THE DEVELOPER'S CONFERENCE

Reconhecimento facial para controle de acesso em tempo real

Victor Herzog Damke
Desenvolvedor Full-Stack
CWI Software

Olá

Victor Herzog Damke
Desenvolvedor Full-Stack
Núcleo de Tecnologia
CWI Software

Engenharia de Controle e Automação
Budapest University of Technology and Economics

Agenda

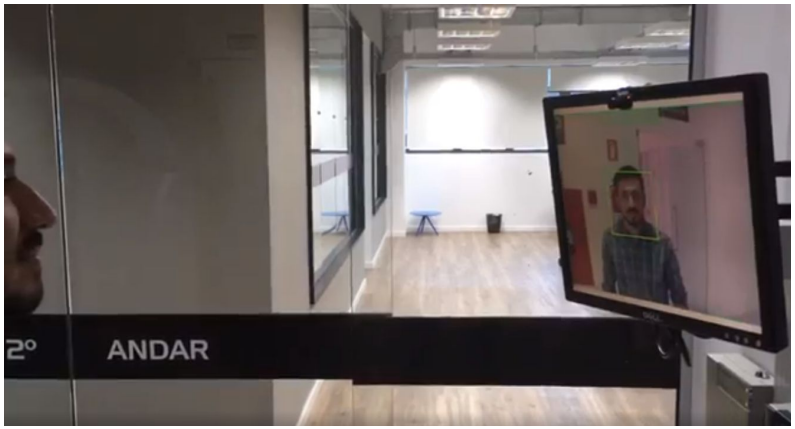
THE DEVELOPER'S CONFERENCE

- Objetivos
- Detecção vs Reconhecimento
- Detecção Facial
- Reconhecimento Facial
- Solução
- Resultados

Objetivos

THE DEVELOPER'S CONFERENCE

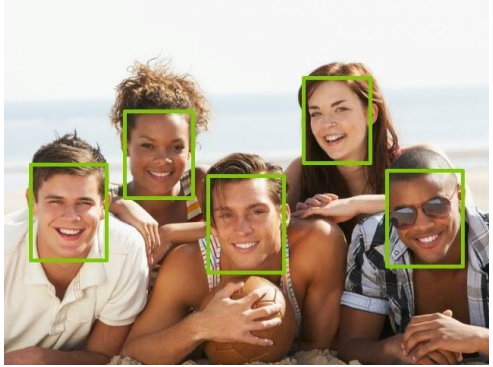
- Controle de acesso
  - Colaboradores da CWI;
  - Pessoas cadastradas previamente (clientes, visitantes, equipe de manutenção);
  - Limitação de dias, faixa de horas e andares;



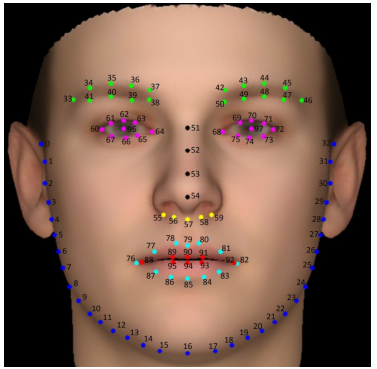
Detecção vs Reconhecimento

THE DEVELOPER'S CONFERENCE

- Detecção Facial
  - Encontrar rostos em uma imagem.
  - Ex.: Autofoco, contagem de pessoas...



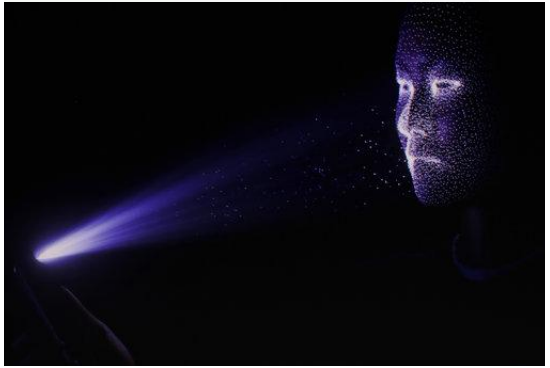
- Reconhecimento Facial
  - Analisar as características de um rosto e compará-las com uma coleção.
  - Identificar ou validar

