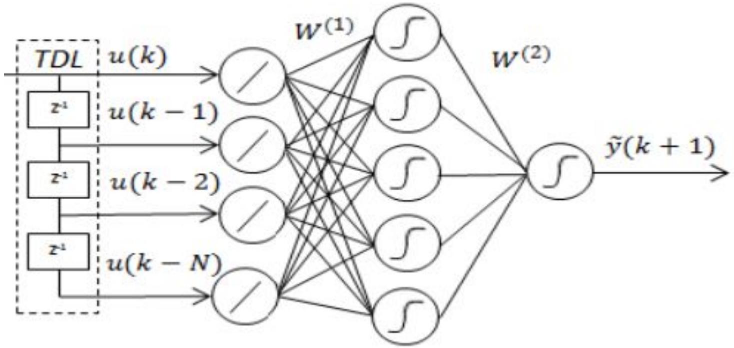


Back-propagation

- Cálculo de gradientes locais:

$$\delta^{OUT} = t(k+1) - \tilde{y}(k+1),$$

$$\delta_j^{HID} = f'(z_j) w_j^{(2)} \delta^{OUT}.$$



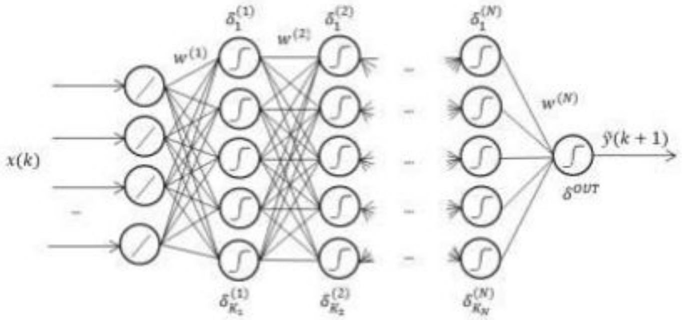
- $$\frac{\partial E(k)}{\partial w_j^{(2)}} = \delta^{OUT} z_j,$$ derivadas

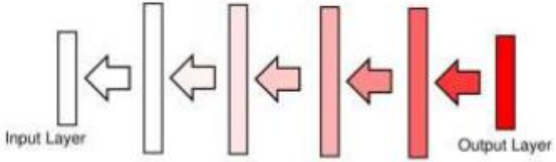
$$\frac{\partial E(k)}{\partial w_{ji}^{(1)}} = \delta_j^{IN} x_i.$$

- Rumelhart, David E.; Hinton, Geoffrey E.; Williams, Ronald J. (8 October 1986). "Learning representations by back-propagating errors". *Nature*. **323** (6088): 533–536.

Efeito ruim do *vanish gradient*

- Multiple-layer Perceptron





$$\delta_j^{(m)} = f_j^{(m-1)\prime} \sum_i w_{ij}^{(m)} \delta_i^{(m+1)},$$

$$\frac{\partial E(k)}{\partial w_{ji}^{(m)}} = \delta_j^{(m)} z_i^{(m-1)},$$

$$\frac{\partial E(k)}{\partial w_{ji}^{(m)}} \to 0 \;\; for\, m \to 1$$

Multiple-layer Perceptron

- Cenário em 1995-2000s
  - Limitadas a 2/3 camadas (*vanish gradient*)
  - Processo de aprendizado caro computacionalmente
  - A escolha da arquitetura é dependente da complexidade do problema
  - Ajuste de parâmetros (*learning rate* & *momentum*)

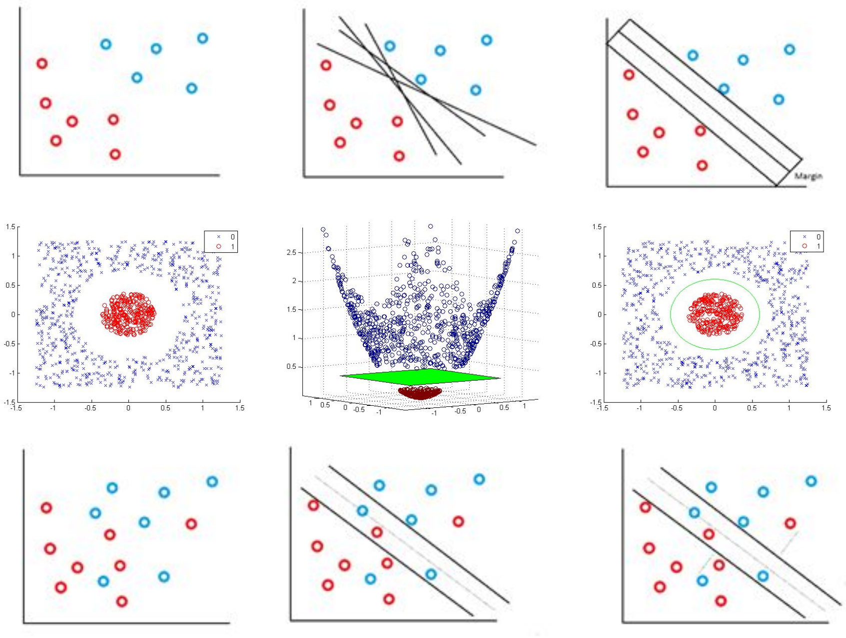
Support Vector Machines

- 1995

Linear (1963)

Kernel Trick (1992)

Soft Margin (1995)



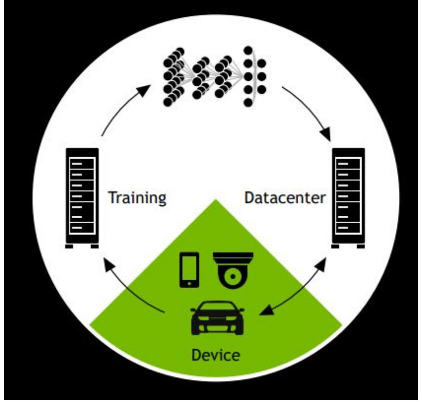
- Cortes, C. & Vapnik, V. N. Support-vector Networks. Machine Learning 20 (3), 1995

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)

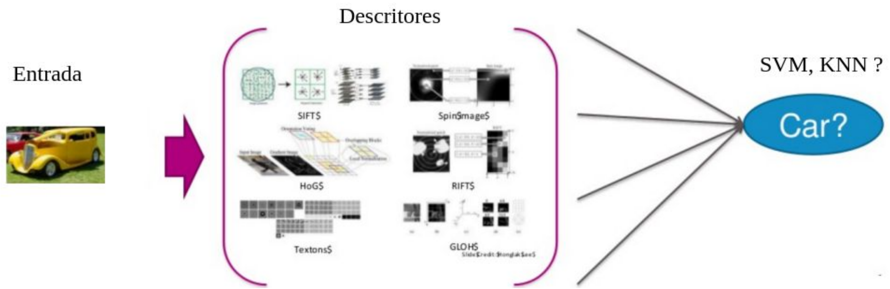
Deep Learning

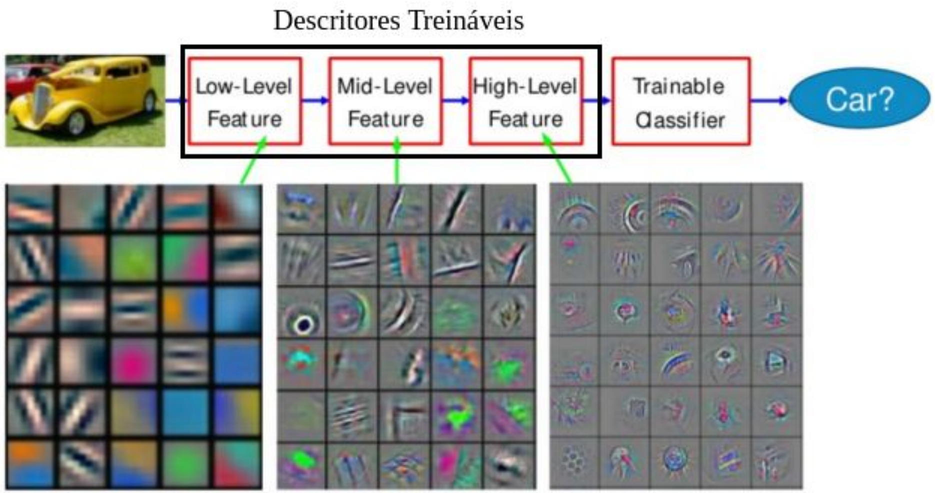


- Fatores que influenciam
  - Desenvolvimento de Algoritmos
  - Hardware capaz de processar grandes dados (GPU)

  - Disponibilidade de dados anotados
    - Imagens no Facebook
    - Voz e Texto no Google

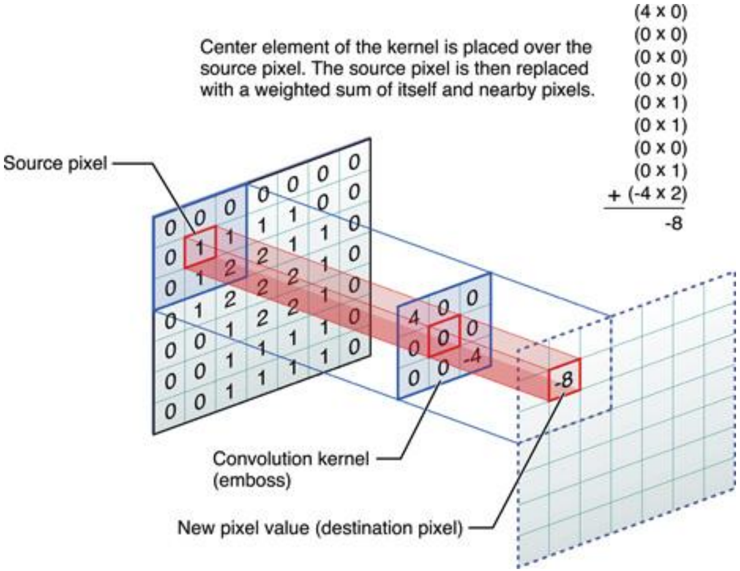
Tradicional x Deep Learning

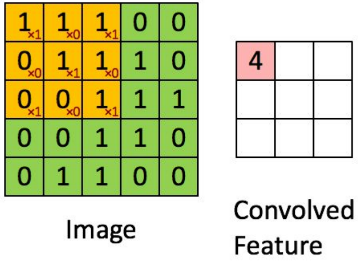




Convolução

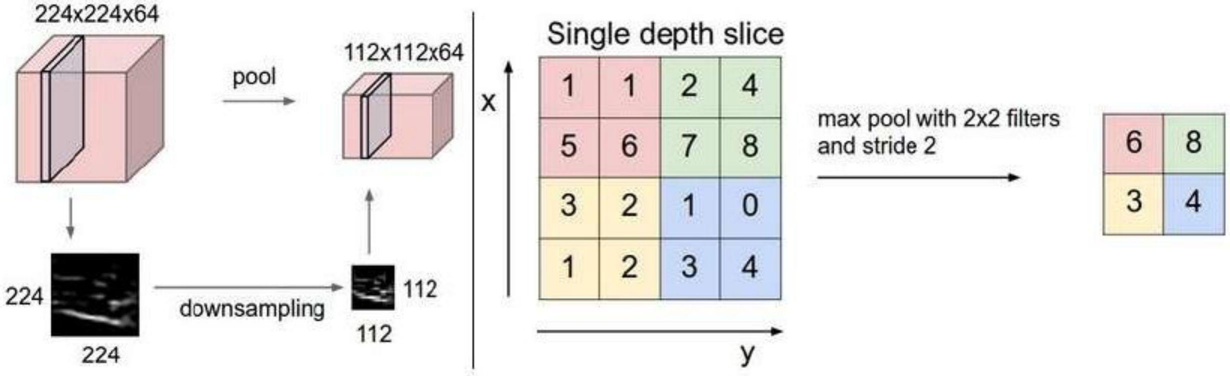
- Chave em *Deep Learning*
  - Pesos compartilhados





