

Reconhecimento de Faces e de Objetos com Python e Dlib

Jones Granatyr

Léo da Silva



Conteúdo do curso

- OpenCV e Dlib
- Instalação das ferramentas
- Detecção de faces
 - Haarcascades
 - HOG (histograma de gradientes orientados)
 - CNN (redes neurais convolucionais)
- Reconhecimento facial
 - Detecção de pontos faciais
 - Reconhecimento facial com Dlib (CNN + KNN)
 - Alinhamento de faces
- Detecção de objetos
 - imglab
 - Preditores de forma (definição de pontos de referência)
 - Detecção de objetos pela webcam

Pré-requisitos

- Lógica de programação (if e for)
- Conhecimentos básicos sobre Python
- Conhecimentos sobre detecção de faces e reconhecimento facial com o OpenCV são desejáveis
- Conhecimentos básicos sobre aprendizagem de máquina são desejáveis, bem como os algoritmos: SVM, KNN e redes neurais
- Nível do curso: intermediário

O que não veremos

- Detalhes muito técnicos sobre os algoritmos (haarcascade, hog, SVM, KNN, redes neurais convolucionais)
- Reconhecimento facial pela webcam
- Aplicativo comercial ou interface gráfica

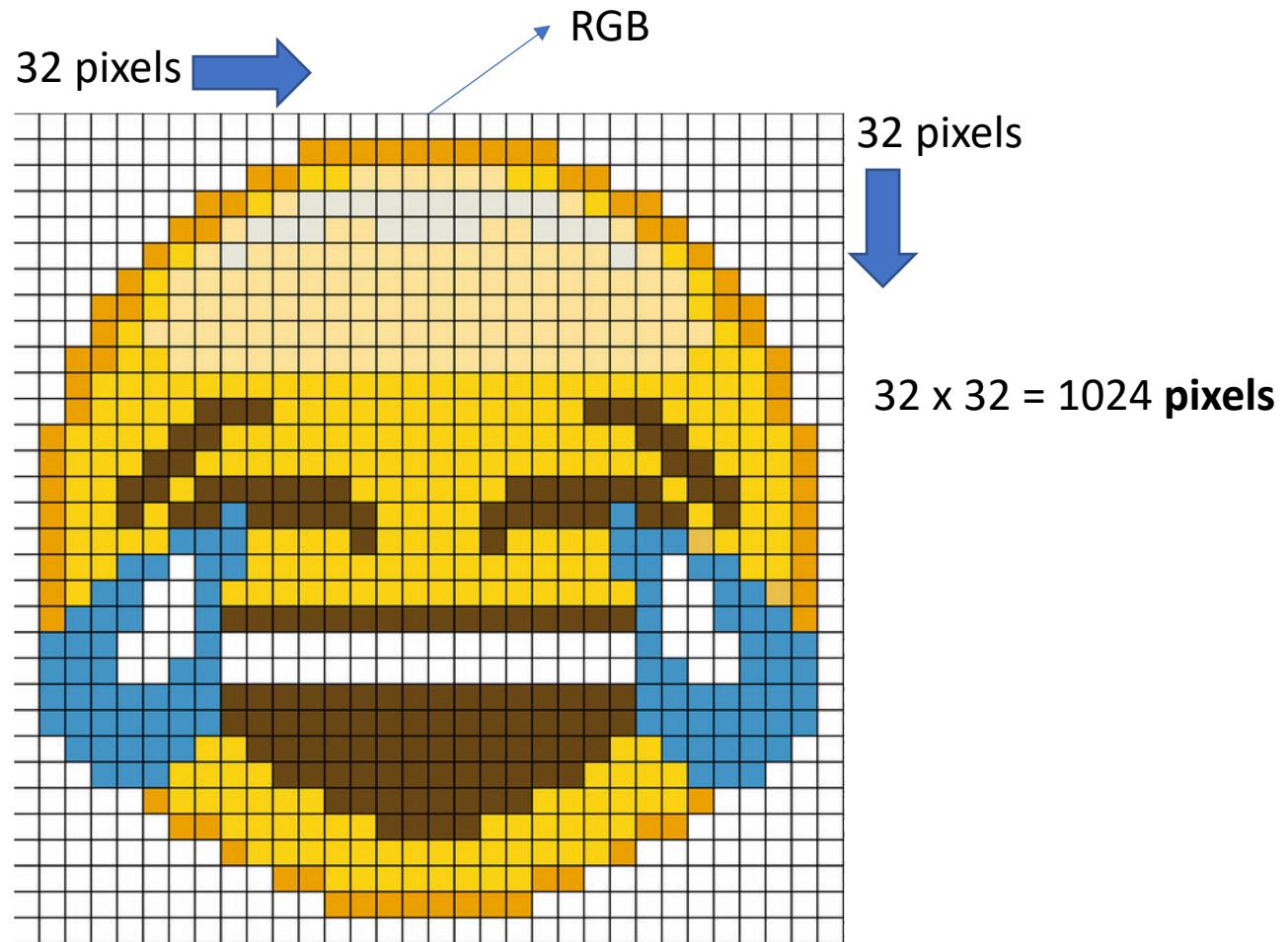
OpenCV

- Biblioteca mais popular para detecção de faces
- OpenCV = Open source computer vision
- Intel 1999
- Escrita em C/C++
- <https://opencv.org/>

Dlib

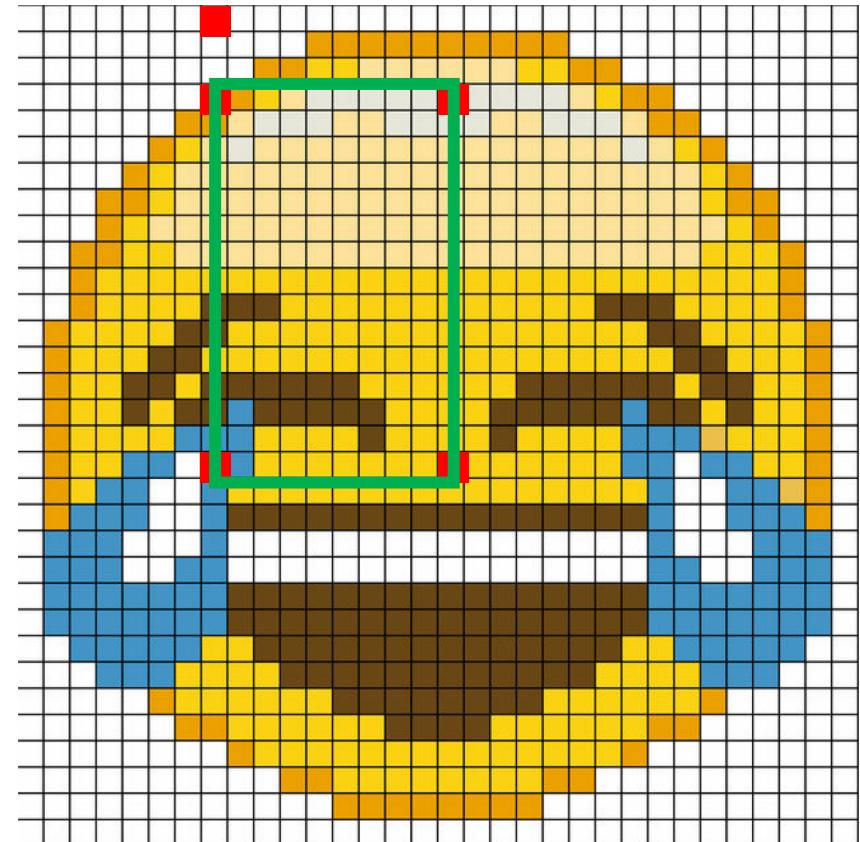
- Escrita em C++
- Criação de softwares complexos para resolver problemas do mundo real
- Aplicações como robótica, sistemas embarcados, smartphones (aplicações de larga escala)
- Documentação excelente
- Machine learning: classificação, regressão, agrupamento
- Máquinas de vetores de suporte (SVM) e deep learning (CNN)
- <http://dlib.net>
- http://dlib.net/ml_guide.svg

