

E-MOTIV

Protótipo de Sistema para Reconhecimento de Expressões Faciais

Aluno: William Leander Seefeld

Orientador: Daniel Theisges dos Santos

Roteiro

- Introdução
- Objetivos
- Fundamentação Teórica
- Trabalhos Correlatos
- Requisitos e Especificação
- Implementação
- Operacionalidade
- Resultados e Discussões
- Conclusões e Sugestões



Introdução

Introdução

- Computação cada vez mais permeável e presente; ubíqua
- Necessidade de tornar a computação mais humana
- Emoções são características fortes de humanos
- Se manifestam através de palavras, tom de voz, **expressões faciais**, comportamentos, entre outros
- Por que não dotar computadores desta capacidade?

Introdução

- Objetivo da Computação Afetiva
- Forma alternativa de entrada de dados
- Algumas expressões faciais de emoções são distinguíveis e universais
- Técnicas de Visão Computacional (VC) permitem que computadores “vejam”
- Técnicas de Inteligência Artificial (IA) permitem que computadores “entendam”

Objetivos

Desenvolver um sistema capaz de reconhecer expressões faciais de emoção em imagens estáticas, usando técnicas de Visão Computacional e Inteligência Artificial

- a) utilizar técnicas de VC para extração de características faciais
- b) implementar um classificador de emoções utilizando uma Rede Neural Artificial
- c) reconhecer as expressões faciais das emoções neutra, felicidade e tristeza



Fundamentação Teórica

Fundamentação Teórica

Expressões Faciais Universais

- Seis expressões faciais de emoção universais (EKMAN; FRIESEN, 1969)
 - Manifestadas através de músculos faciais específicos
 - Codificadas através do sistema FACS (Sistema de Codificação da Ação Facial)

Fundamentação Teórica

Expressões Faciais Universais

- Seis expressões faciais de emoção universais (EKMAN; FRIESEN, 1969)
- **Aversão**



Fundamentação Teórica

Expressões Faciais Universais

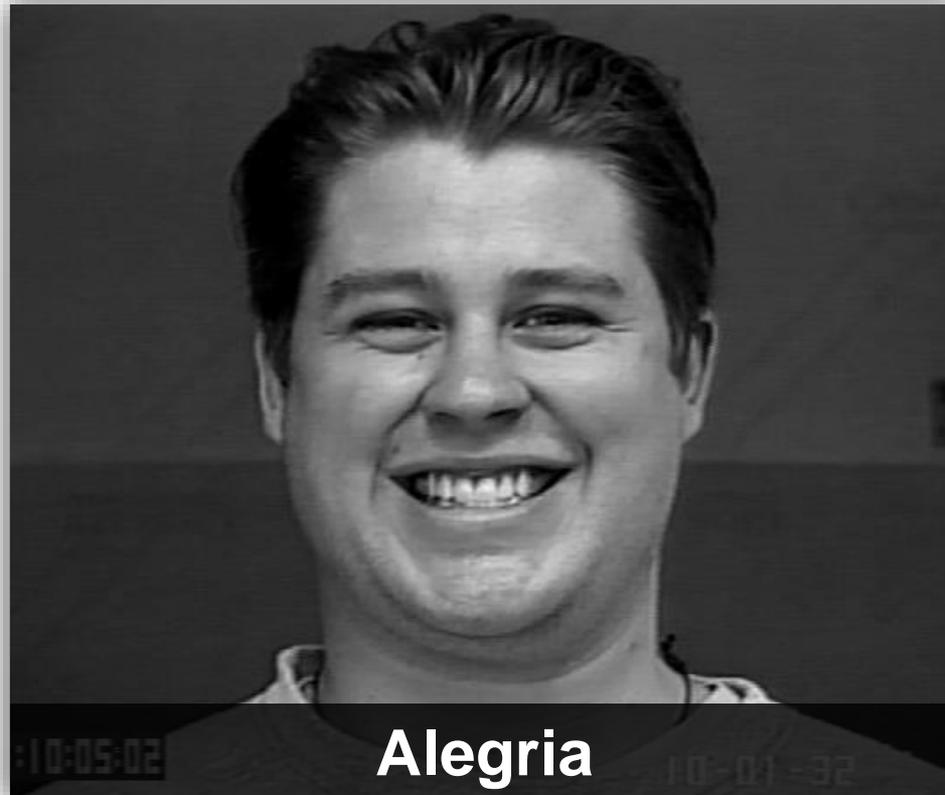
- Seis expressões faciais de emoção universais (EKMAN; FRIESEN, 1969)
 - Aversão
 - **Medo**