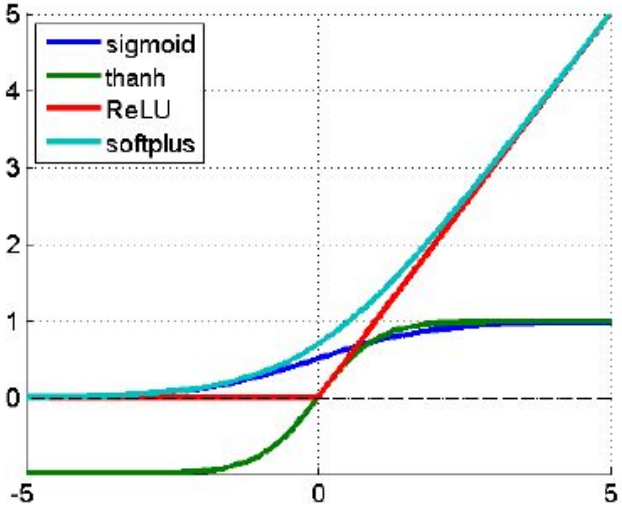
Pooling

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



Funções de ativação

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



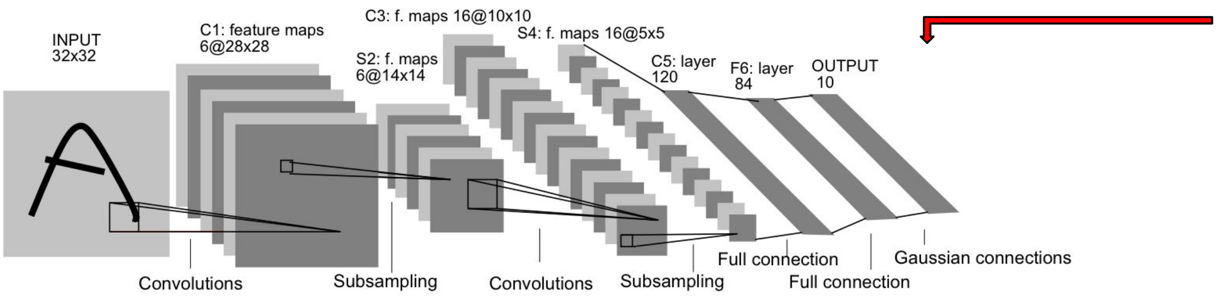
$$f(x) = \max(0, x)$$

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

- 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 (1998)
  - 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**

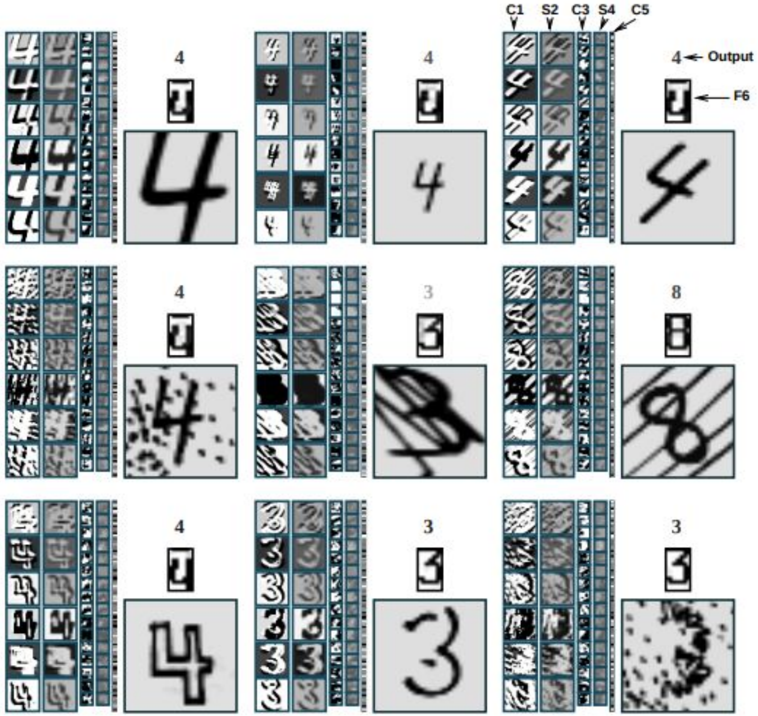
  - Out: fully: 84x10

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

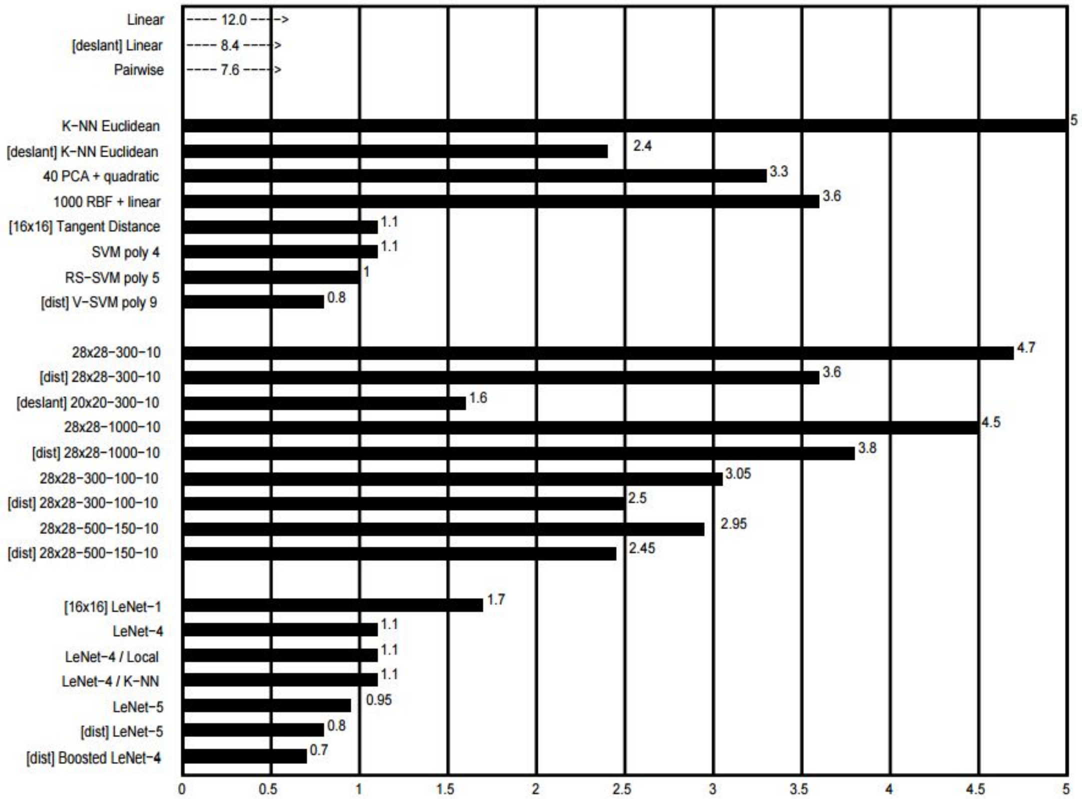
Camada F6

LeNet - Funcionamento

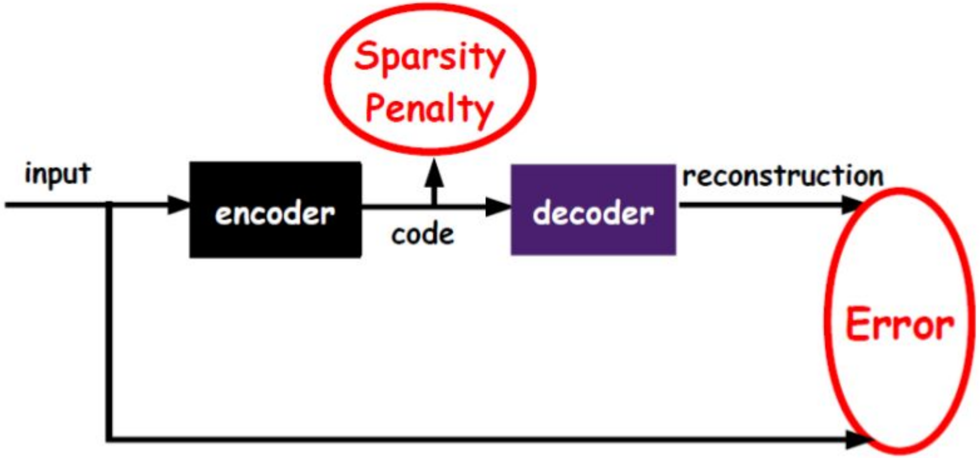


Robusto a ruídos, inclinação e forma

LeNet vs *State of the Art*



Sparse Autoencoders

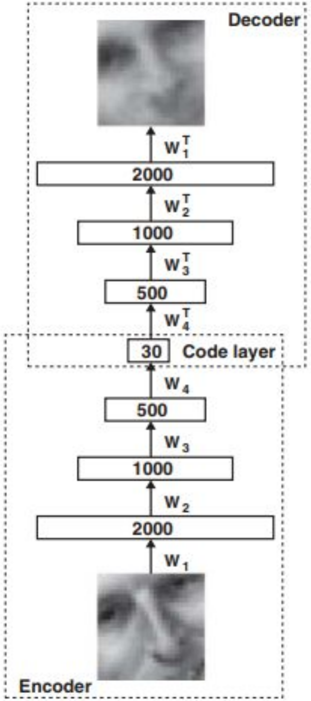


- input: X code: $h = W^T X$

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

Eventos

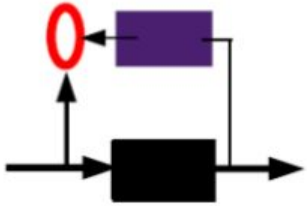
- **2006 - Autoencoders**
  - Redução de Dimensional.
  - Uso de RBM empilhados como um deep auto-encoder
    i. Treinar RBM usando imagens como saída e entrada