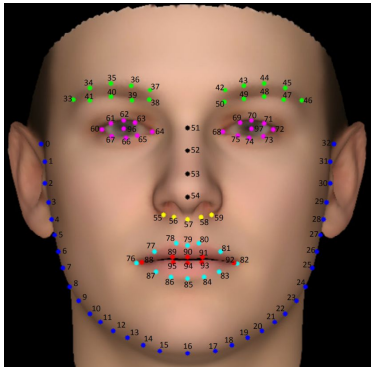


Detecção vs Reconhecimento

THE DEVELOPER'S CONFERENCE

- Reconhecimento 2d e 3d
  - Matriz de pontos infravermelhos.
  - Hardware especial.
- Reconhecimento Facial
  - Analisar as características de um rosto e compará-las com uma coleção.
  - Identificar ou validar





Detecção Facial

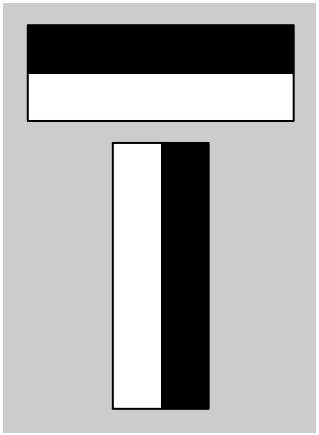
Viola-Jones

Haar features

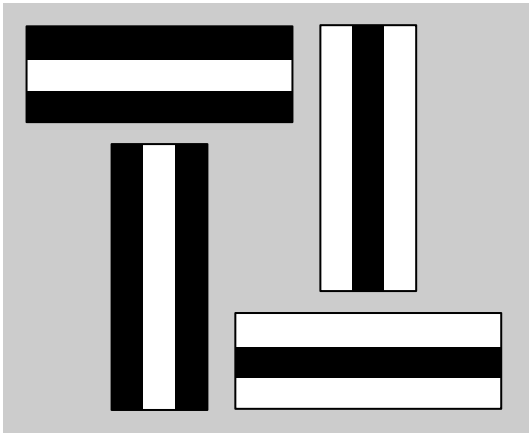
características de Haar

THE DEVELOPER'S CONFERENCE

Edge features



Line features



Four rectangle feature

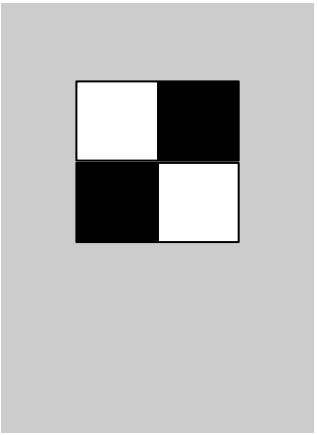
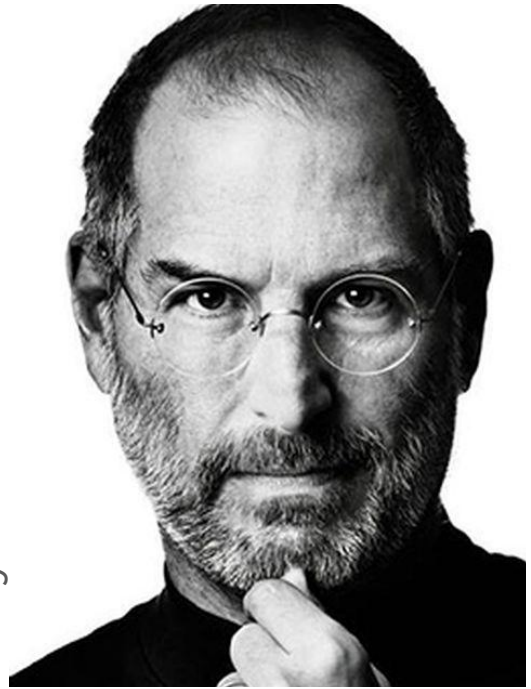