Fundamentação Teórica

Expressões Faciais Universais

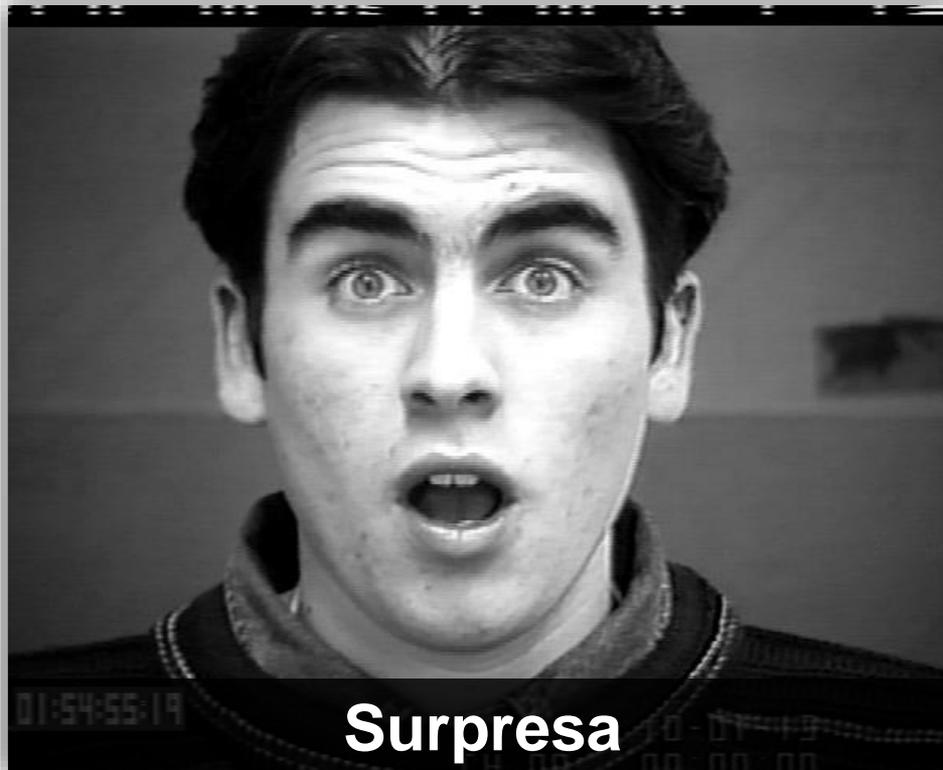
- Seis expressões faciais de emoção universais (EKMAN; FRIESEN, 1969)
 - Aversão
 - Medo
 - **Alegria**



Fundamentação Teórica

Expressões Faciais Universais

- Seis expressões faciais de emoção universais (EKMAN; FRIESEN, 1969)
 - Aversão
 - Medo
 - Alegria
 - **Surpresa**



Fundamentação Teórica

Expressões Faciais Universais

- Seis expressões faciais de emoção universais (EKMAN; FRIESEN, 1969)