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

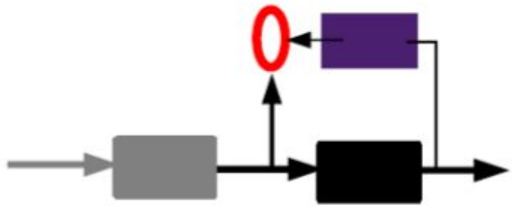
Unsupervised Pre-training

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



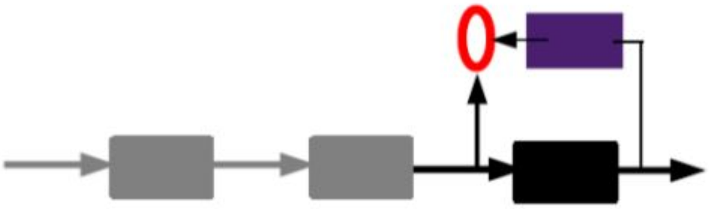
Unsupervised Pre-training

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



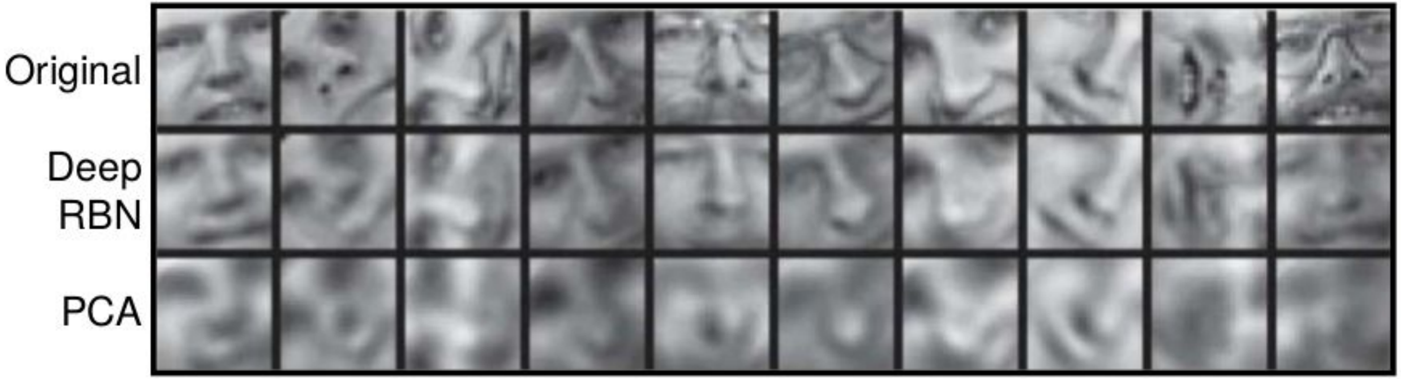
Unsupervised Pre-training

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



Eventos

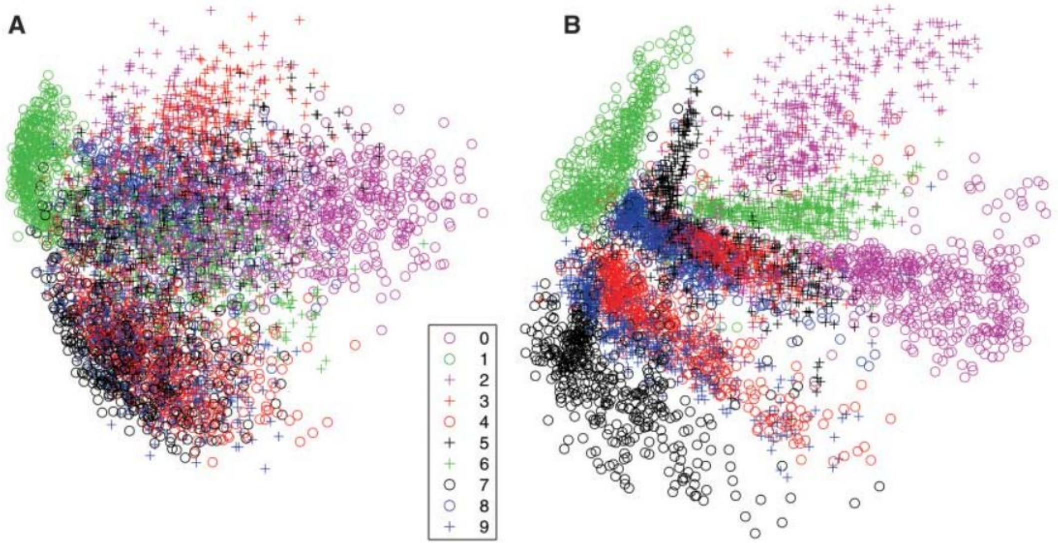
Olivetti face data, 25x25 pixel images reconstructed from 30 dimensions (625 → 30)



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

Eventos

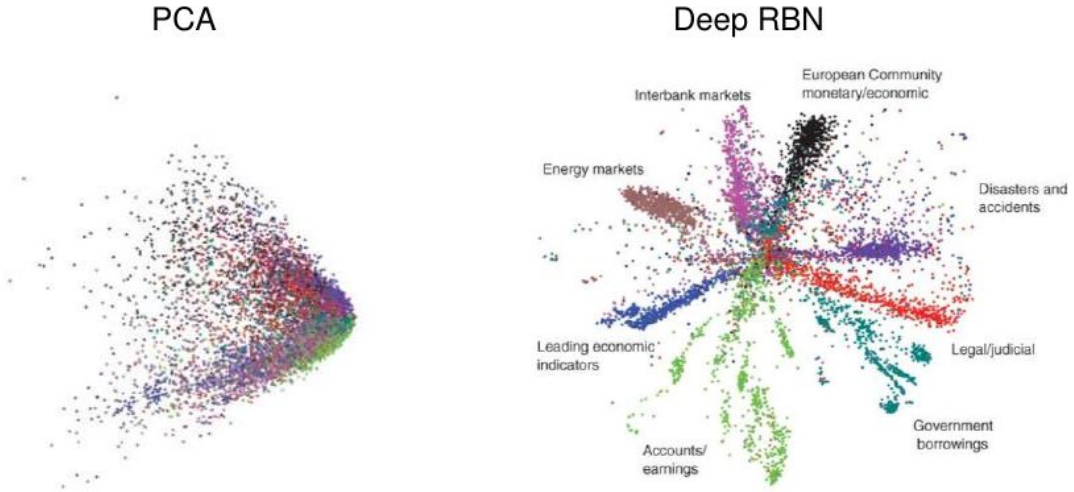
- **2006 - Autoencoders**
  - PCA vs (784-1000-500-250-2) autoencoder   * $28^2 = 784$



- 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



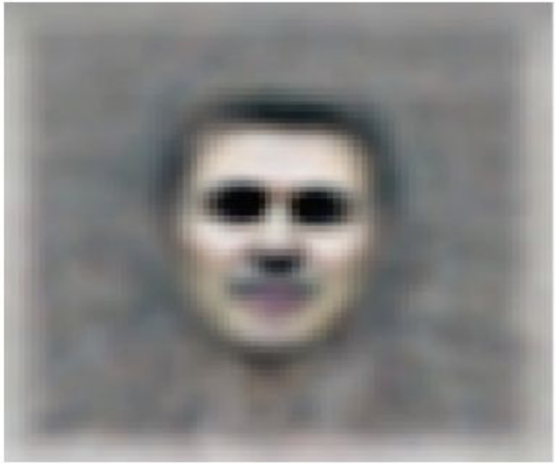
Eventos

- 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)

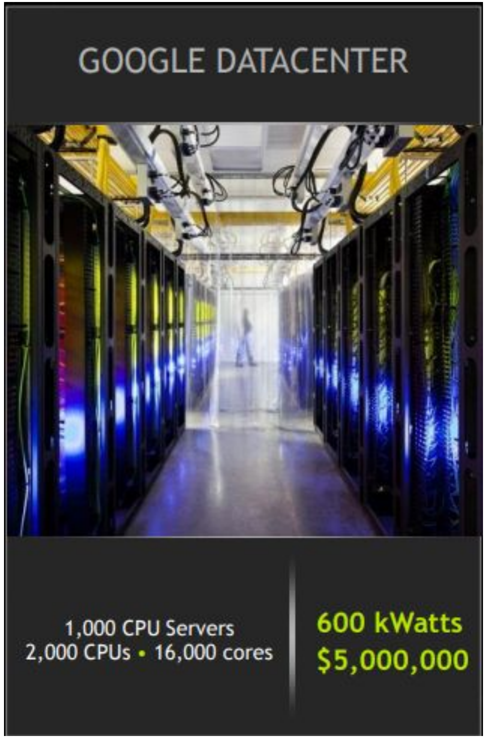


- Q. V. Le, M. Ranzato, R. Monga, M. Devin, K. Chen, G. S. Corrado, J. Dean, A. Ng. Building High-Level Features using Large Scale Unsupervised Learning in ICML 2012 (citado por 1610 jul/2018)

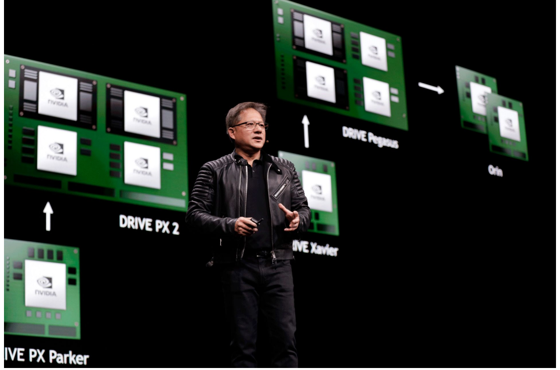
Eventos

- **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

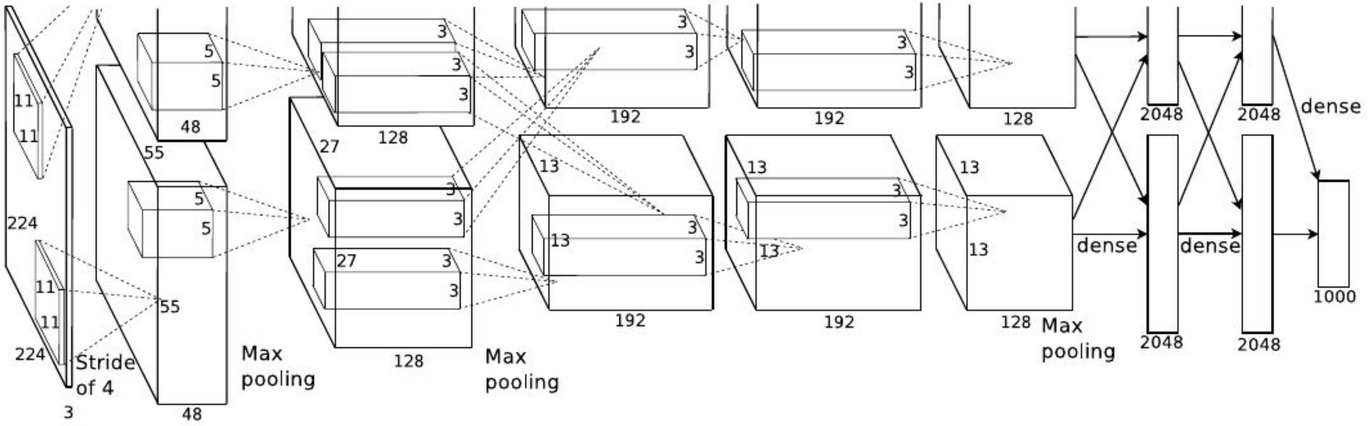
- **2013 - NVIDIA**
  - GTC (2014)
  - 3 Servidor (12 GPUs)
    - 18.432 cores
  - 4 KWatts
  - U$ 33.000