Imagem Integral

THE DEVELOPER'S CONFERENCE

Grayscale

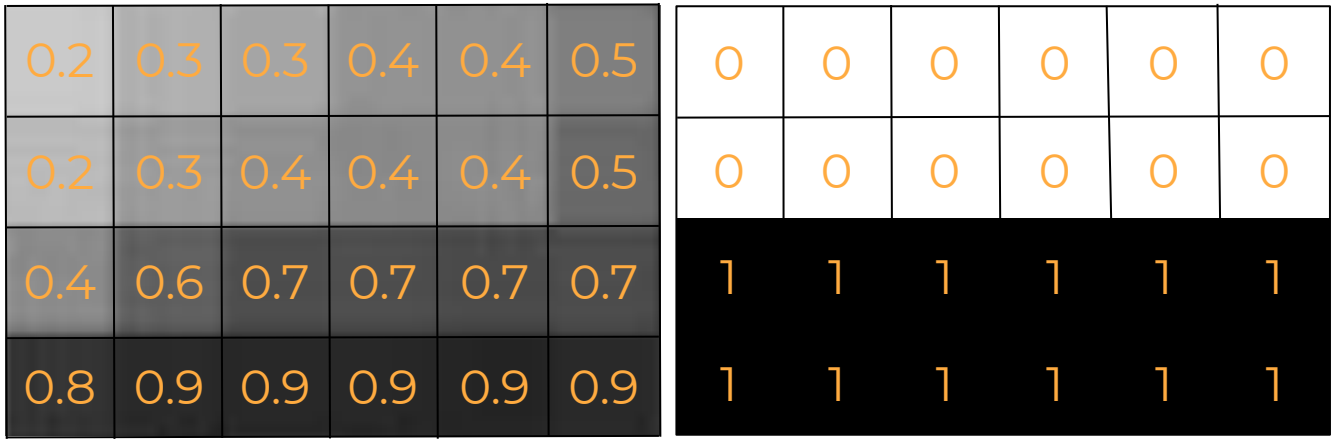


Integral

Viola-Jones

+ 6.000 features

Haar features



Diferença entre as intensidades:

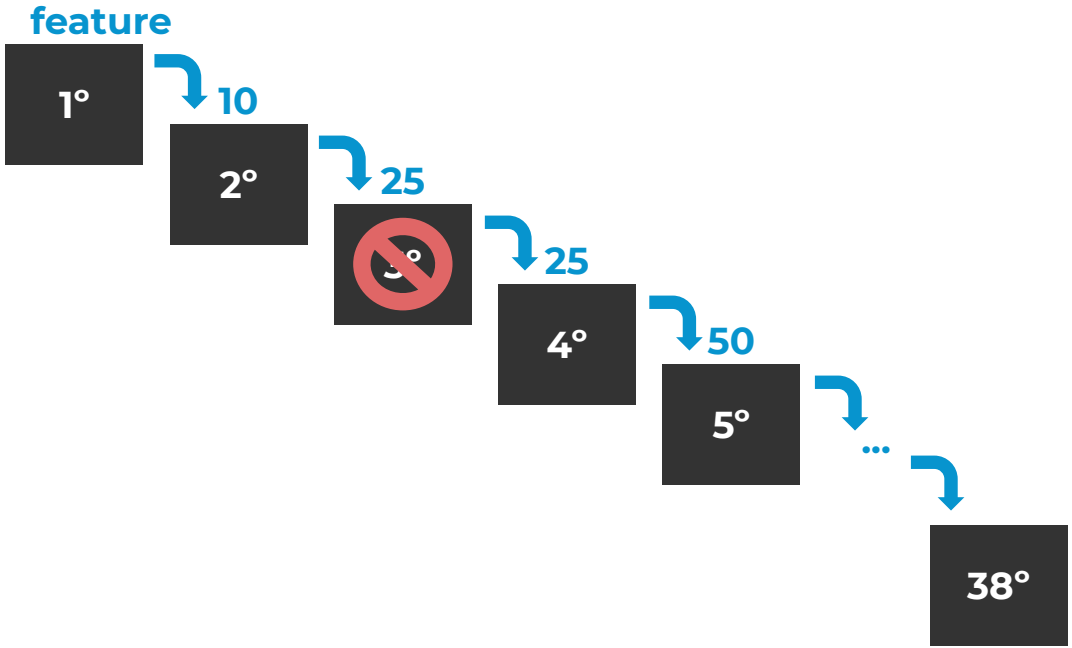
$$\Delta = dark - white = \frac{1}{n}\sum_{dark}^{n} I(x) - \frac{1}{n}\sum_{white}^{n} I(x)$$