Pixels



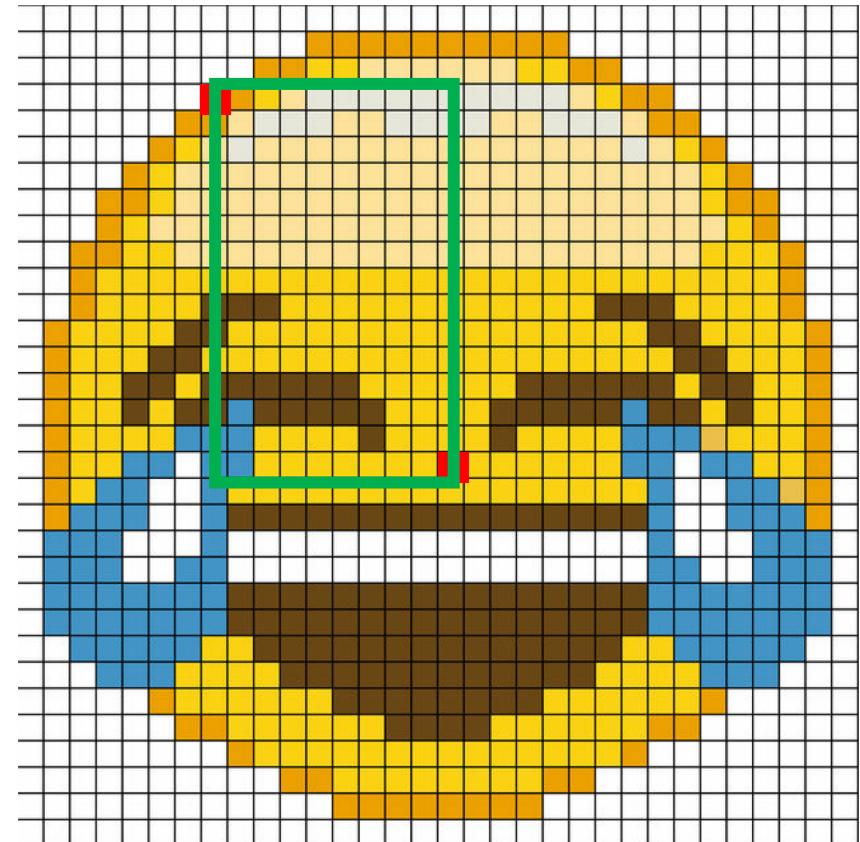
Bounding box (caixas delimitadoras)

- Haarcascade (OpenCV)
 - Left, top, width, height
 - 8, 4, 10, 15



Bounding box (caixas delimitadoras)

- HOG (Dlib)
 - Left, top, right, bottom
 - 8, 4, 17, 15



Técnicas para detecção de faces

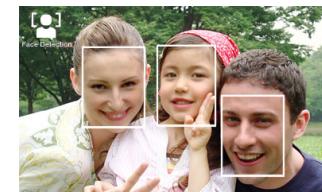
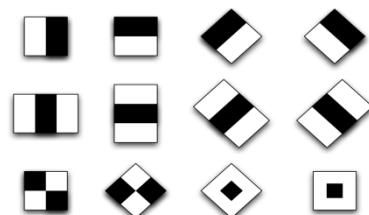
- Haarcascade
- HOG (histograma de gradientes orientados)
- CNN (redes neurais convolucionais)