- 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

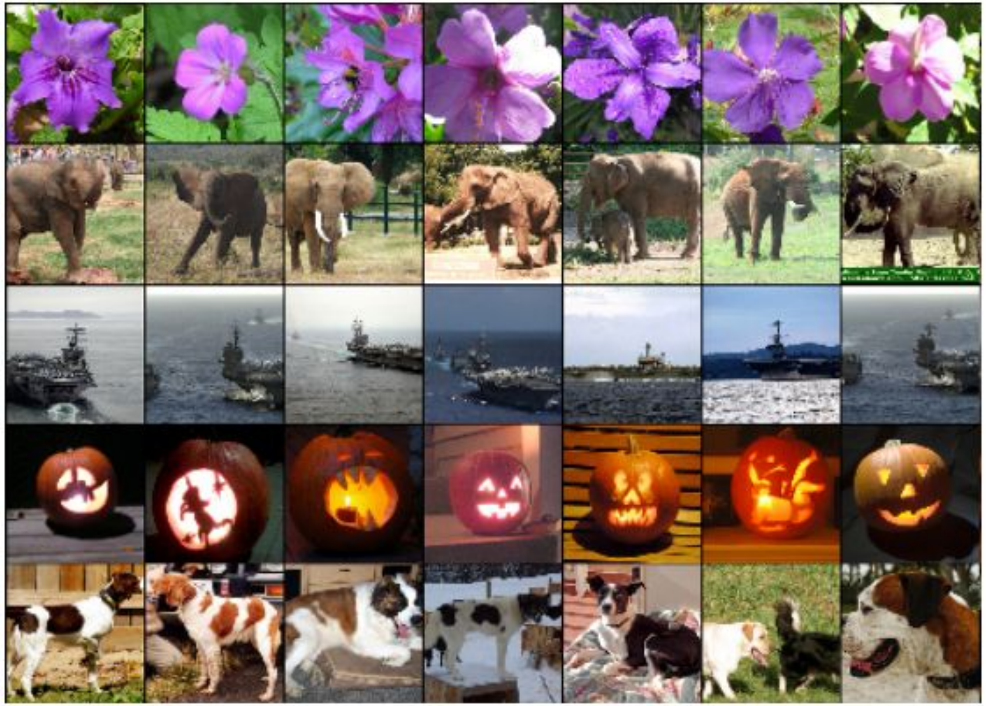


  - 60M parametrôs / 650K neurônios
  - 3 GPUs - limitado tamanho modelo pela RAM das GPUs
  - ImageNet (15M images / 22k classes) - erro caiu para 15%

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

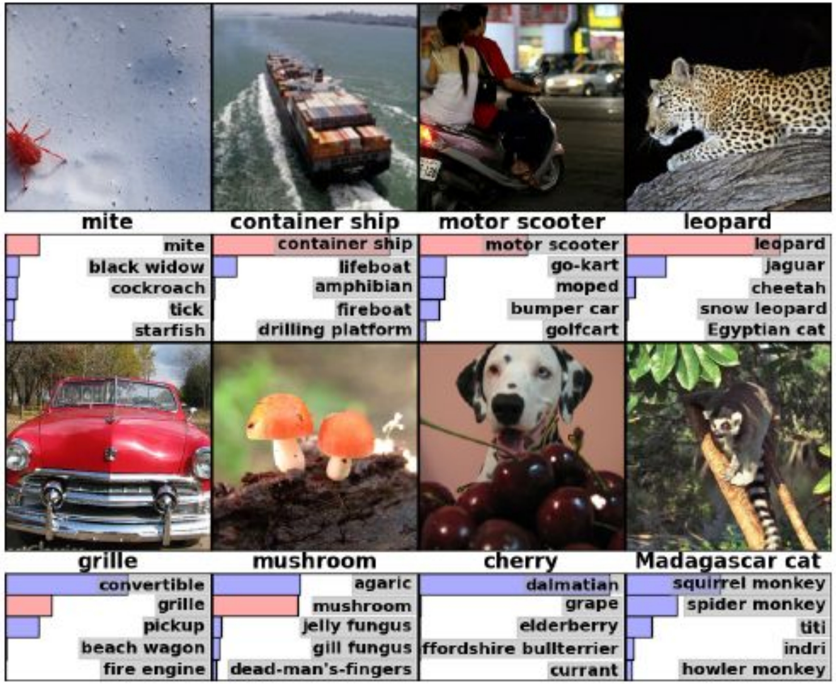
AlexNet (ImageNet)

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



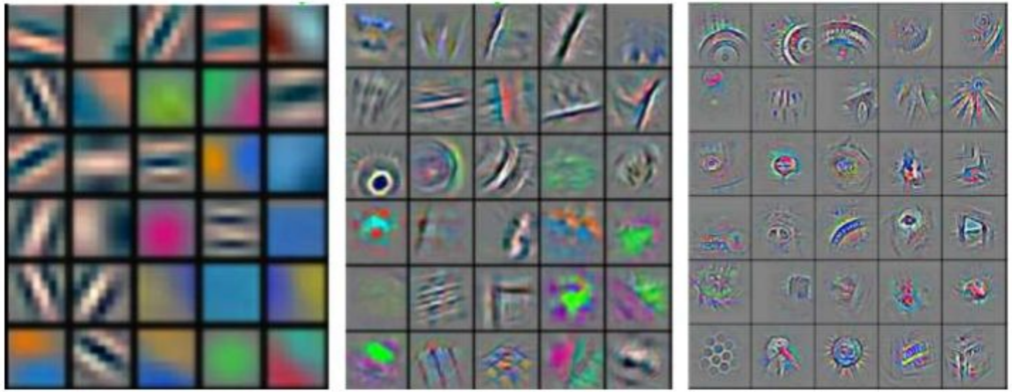
AlexNet (ImageNet)

- **Classes *Top 5***



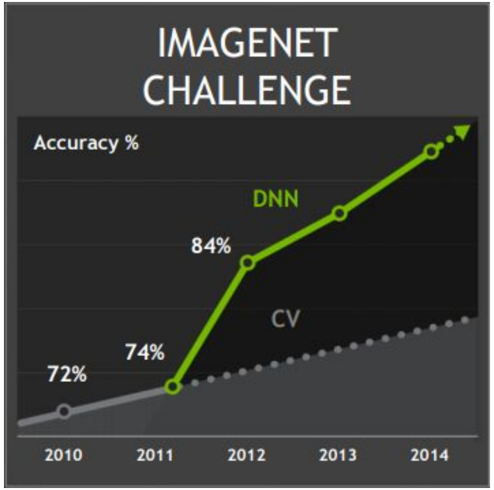
Hierarquia das Representações Treinadas

Características de baixo nível → Características de médio nível → Características de alto nível



- 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

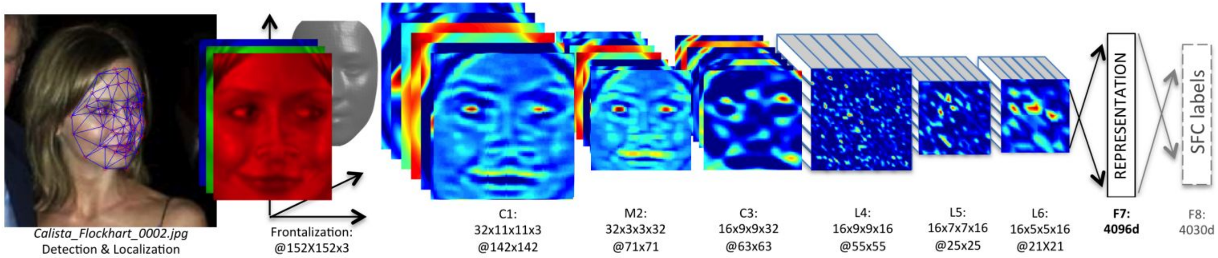
"Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariant Shift"

- 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**



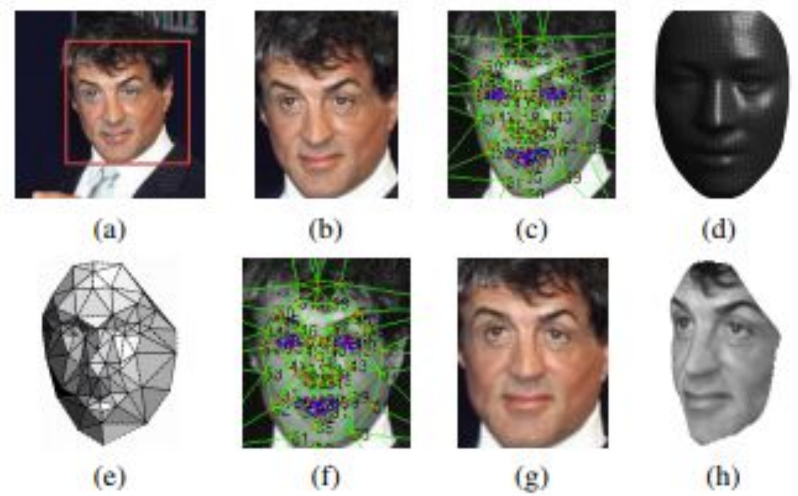
Modelo	# de parâmetros	Acurácia (%)
Deep Face	128 M	97,35
State-of-the-art [2013]	~	79,70
Nível Humano	N / A	97,50

– Treinada com 4M imagens faciais

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

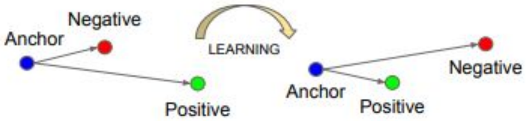
Eventos

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

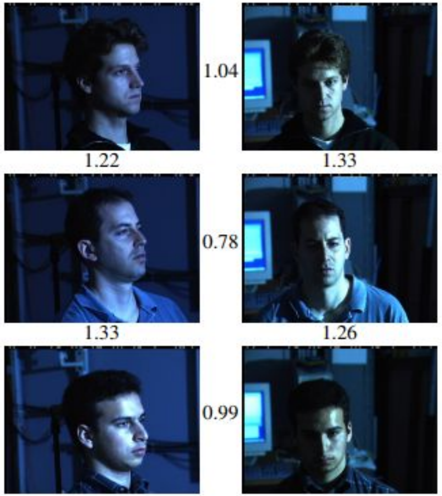


Eventos

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



  - Triplet loss



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

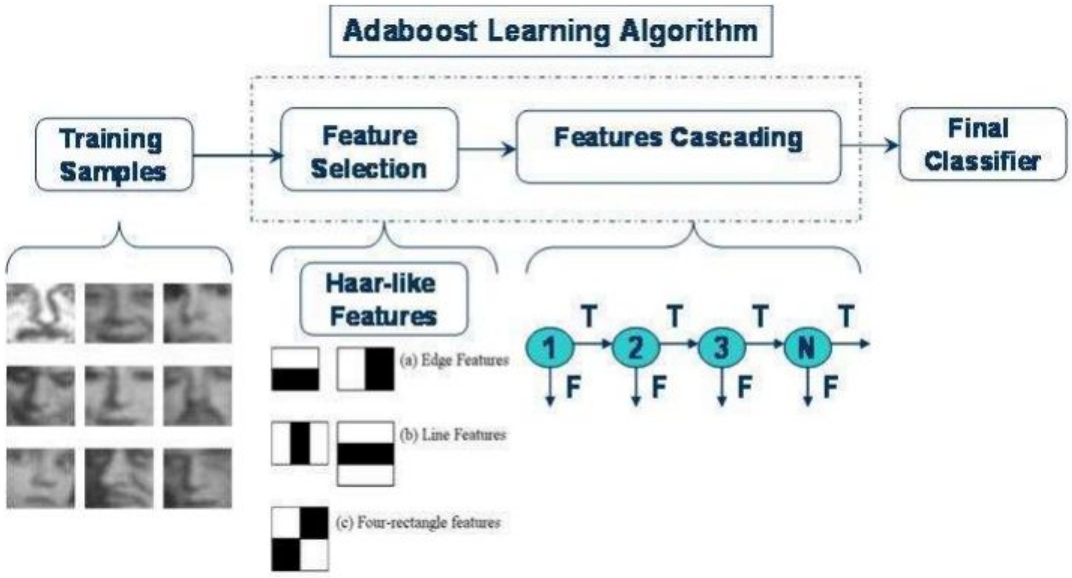
- H. Sak, A. Senior, F. Beaufays. Long Short-Term Memory Recurrent Neural Network Architectures for Large Scale Acoustic Modeling // INTERSPEECH'2014.
- K. Vesely, A. Ghoshal, L. Burget, D. Povey. Sequence-discriminative training of deep neural networks // INTERSPEECH'2014.