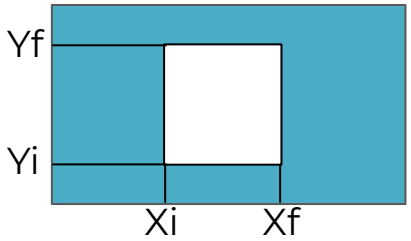
$\triangle$ = (9,1 / 12) - (4,3 / 12)

$\triangle$ = 0,7583 - 0,3583

$\triangle$ = **0,4**

Cascata de classificadores





[[0, 0, 150, 150], [100, 200, 300, 500]]

Reconhecimento Facial

THE DEVELOPER'S CONFERENCE

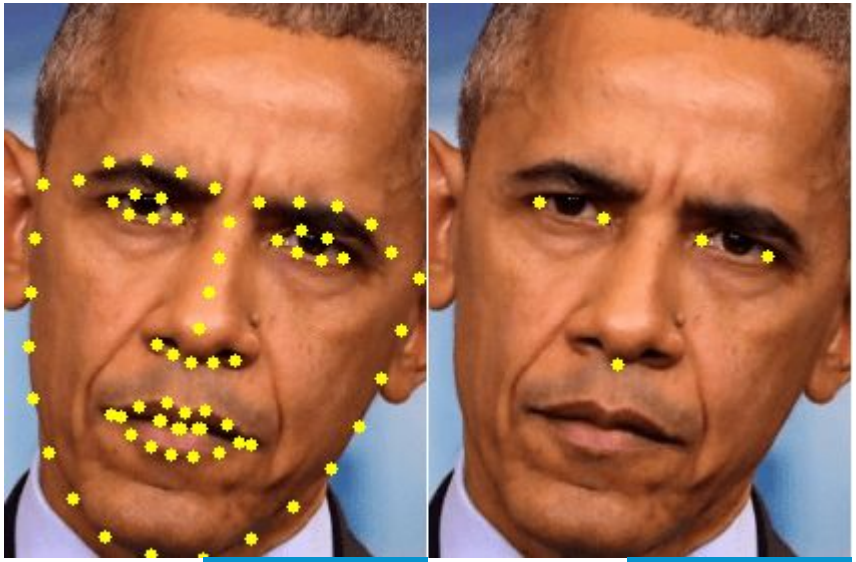
- Face Landmarks
  - Diferença de intensidade dos pixels.
  - Ex.: Reconhecimento facial, animação, reconhecimento de emoções, fotografia...
- Encodings
- Comparação