Classificador Cascade



Treinamento
com AdaBoost

Seleção das
características



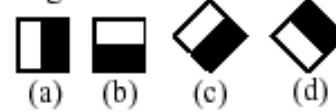
Componentes

- Haar cascades
- AdaBoost
- Cascade

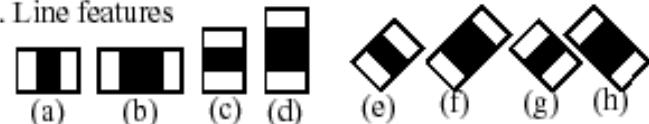
Haar cascades

- Combinação de features haar para formar um classificador
- Padrão retangular nos dados
- Diferenças na intensidade das regiões retangulares da imagem

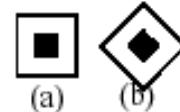
1. Edge features



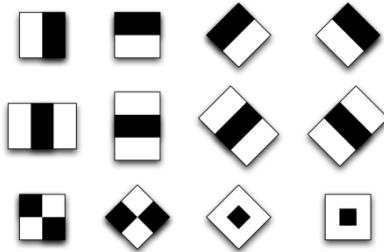
2. Line features



3. Center-surround features



Haar cascades



Soma pixels brancos –
soma pixels pretos

Mais de 160.000
combinações em uma
imagem 24 x 24!



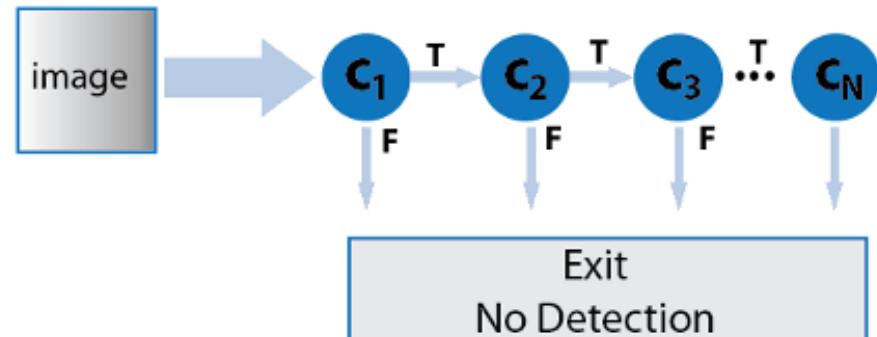


AdaBoost

- AdaBoost remove as características não necessárias
- Combina vários classificadores fracos em um classificador forte

Cascade

- Desliza pela imagem
- Computa a média dos valores dos pixels na área branca e preta
- Se a diferença entre as áreas é abaixo de um limiar, a característica coincide (match)
- Aprendizagem supervisionada