Inteligência Artificial

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

Eventos

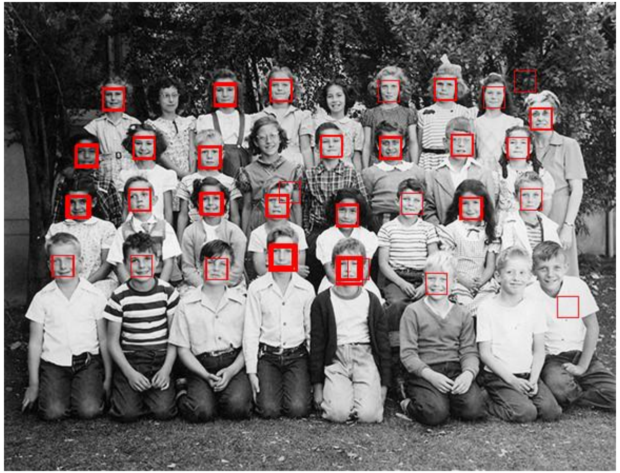
- **2001 - Face Detection - Viola & Jones**



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

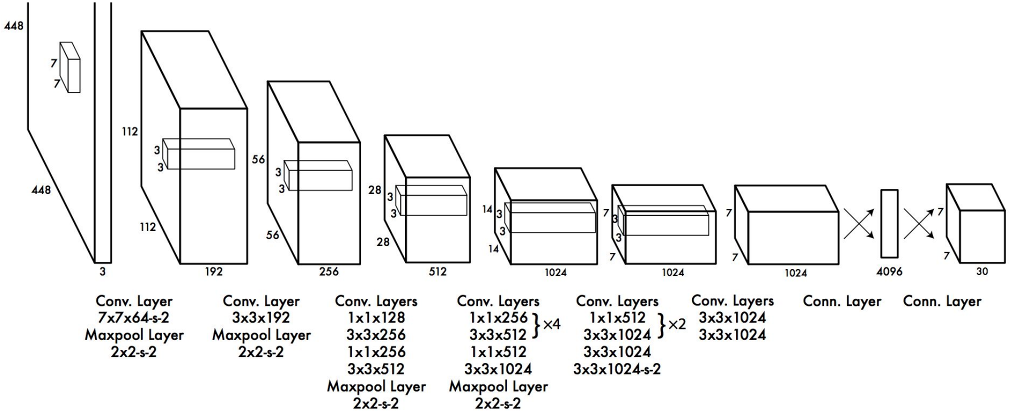
Eventos

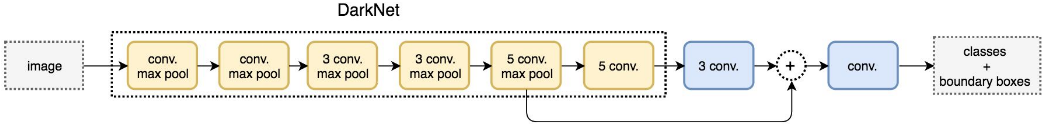
- **2001 - Object Detection - Viola & Jones**
  - Semanas para treinar
    - Características simples
  - Imagens de 384x288 pixels
  - 15 fps 700 Mhz
    - Pentium III
  - ~94 de acurácia
  - Máquinas Fotogr.



Eventos

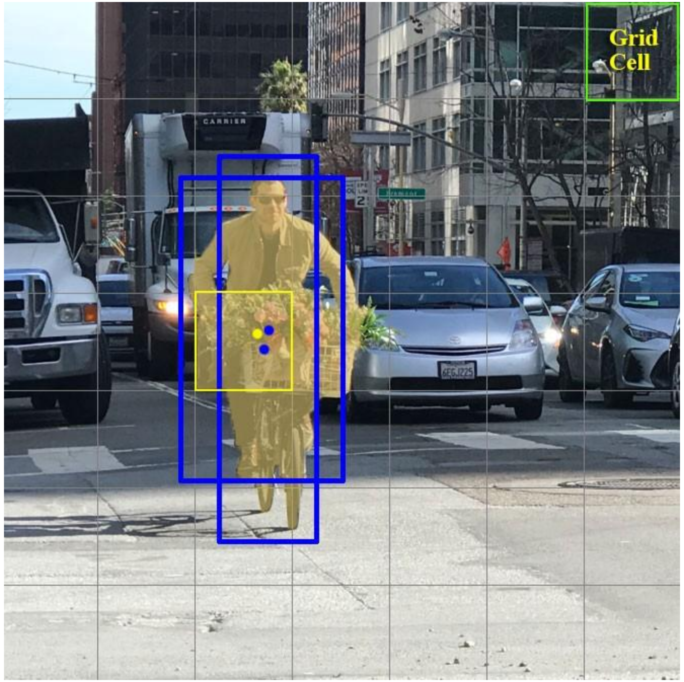
- **2015-2017 Yolo (Object Detection)**





Eventos

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



Object Detection

- **Demonstração YoloV2 & YoloV3 & YoloV4**
  - **Vídeos ~ 1 minutos**

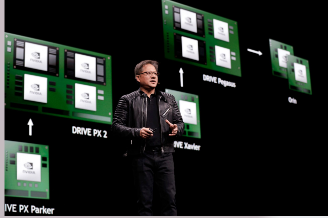
Eventos

- **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



Eventos

NVIDIA GPU Technology Conference GTC 2021

- https://www.youtube.com/watch?v=jhDiaUL_RaM
- A conversational AI (Avatars - ~28 minutes)
  - Toy-Me - tudo em tempo real
    - Compreensão de Fala
    - Processamento de Linguagem Natural
    - Síntese de Voz (com a sua voz)
    - Computação Gráfica (*ray traced*)

Minhas Pesquisas

- **2008-2021 Identificação de Veículos**
  - Câmera e Veículo em movimento

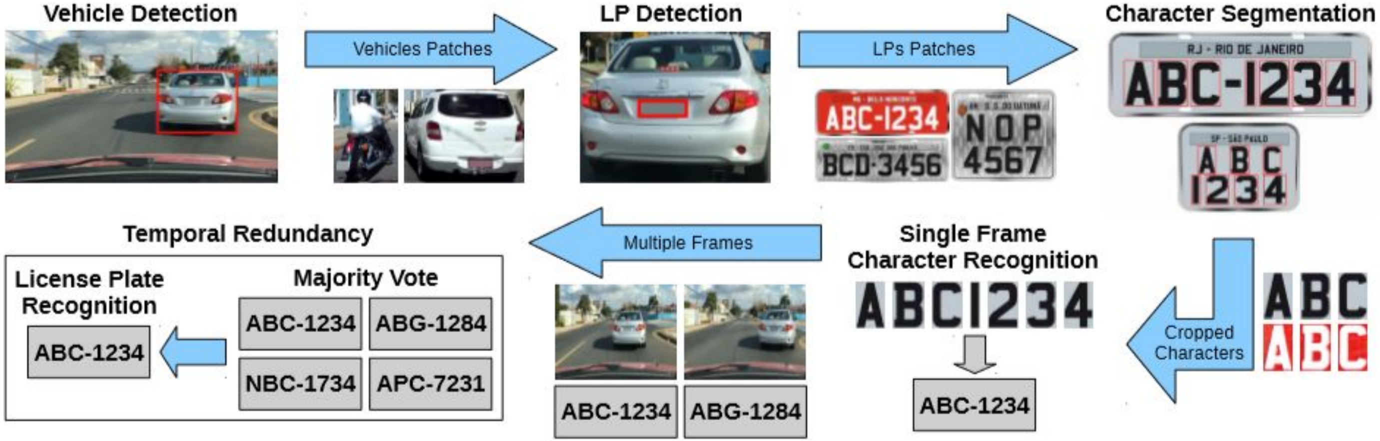
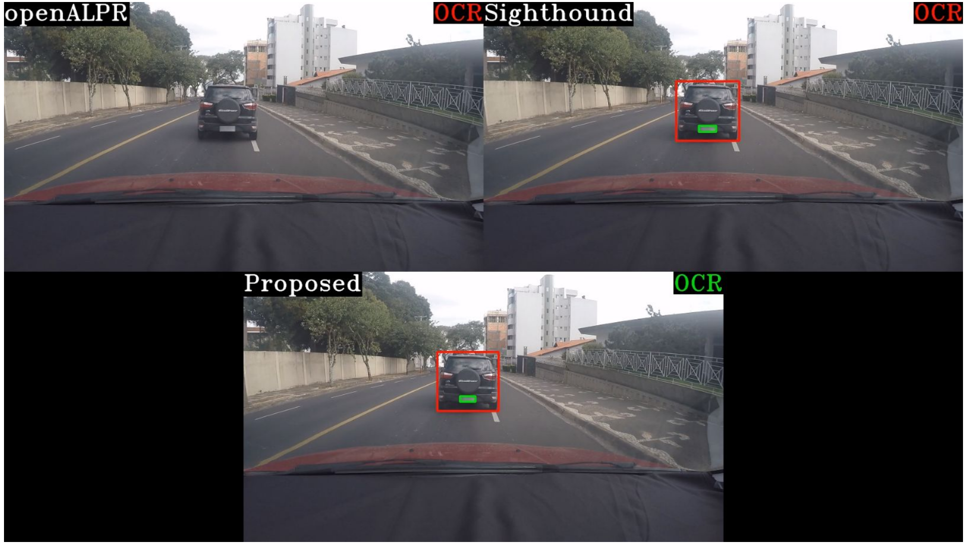


Fig. 5. An usual ALPR pipeline having temporal redundancy at the end.

- 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.