- Aversão
- Medo
- Alegria
- Surpresa
- **Tristeza**

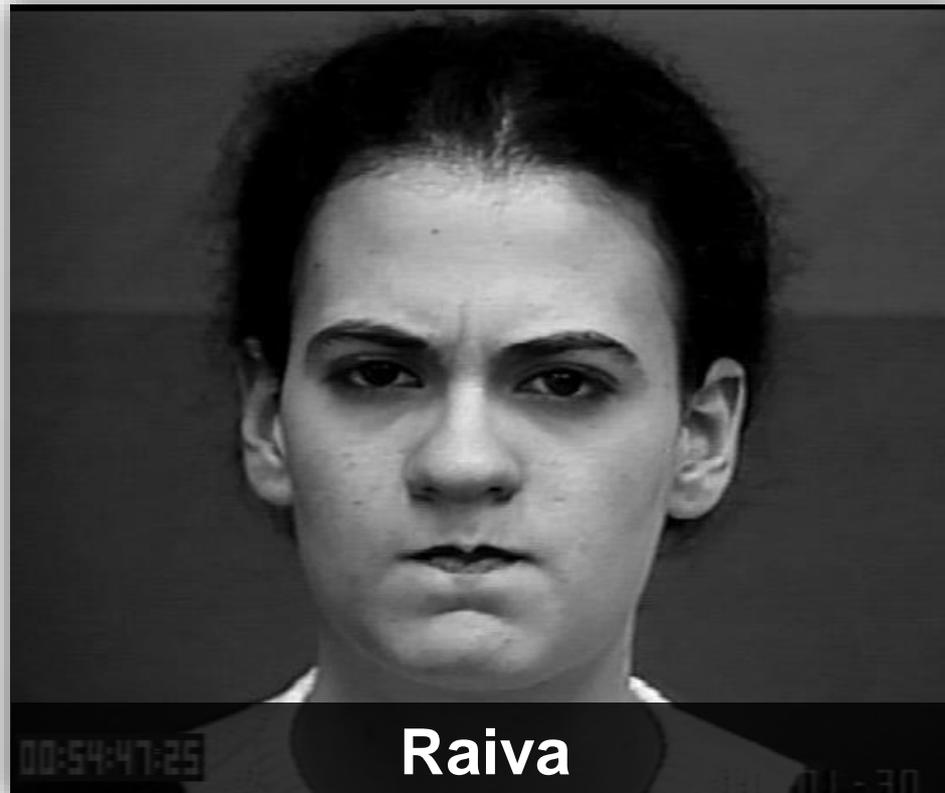


Fundamentação Teórica

Expressões Faciais Universais

- Seis expressões faciais de emoção universais (EKMAN; FRIESEN, 1969)

- Aversão
- Medo
- Alegria
- Surpresa
- Tristeza
- **Raiva**



Fundamentação Teórica

Expressões Faciais Universais

- Seis expressões faciais de emoção universais (EKMAN; FRIESEN, 1969)