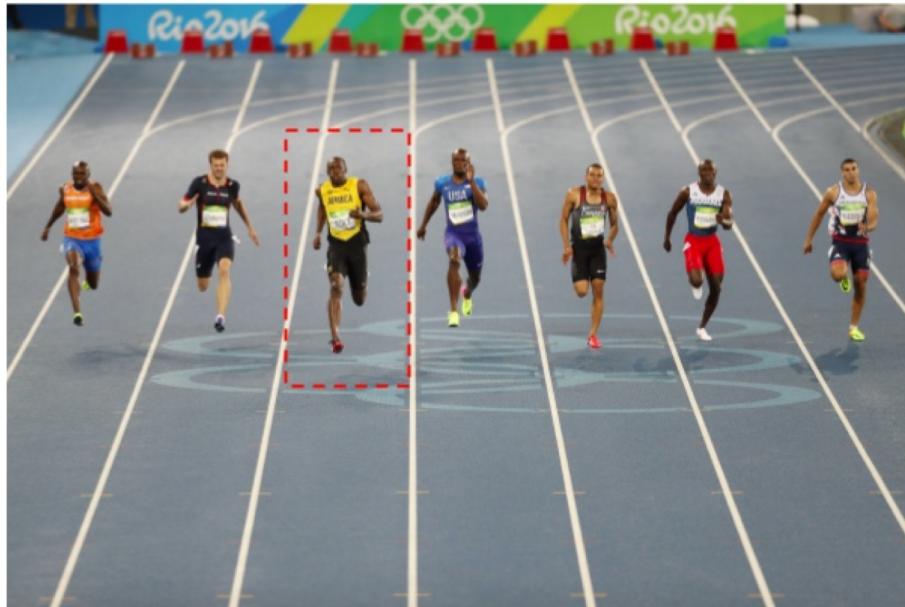


HOG

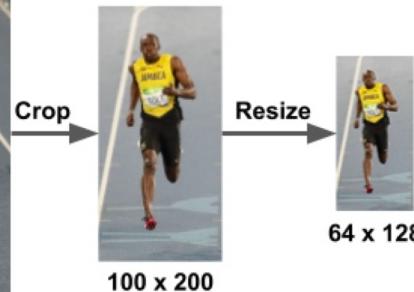


- Descritor de características: forma que o computador usa para selecionar as partes principais de uma imagem
- Descritor de forma
- Outros descritores: cor, textura, haar
- HOG: forma + textura

HOG



Original Image : 720 x 475



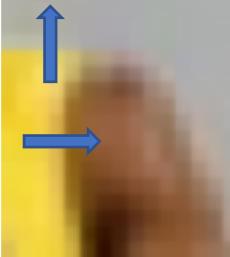
Derivada permite medir a taxa de variação (derivada zero, derivada pequena e derivada alta)

Vetor gradiente (direção e sentido que os valores aumentam)

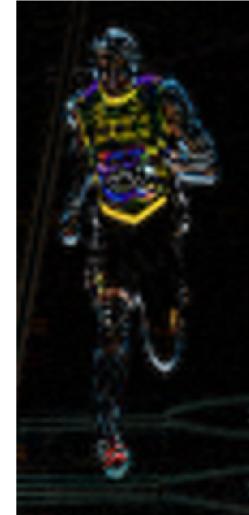
Fonte: <https://www.learnopencv.com/histogram-of-oriented-gradients/>



HOG

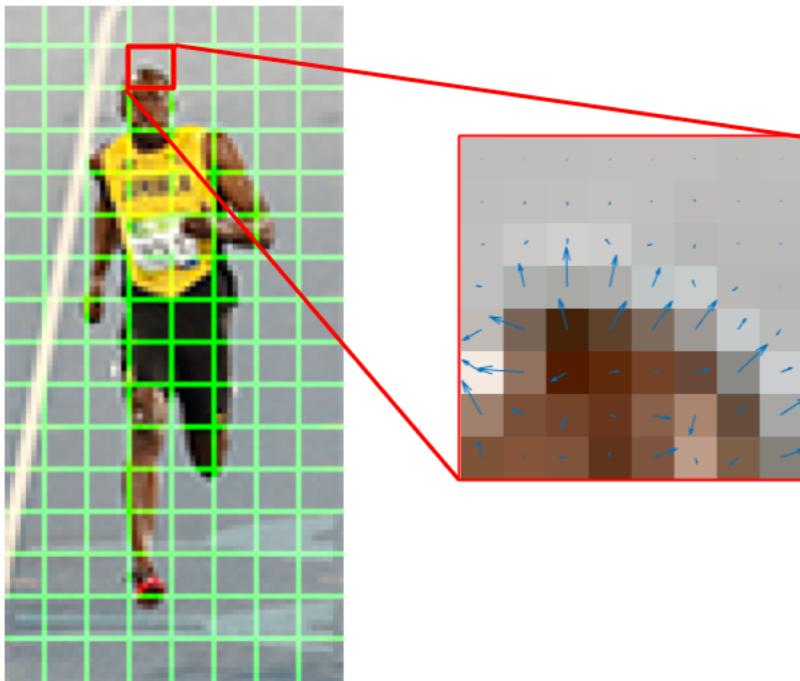


- Variação da cor de um pixel para seu vizinho (gradiente)
 - Amarelo e amarelo a derivada é zero
 - Amarelo e cinza a derivada pode ser pequena
 - Amarelo e marrom a derivada pode ser maior
- Onde não tem variação fica preto e onde tem variação varia “liga-se” um pixel (contorno ou forma)