Minhas Pesquisas

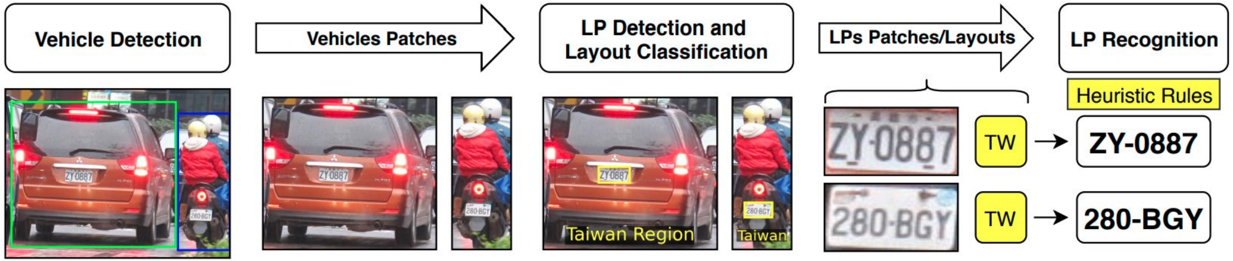
- **2018 Identificação de Veículos**
  - Vídeo



- 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.

Minhas Pesquisas

- **2021 Identificação de Veículos**



Dataset \ Approach	[58]	[38]	[7]	[37]	[17]	Sighthound	OpenALPR	No Vehicle Detection*	No Layout Classification†	Proposed
Caltech Cars	–	–	–	–	–	95.7 ± 2.7	$\mathbf{99.1 \pm 1.2}$	98.3 ± 1.8	96.1 ± 1.8	98.7 ± 1.2
EnglishLP	**97.0**	–	–	–	–	92.5 ± 3.7	78.6 ± 3.6	95.3 ± 1.6	95.5 ± 2.4	95.7 ± 2.3
UCSD-Stills	–	–	–	–	–	**98.3**	**98.3**	98.0 ± 0.7	97.3 ± 1.9	98.0 ± 1.4
ChineseLP	–	–	–	–	–	90.4 ± 2.4	92.6 ± 1.9	97.0 ± 0.7	95.4 ± 1.1	$\mathbf{97.5 \pm 0.9}$
AOLP	–	**99.8**‡	–	–	–	87.1 ± 0.8	–	98.8 ± 0.3	98.4 ± 0.7	99.2 ± 0.4
OpenALPR-EU	–	–	93.5	85.2	–	93.5	91.7	$\mathbf{97.8 \pm 0.5}$	96.7 ± 1.9	$\mathbf{97.8 \pm 0.5}$
SSIG-SegPlate	–	–	88.6	89.2	85.5	82.8	92.0	96.5 ± 0.9	96.9 ± 0.5	$\mathbf{98.2 \pm 0.5}$
UFPR-ALPR	–	–	–	–	64.9	62.3	82.2	59.6 ± 0.9	82.5 ± 1.1	$\mathbf{90.0 \pm 0.7}$
Average	–	–	–	–	–	87.8 ± 2.4	90.7 ± 2.3	92.7 ± 0.9	94.8 ± 1.4	$\mathbf{96.9 \pm 1.0}$

* A modified version of our approach in which the LPs are detected (and their layouts classified) directly in the original image (i.e., without vehicle detection).
† The proposed ALPR system assuming that all LP layouts were classified as undefined (i.e., without layout classification and heuristic rules).
‡ The LP patches for the LP recognition stage were cropped directly from the ground truth in [38].

Minhas Pesquisas

- **2020/2021 Identificação de Veículos**
  - Prêmio SBC 3a Dissertação
  - NVIDIA *Applitcations* / +300 citações / +800 *downloads*

Dataset \ Approach	[58]	[38]	[7]	[37]	[17]	Sighthound	OpenALPR	No Vehicle Detection*	No Layout Classification†	Proposed
Caltech Cars	–	–	–	–	–	95.7 ± 2.7	$\mathbf{99.1 \pm 1.2}$	98.3 ± 1.8	96.1 ± 1.8	98.7 ± 1.2
EnglishLP	**97.0**	–	–	–	–	92.5 ± 3.7	78.6 ± 3.6	95.3 ± 1.6	95.5 ± 2.4	95.7 ± 2.3
UCSD-Stills	–	–	–	–	–	**98.3**	**98.3**	98.0 ± 0.7	97.3 ± 1.9	98.0 ± 1.4
ChineseLP	–	–	–	–	–	90.4 ± 2.4	92.6 ± 1.9	97.0 ± 0.7	95.4 ± 1.1	$\mathbf{97.5 \pm 0.9}$
AOLP	–	**99.8**‡	–	–	–	87.1 ± 0.8	–	98.8 ± 0.3	98.4 ± 0.7	99.2 ± 0.4
OpenALPR-EU	–	–	93.5	85.2	–	93.5	91.7	$\mathbf{97.8 \pm 0.5}$	96.7 ± 1.9	$\mathbf{97.8 \pm 0.5}$
SSIG-SegPlate	–	–	88.6	89.2	85.5	82.8	92.0	96.5 ± 0.9	96.9 ± 0.5	$\mathbf{98.2 \pm 0.5}$
UFPR-ALPR	–	–	–	–	64.9	62.3	82.2	59.6 ± 0.9	82.5 ± 1.1	$\mathbf{90.0 \pm 0.7}$
Average	–	–	–	–	–	87.8 ± 2.4	90.7 ± 2.3	92.7 ± 0.9	94.8 ± 1.4	$\mathbf{96.9 \pm 1.0}$

* A modified version of our approach in which the LPs are detected (and their layouts classified) directly in the original image (i.e., without vehicle detection).
† The proposed ALPR system assuming that all LP layouts were classified as undefined (i.e., without layout classification and heuristic rules).
‡ The LP patches for the LP recognition stage were cropped directly from the ground truth in [38].

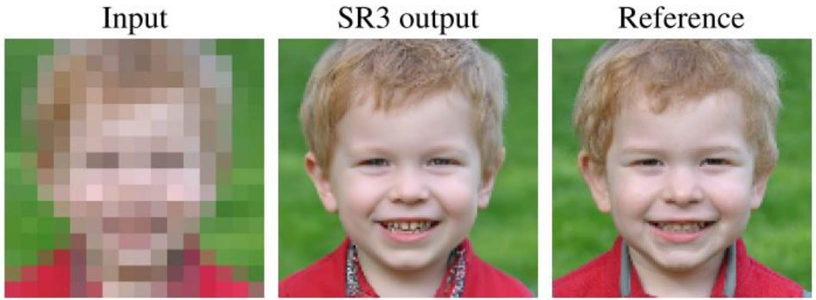
- R. Laroca, L. A. Zanlorensi, G. R. Gonçalves, E. Todt, W. R. Schwartz, D. Menotti, “An Efficient and Layout-Independent Automatic License Plate Recognition System Based on the YOLO Detector,” IET Intelligent Transport Systems, 2021.

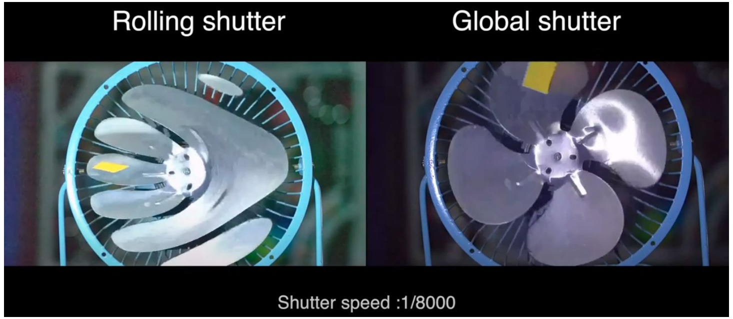
Minhas Pesquisas

- **2020 Rec. Automático Medidores de leitura**

Minhas Pesquisas

- **Face & ALPR - Superresolução - em desenvolvimento**





Minhas Pesquisas (Capes & PF)

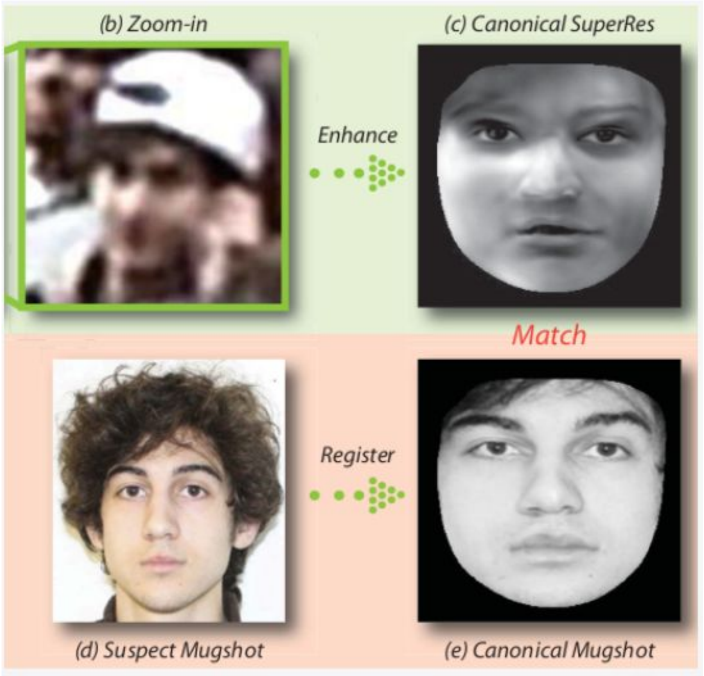
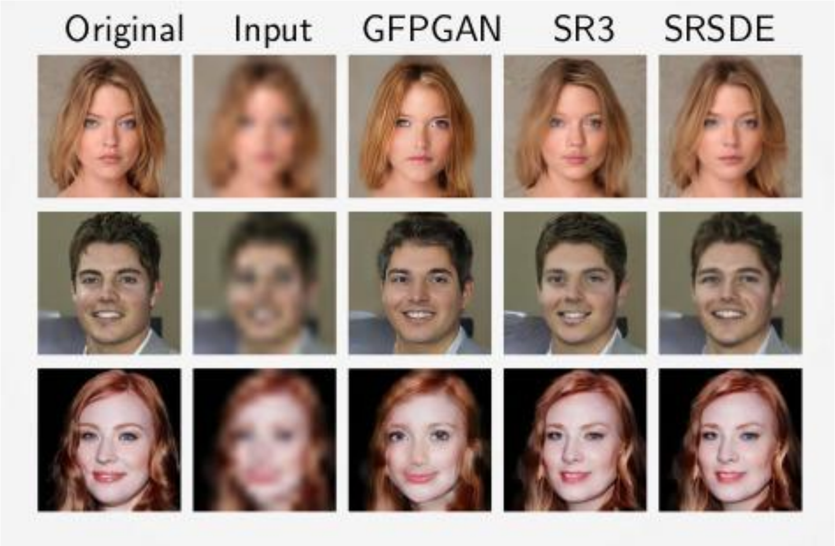
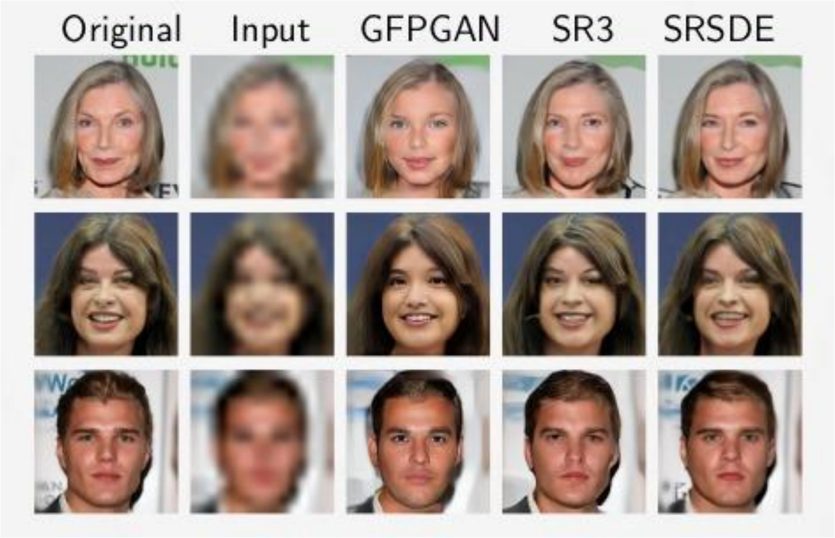
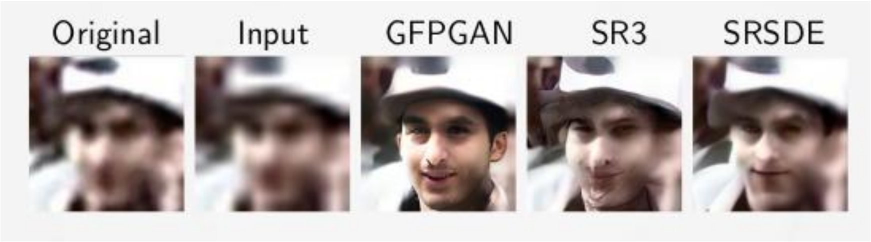