- Aversão
- Medo
- Alegria
- Surpresa
- Tristeza
- Raiva
- **Neutra**



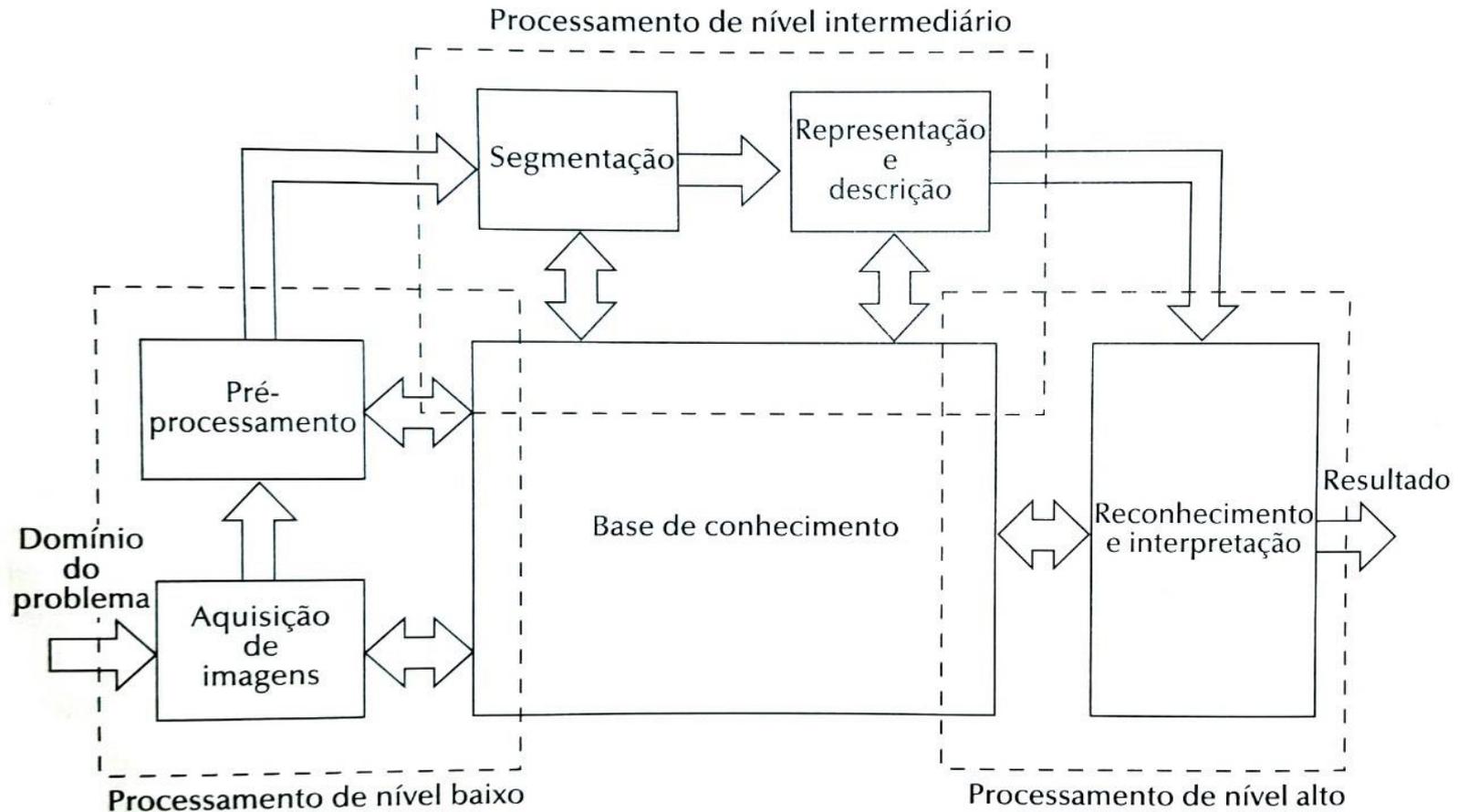
Fundamentação Teórica

Visão Computacional

- Interpretação de informações visuais por computadores
 - Reconhecimento facial
 - Identificação de sinais de trânsito
 - Reconhecimento de caracteres
- Interpretação de cenas e tomada de decisão
 - Liberar acessos
 - Mudar a direção de um veículo
 - Processar pagamentos

Fundamentação Teórica

Visão Computacional



Fonte: Gonzalez e Woods (2000, p. 408)

Fundamentação Teórica

Processamento de Imagens Digitais

- Melhoria de imagens:
 - Para humanos
 - Para computadores
- Operações de manipulação de *pixels*
- Extração de características relevantes
- Reconhecimento e Interpretação

Fundamentação Teórica

Aprendizagem de Máquina

- Técnica de Inteligência Artificial
 - Área que visa dotar máquinas de inteligência similar a de seres humanos
- Permite ao computador aprender
- Aprendizado pode ser supervisionado, não-supervisionado, por reforço, entre outras abordagens