68 pontos

Face Landmarks

THE DEVELOPER'S CONFERENCE

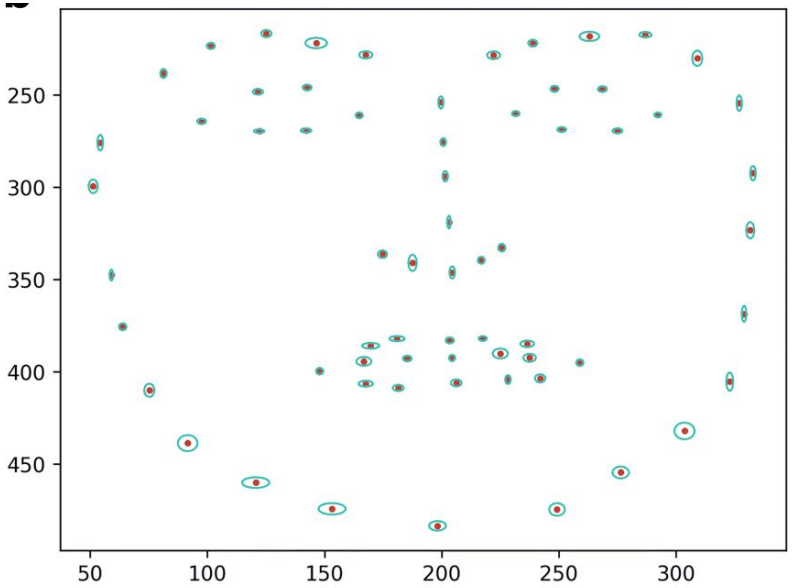
modelos



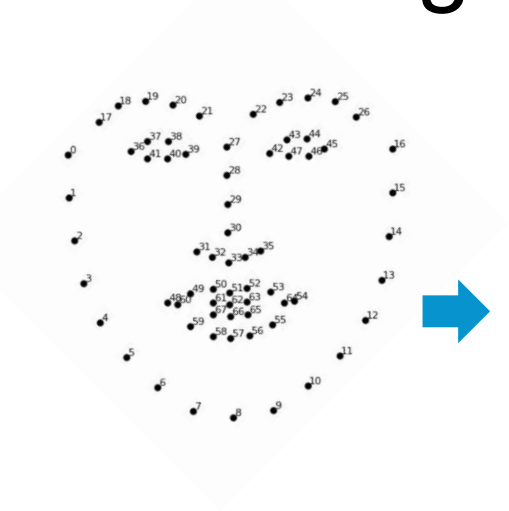
68 pts

5 pts

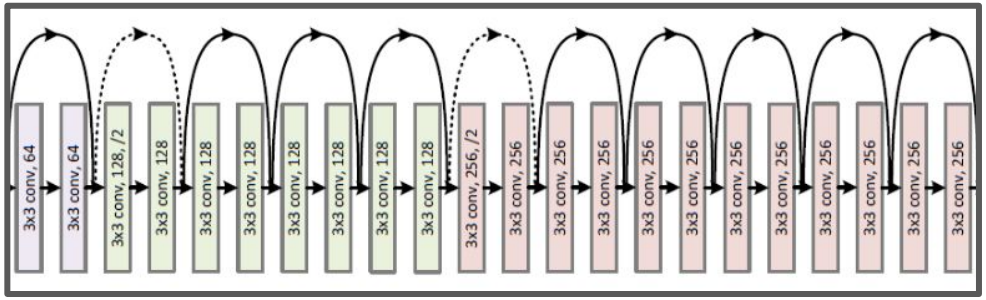
desvio



Encodings



Residual Network (ResNet)



[-0.0738273933529854,
0.00889372732490301,
-0.00961793214082718,
-0.0524267368018627,
-0.0772915706038475,
-0.017961286008358,
-0.0307803973555565,
-0.152948439121246,
...
0.0832497179508209,
-0.0913219675421715,
-0.258715033531189,
-0.0655047446489334,
0.0492074079811573,
-0.0975103974342346,
0.00199389457702637,
0.039286844432354]

3 milhões de rostos

vários datasets

29 camadas de convoluções

Comparação

[-0.0738273933529854, 0.00889372732490301, -0.00961793214082718, ... 0.0492074079811157, -0.097510397434234, 0.00199389457702631]

[-0.0695742517709732, 0.0315547659993172, -0.043367862701416, ... -0.0153855439275503, 0.0754944831132889, -0.0277792438864708]