Imagem de vigilância do suspeito de atentado a bomba na Maratona de Boston em 2013 - resultado da literatura de 2019

Minhas Pesquisas (Capes & PF)







Minhas Pesquisas (Capes & PF)

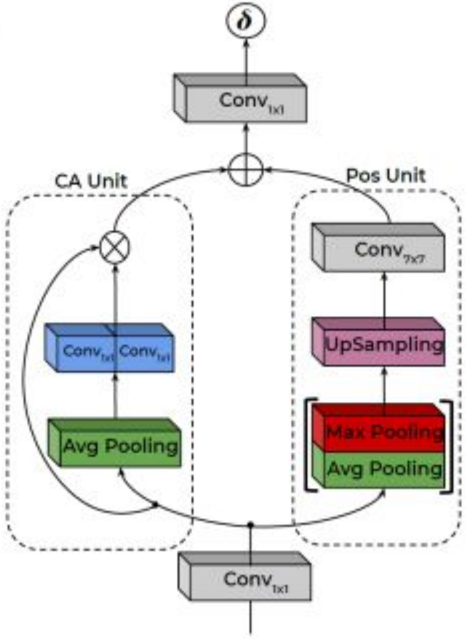
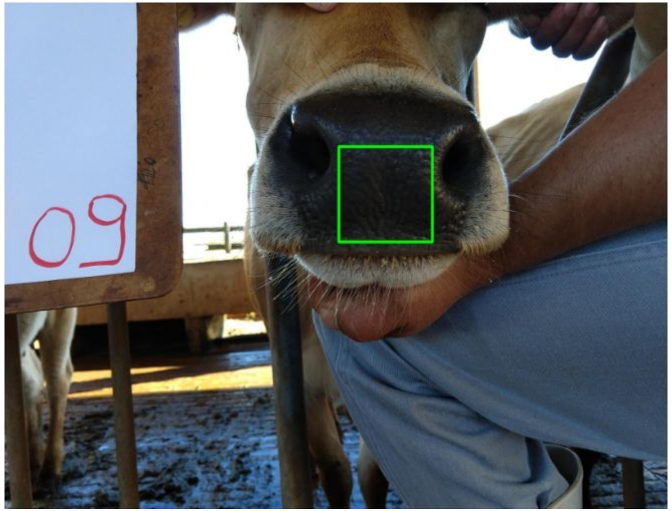


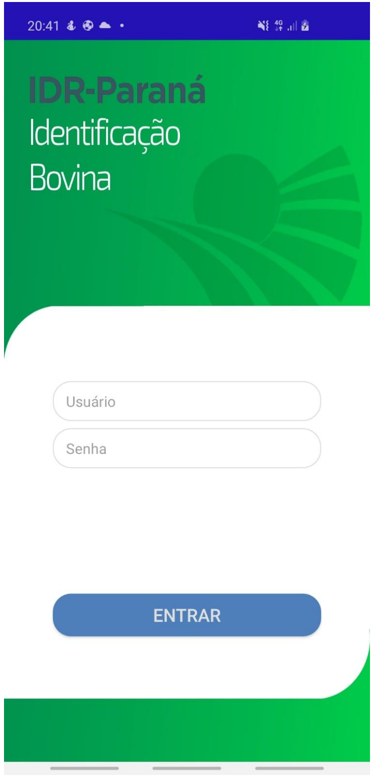
Figure 1.1: Two-Fold Attention Mechanism. Source: (Mehri et al., 2020)

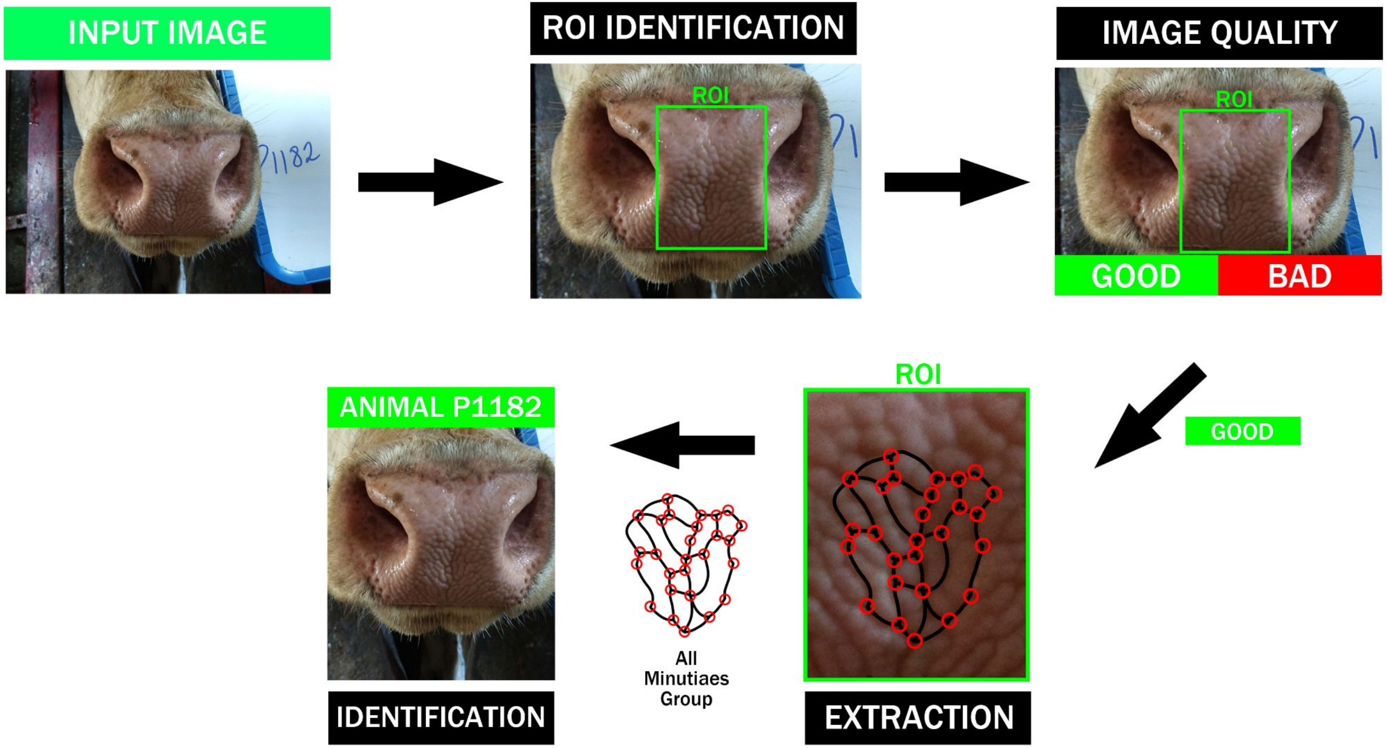
Minhas Pesquisas (Capes & PF)

Minhas Pesquisas (IDR-PR)

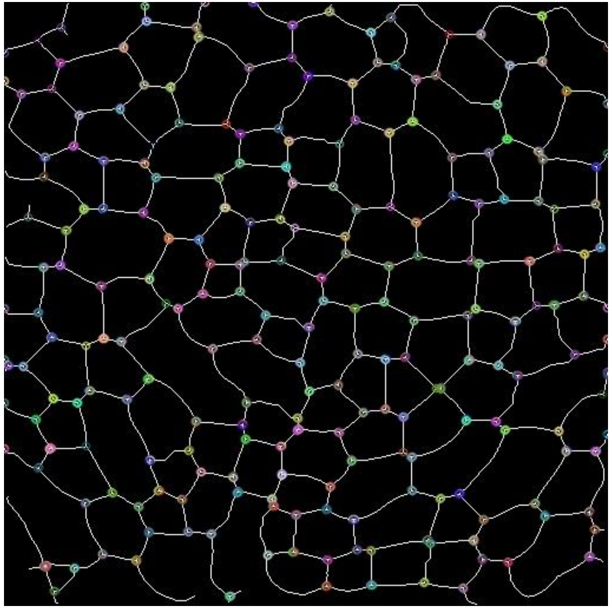


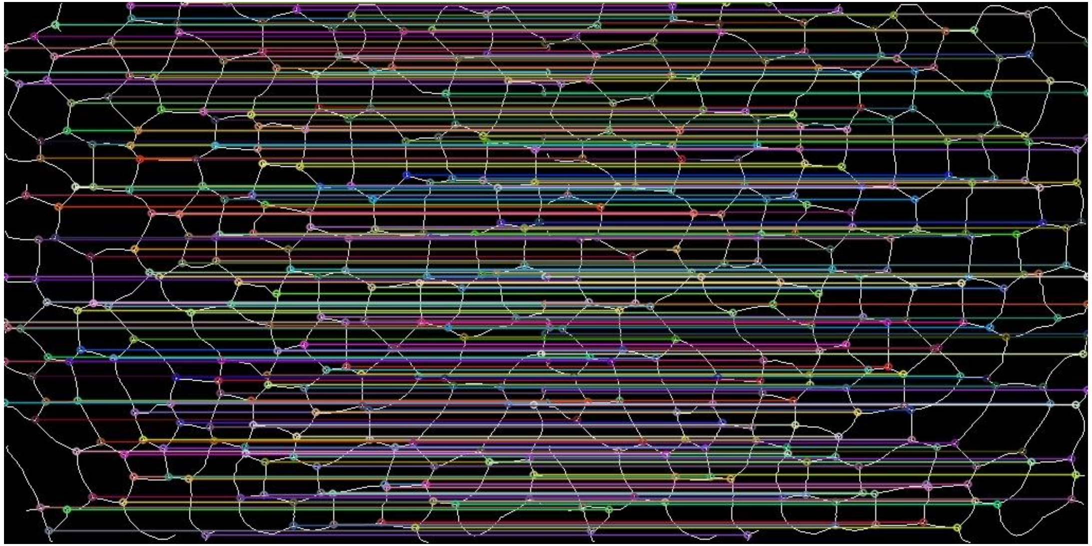
Minhas Pesquisas (IDR-PR)