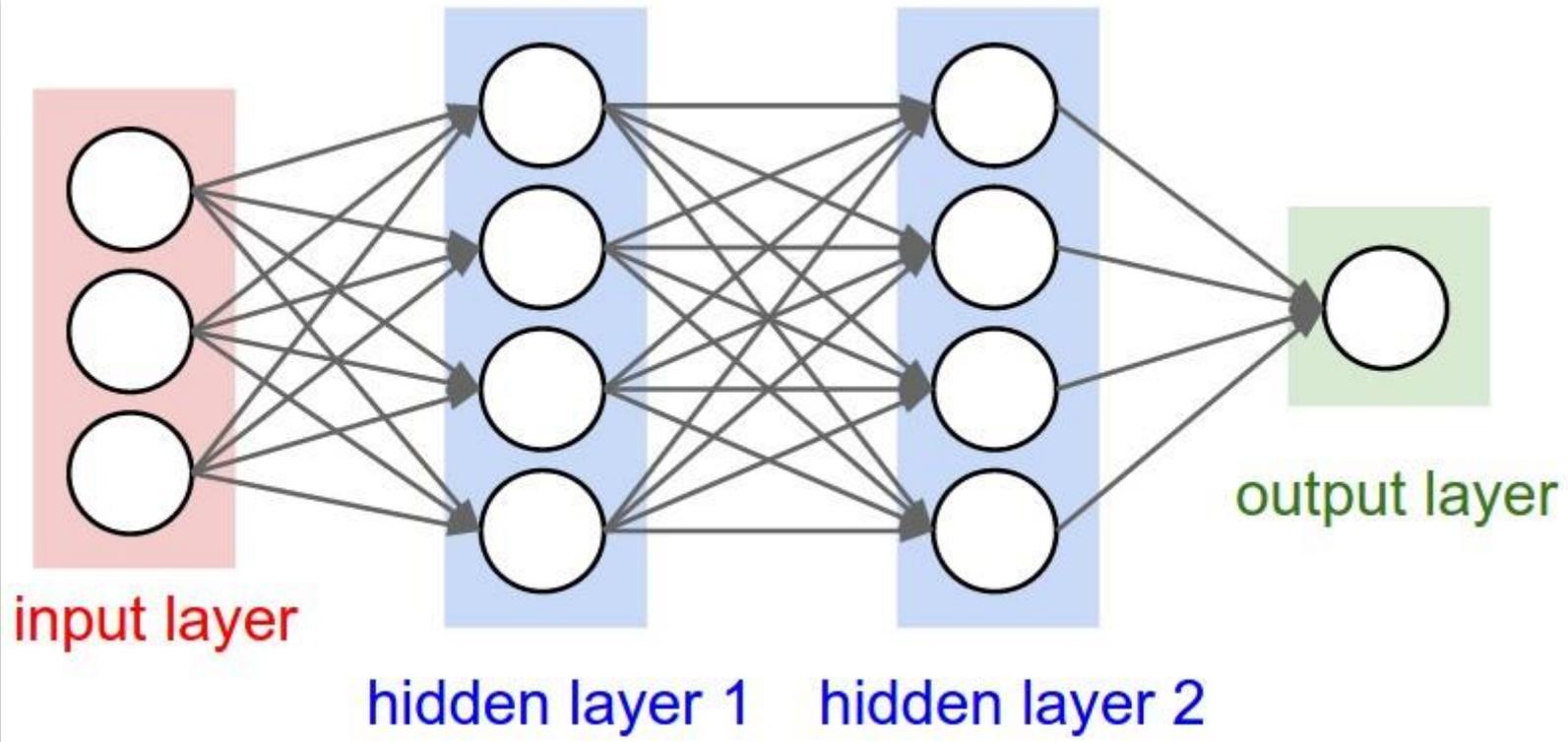
Fundamentação Teórica

Redes Neurais Artificiais

- Segundo Marques (2005, p. 163):
“as redes neurais artificiais são constituídas por elementos simples interligados, com capacidade de aprendizagem a partir dos dados”
- Um elemento simples pode ser um *perceptron*, tratado como um neurônio
- Redes *perceptron* multicamadas são possíveis com retropropagação de erro

Fundamentação Teórica

Redes Neurais Artificiais



Fonte: Stanford (2016)



Trabalhos Correlatos

Trabalhos Correlatos

Oliveira e Jaques (2013)

Classificação de emoções básicas através de imagens capturadas em vídeos de baixa resolução

- Reconhecimento em vídeo
- Vetor de Pontos Extremos
- Rede Neural – taxa de acerto de 89,9%