[-0.0989925637841225, 0.0233747325837612, 0.028356870636344, ... -0.00901452358812094, 0.0234573371708393, 0.0483754687011242]

[-0.0759359374642372, 0.0590272210538387, -0.00541636766865849, ... -0.0401360727846622, 0.0457880757749081, 0.0317194014787674]

[-0.147801905870438, 0.0267262309789658, 0.0401979647576809, ... 0.102912932634354, 0.0626524537801743, 0.0218381956219673]

[-0.0943850129842758, 0.0535568334162235, 0.0536609143018723, ... -0.00924267619848251, 0.0557434819638729, -0.0243553519248962]

[-0.105790950357914,0.0607326999306679,0.00598054332658648, ... -0.03774369135499, 0.104004122316837, -0.0120019689202309]

[-0.0902278944849968, 0.0412012711167336, 0.059528473764658, ... -0.0316068157553673, 0.105147130787373, 0.0557747632265091]

Comparação

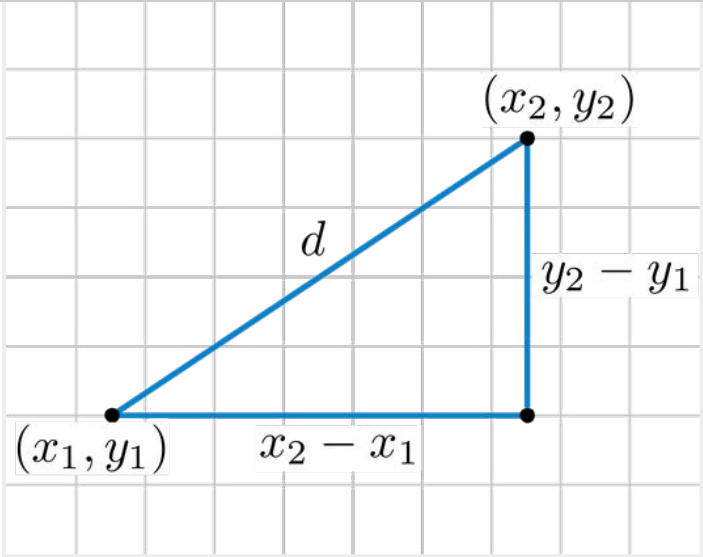
THE DEVELOPER'S CONFERENCE

[-0.0738273933529854, 0.00889372732490301, -0.00961793214082718, ... 0.049207407981157, -0.097510397434234, 0.0019938945770263]

[-0.0695742517709732, 0.0315547659993172, -0.043367862701416, ... -0.0153855439275503, 0.0754944831132889, -0.0277792438864708]

distância euclidiana

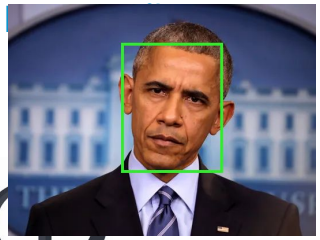
$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$



Solução

Câmera

Detecta rosto



Reconhecimento

Landmarks

Encodings

[0.067, 0.039, … , 0.002, -0.061]