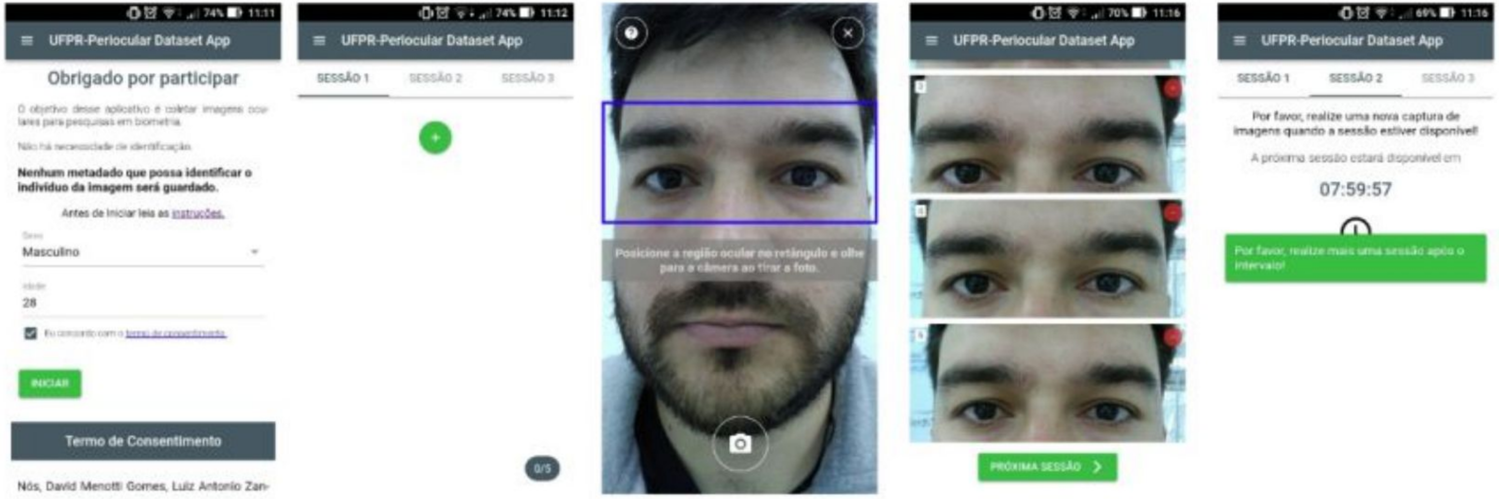
Minhas Pesquisas (IDR-PR)





Minhas Pesquisas

- **Reconhecimento Ocular - Sensores e Sessões**



- Zanlorensi, Laroca, Lucio, Santos, Britto & Menotti. **UFPR-Periocular**: A Periocular Dataset Collected by Mobile Devices in Unconstrained Scenarios, arXiv:2011.12427, 2020.

Minhas Pesquisas

- **Reconhecimento Ocular - Sensores e Sessões**

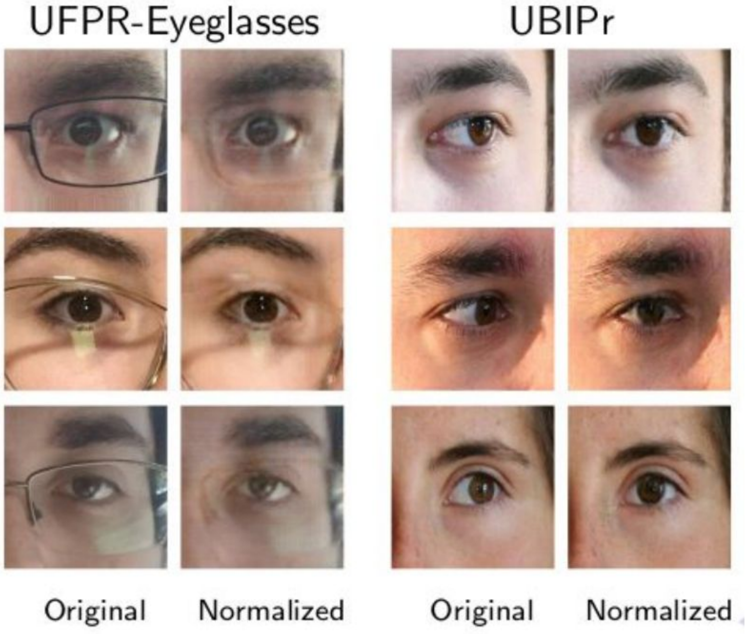
TABLE VI
BENCHMARK RESULTS IN THE CLOSED-WORLD PROTOCOL FOR THE IDENTIFICATION AND VERIFICATION TASKS.

Model	Identification (1:N)		Verification (1:1)		
	Rank 1 (%)	Rank 5 (%)	AUC (%)	EER (%)	Decidability
VGG16	50.56 ± 3.30	68.73 ± 3.01	99.41 ± 0.11	3.59 ± 0.32	4.4544 ± 0.1502
VGG16-Face	56.29 ± 1.62	73.84 ± 1.48	99.43 ± 0.08	3.44 ± 0.28	4.5069 ± 0.1379
Xception	57.43 ± 1.43	75.88 ± 1.52	99.77 ± 0.04	2.19 ± 0.18	4.2470 ± 0.0538
ResNet50V2	63.18 ± 2.14	77.79 ± 1.81	99.74 ± 0.04	2.24 ± 0.18	4.9382 ± 0.1184
InceptionResNet	65.16 ± 2.45	81.53 ± 1.99	99.78 ± 0.15	1.85 ± 0.40	4.5561 ± 0.1183
ResNet50	71.06 ± 1.14	85.22 ± 0.82	99.89 ± 0.02	1.41 ± 0.10	5.1242 ± 0.0634
ResNet50-Face	73.76 ± 1.43	86.86 ± 1.02	99.83 ± 0.03	1.74 ± 0.12	5.2400 ± 0.0837
DenseNet121	75.54 ± 1.36	88.53 ± 0.97	99.93 ± 0.02	1.11 ± 0.09	5.1730 ± 0.0497
MobileNetV2	77.98 ± 1.08	90.19 ± 0.79	99.93 ± 0.01	1.13 ± 0.07	5.2477 ± 0.0650
Multi-task	$\mathbf{84.32 \pm 0.71}$	$\mathbf{94.55 \pm 0.58}$	$\mathbf{99.96 \pm 0.01}$	$\mathbf{0.81 \pm 0.06}$	5.1978 ± 0.0340
Siamese	–	–	98.94 ± 0.22	4.86 ± 0.44	3.0005 ± 0.1871
Pairwise	–	–	99.44 ± 0.66	3.06 ± 1.84	$\mathbf{6.4503 \pm 1.2270}$

- Zanlorensi, Laroca, Lucio, Santos, Britto & Menotti. **UFPR-Periocular**: A Periocular Dataset Collected by Mobile Devices in Unconstrained Scenarios, arXiv:2011.12427, 2020.

Minhas Pesquisas

- **Reconhecimento Ocular - Óculos Gaze**



- Zanlorensi, Proença & Menotti, “Unconstrained Periocular Recognition: Using Generative Deep Learning Frameworks for Attribute Normalization,”. IEEE ICIP, 2020.

Minhas Pesquisas

- **Identificação de Pessoas - Ocular**
  - 5 a 50% de erro no reconhecimento facial

- NIST Launches Studies into Masks' Effect on Face Recognition Software, July 27, 2020, https://www.nist.gov/news-events/news/2020/07/nist-launches-studies-masks-effect-face-recognition-software
- Projeto BOVIFOCR: Biometria Ocular, Vivacidade de Imagens Faciais e Reconhecimento de Texto (OCR) em Documentos Oficiais - https://www.inf.ufpr.br/menotti/bovifocr/ **(R$ 1M - 3 anos)**

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

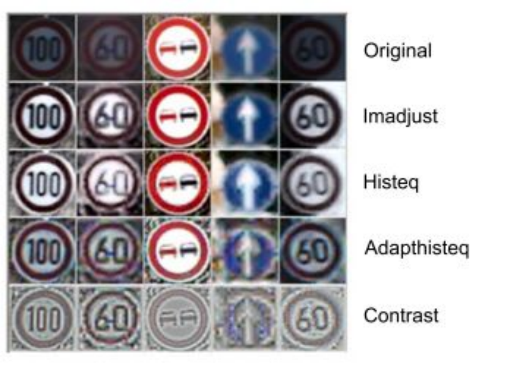
Bigshots - Machine Learning

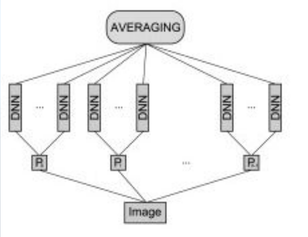
- **Yann Lecun - Facebbok / NYU**
  - "Pai" da Rede de Convolução
  - 1989 Trabalho de graduação (ESIEE-Paris)
  - 1998 - Convolução para Back-propagation
  - 1991 PostDoc (University of Toronto)
    - Hinton
  - 1996 AT&T Labs - Image Processing Head
  - 2003 NYU (*New York University*)
  - 2013 - Chief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

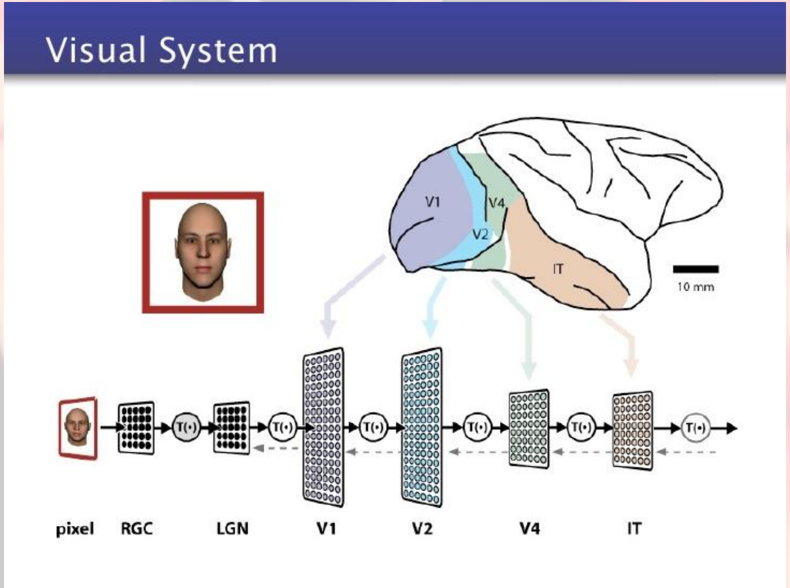
- **Andrew NG - (Stanford University)**
  - – Aprendizado não-supervisionado dos pesos
    - Clusters (k-means) originam pesos
      - – O próximo ***breakthrough*** está aqui ???
  - – Excelente curso MOOC

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**