Trabalhos Correlatos

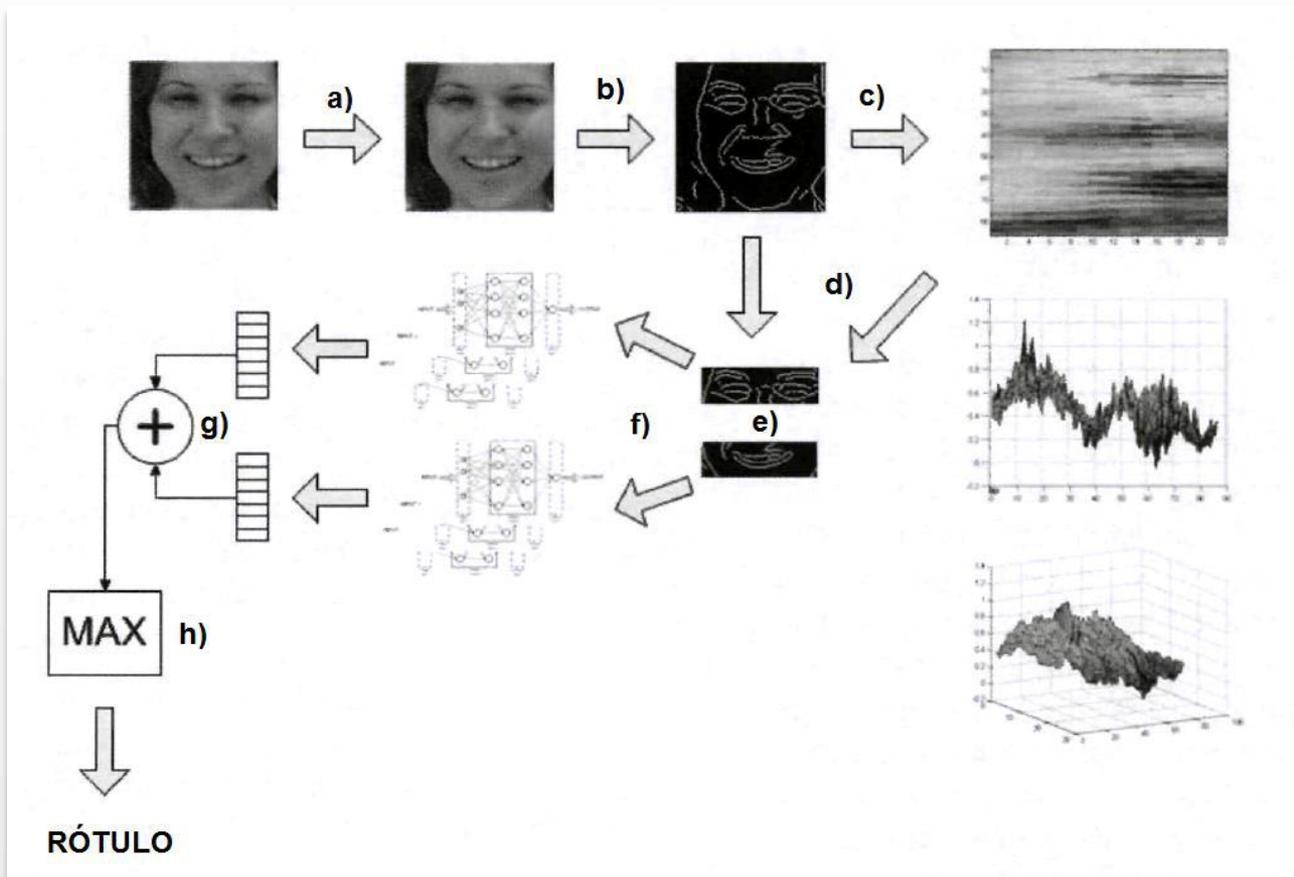
Filko e Martinović (2013)

Emotion recognition system by a neural network based facial expression analysis

- Imagens estáticas
- *Principal Component Analysis e Eigenfaces*
- Conjunto de Redes Neurais – acerto médio de 70%

Trabalhos Correlatos

Filko e Martinović (2013)



Fonte: Adaptado de Filko e Martinović (2013)

Trabalhos Correlatos

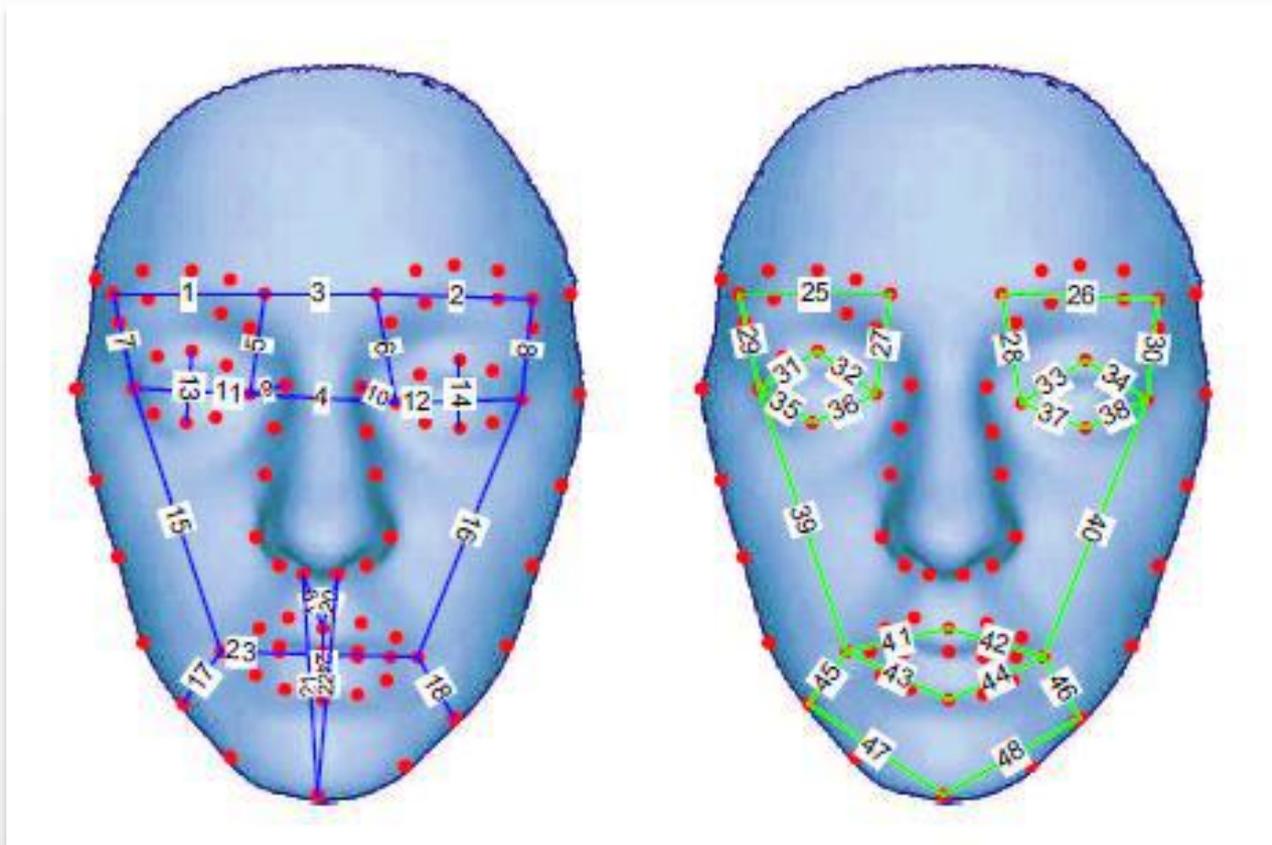
Tang e Huang (2008)