HOG

- Indicar para qual vizinho o pixel variou e com qual intensidade



2	3	4	4	3	4	2	2
5	11	17	13	7	9	3	4
11	21	23	27	22	17	4	6
23	99	165	135	85	32	26	2
91	155	133	136	144	152	57	28
98	196	76	38	26	60	170	51
165	60	60	27	77	85	43	136
71	13	34	23	108	27	48	110

Gradient Magnitude

80	36	5	10	0	64	90	73
37	9	9	179	78	27	169	166
87	136	173	39	102	163	152	176
76	13	1	168	159	22	125	143
120	70	14	150	145	144	145	143
58	86	119	98	100	101	133	113
30	65	157	75	78	165	145	124
11	170	91	4	110	17	133	110

Gradient Direction

HOG

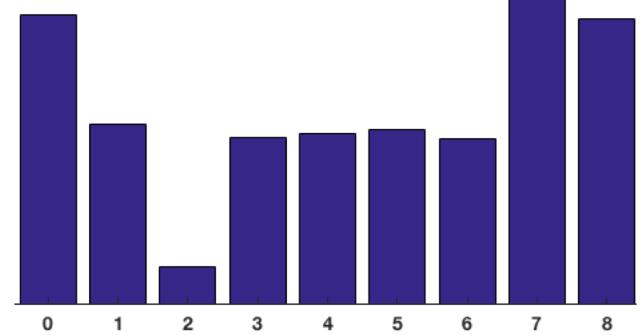
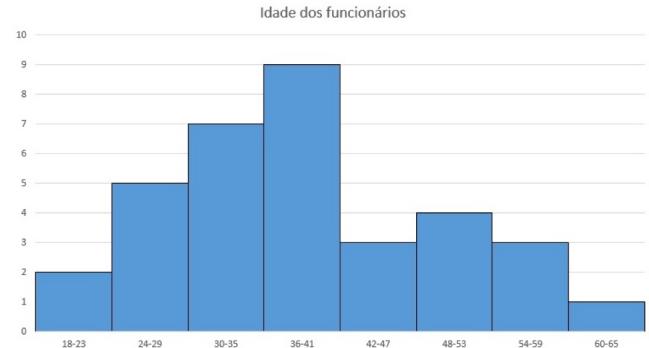
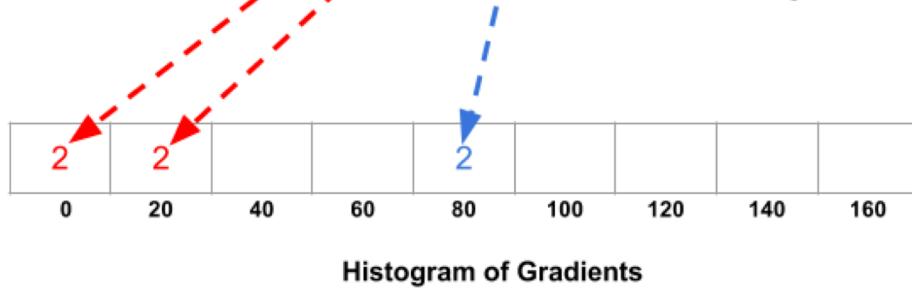
- Histograma

80	36	5	10	0	64	90	73
37	9	9	179	78	27	169	166
87	136	173	39	102	163	152	176
76	13	1	168	159	22	125	143
120	70	14	150	145	144	145	143
58	86	119	98	100	101	133	113
30	65	157	75	78	165	145	124
11	170	91	4	110	17	133	110