Banco de Dados

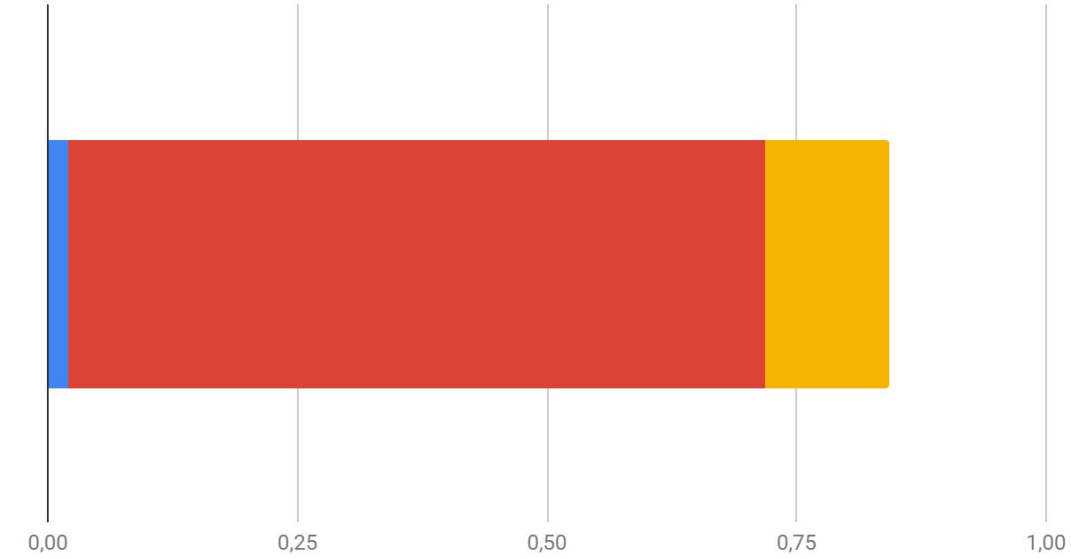
[0.012, -0.043, … , 0.032, 0.076]	0,35
[-0.056, 0.092, … , 0.021, 0.073]	0,22
[0.086, 0.015, … , -0.097, 0.035]	0,27
[0.068, 0.037, … , 0.002, -0.062]	0,18

Distância

0,18 < DIST_MIN

AUTORIZADO

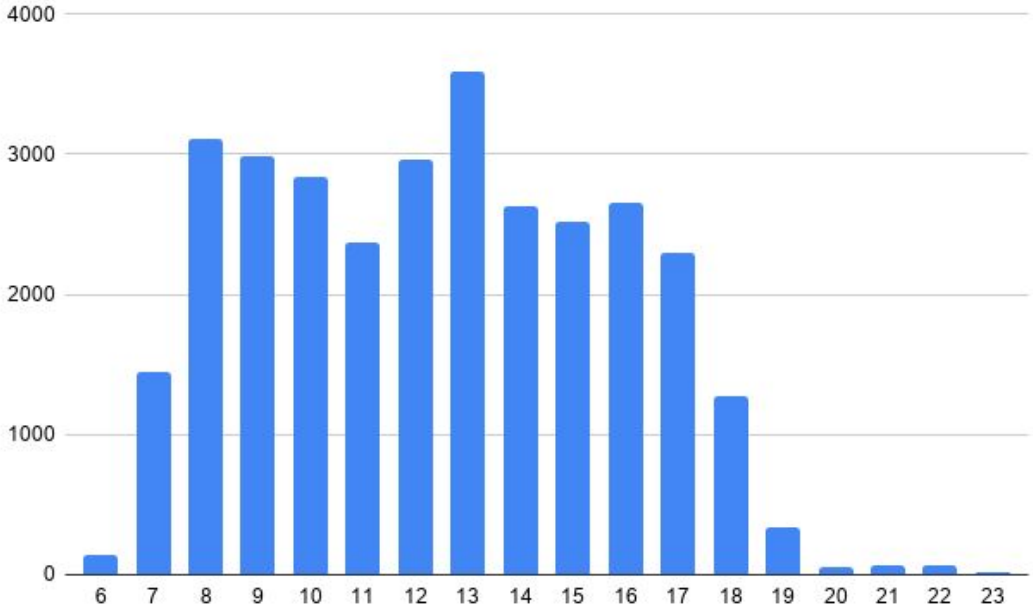
Resultados



	v1	v2	
Detecção	0,36	0,02s	- 0,34
Reconhecimento	1,19	0,70s	- 0,49
Banco de Dados	0,08	0,12s	+ 0,04
Total	**1,63s**	**0,85s**	

0.78245s	mínimo
0.84916s	**médio**
1.06827s	máximo

Resultados



29.621	entradas
221	usuários ativos
230	usuários cadastrados
0,183	média da distância nas entradas

THE
DEVELOPER'S
CONFERENCE