3D facial expression recognition based on properties of line segments connecting facial feature points

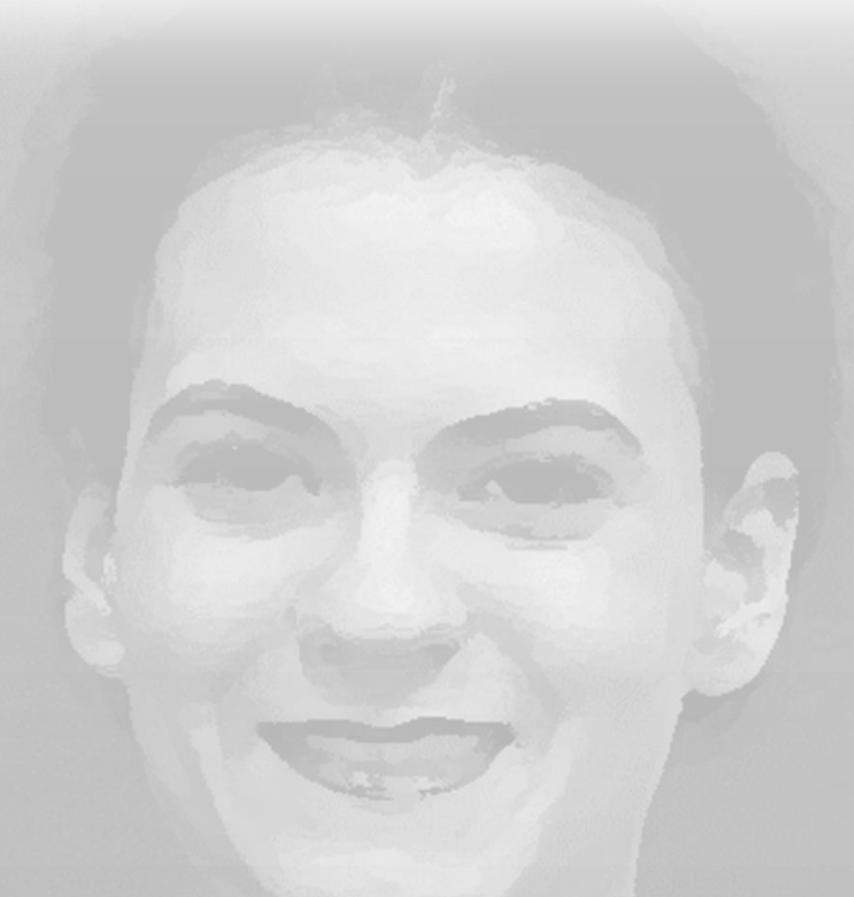
- Imagens em 3D
- Distância entre pontos faciais específicos
- *Support Vector Machine* – acerto médio de 87,1%

Trabalhos Correlatos

Tang e Huang (2008)