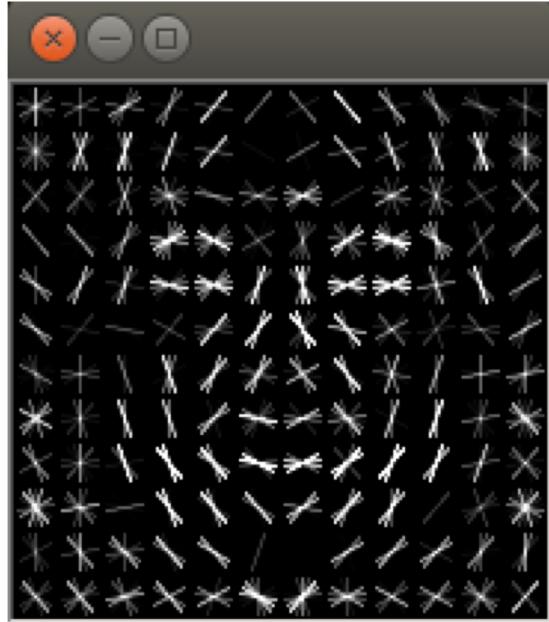
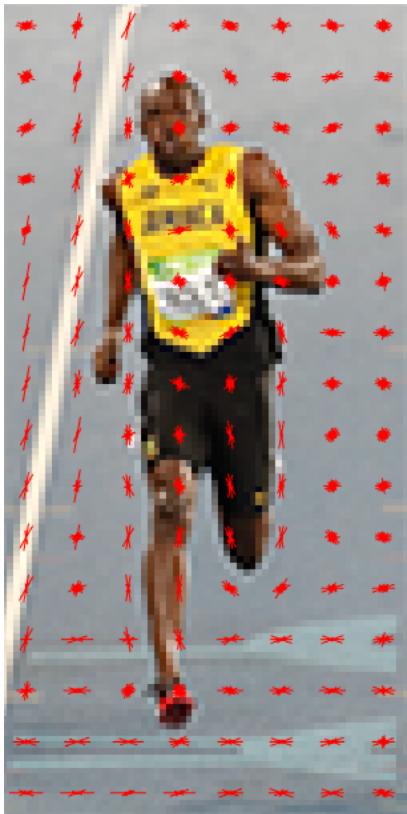
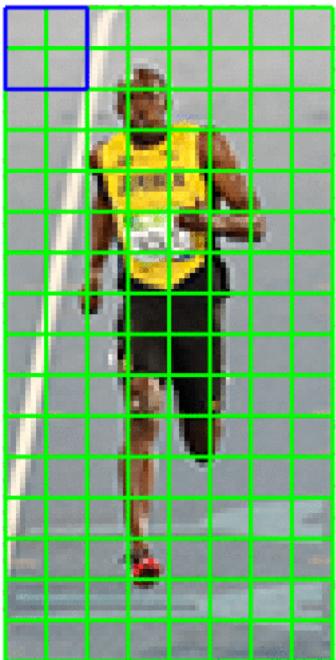
Gradient Direction

2	3	4	4	3	4	2	2
5	11	17	13	7	9	3	4
11	22	23	27	22	17	4	6
23	99	165	135	85	32	26	2
91	155	133	136	144	152	57	28
98	196	76	38	26	60	170	51
165	60	60	27	77	85	43	136
71	13	34	23	108	27	48	110

Gradient Magnitude



HOG

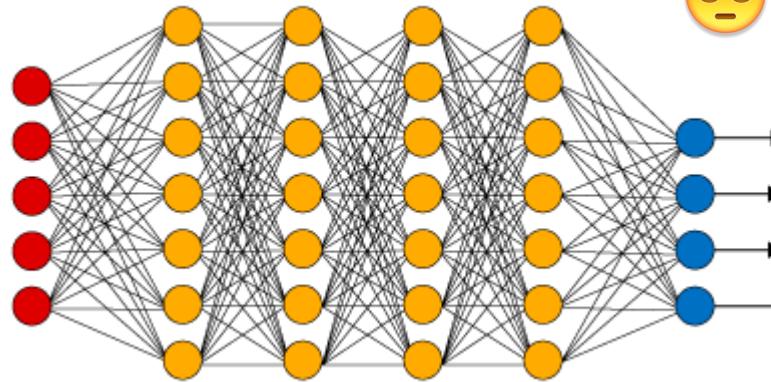
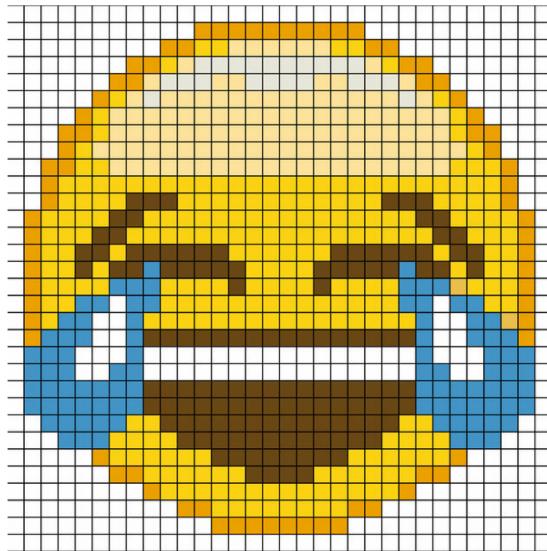


Redes neurais convolucionais (CNN)

- Usado para visão computacional
- Carros autônomos, detecção de pedestres (umas das razões por deep learning funcionar bem)
- Em geral, melhor do que SVM

Redes neurais densas x convolucionais

$32 \times 32 = 1.024 \times 3 = \mathbf{3.072}$ entradas

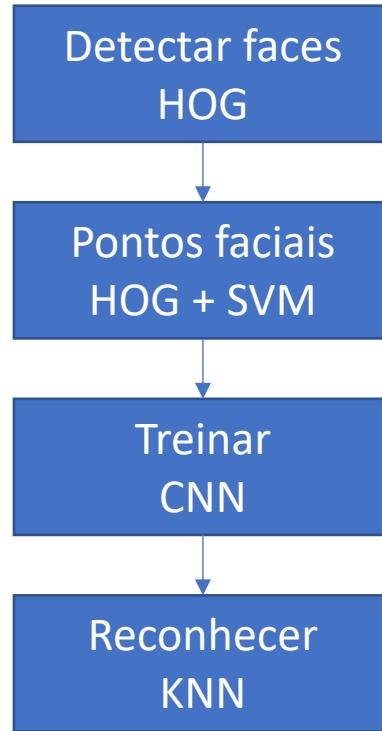


- Não usa todas as entradas (pixels)
- Usa uma rede neural tradicional, mas no começo transforma os dados na camada de entrada
- Quais são as características mais importantes?

Redes neurais convolucionais (CNN)

- Quais características utilizar?
- Para faces
 - Localização do nariz
 - Distância entre os olhos
 - Localização da boca
- Como diferenciar uma face humana de um animal?
- CNN descobre as características

Etapas para o reconhecimento facial



Reconhecimento facial OpenCV

- EigenFaces
- FisherFaces
- LBPH

Detecção de pontos de referência facial



- Acha os pontos dentro da bouding box (ROI – region of interest)
- Tentam localizar as seguintes regiões na face: queixo, lábio externo, lábio interno, sobrancelha direita, sobrancelha esquerda, olho direito, olho esquerdo e nariz
- Aplicações
 - Estimativa da pose da cabeça
 - Face morphing (mudança gradual de uma face para outra, face média, troca de face, detecção de piscada e sonolência de motorista, filtros de rosto)

Conclusão