Fonte: Tang e Huang (2008)



Requisitos

Requisitos

Funcionais:

RF01 - utilizar as técnicas apresentadas para detecção de múltiplas faces em uma imagem;

RF02 - utilizar as técnicas apresentadas para extração de características faciais;

RF03 - extrair as características relevantes para o reconhecimento das expressões faciais;

RF04 - utilizar redes neurais artificiais para a classificação de expressões faciais

Requisitos

Não Funcionais (RNF):

RNF01 - ser implementado utilizando o ambiente de desenvolvimento Microsoft Visual Studio;

RNF02 - ser implementado utilizando a linguagem de programação C++;

RNF03 - ser implementado utilizando a biblioteca *dlib*;

RNF04 - reconhecer adequadamente imagens no formato JPG e PNG;

RNF05 - otimizar o uso do processador através de paralelismo para extrair as características faciais;

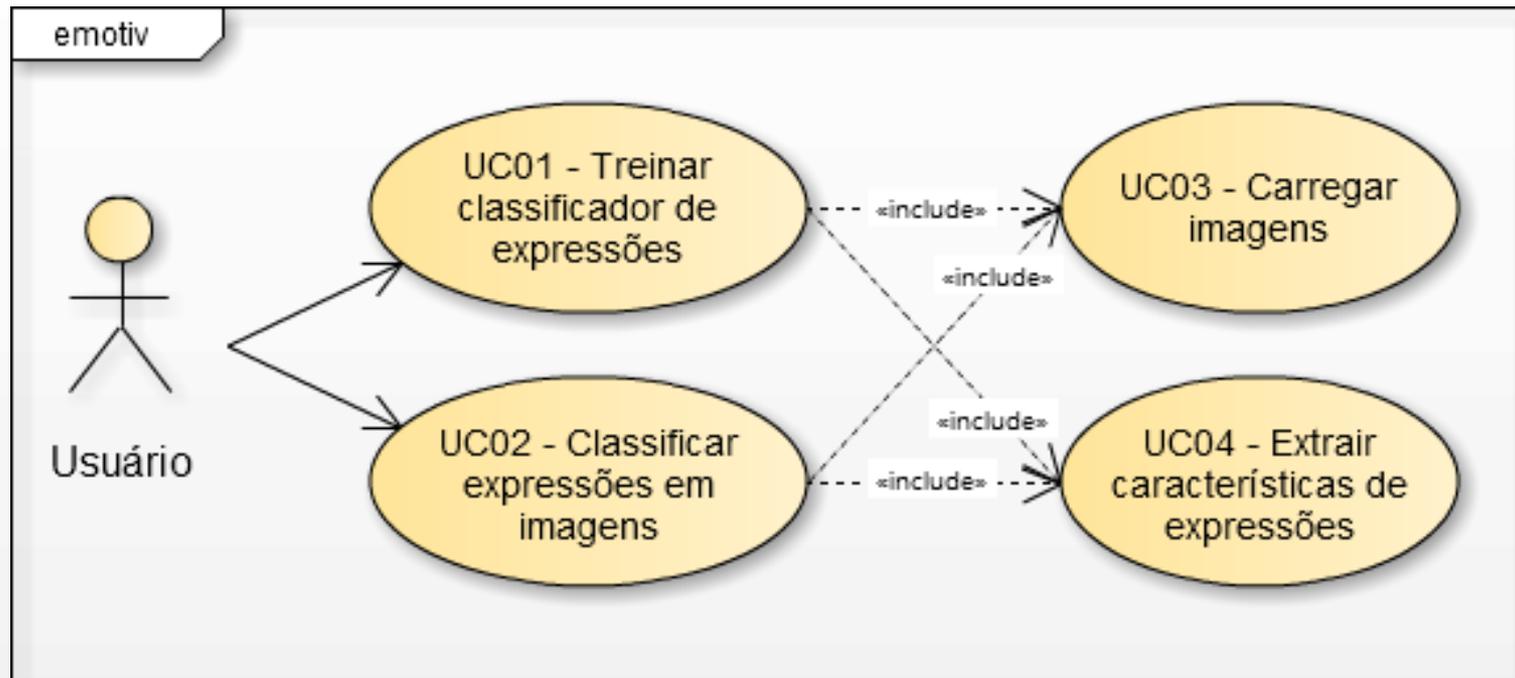
RNF06 - disponibilizar uma interface de linha de comando para interação



Especificação

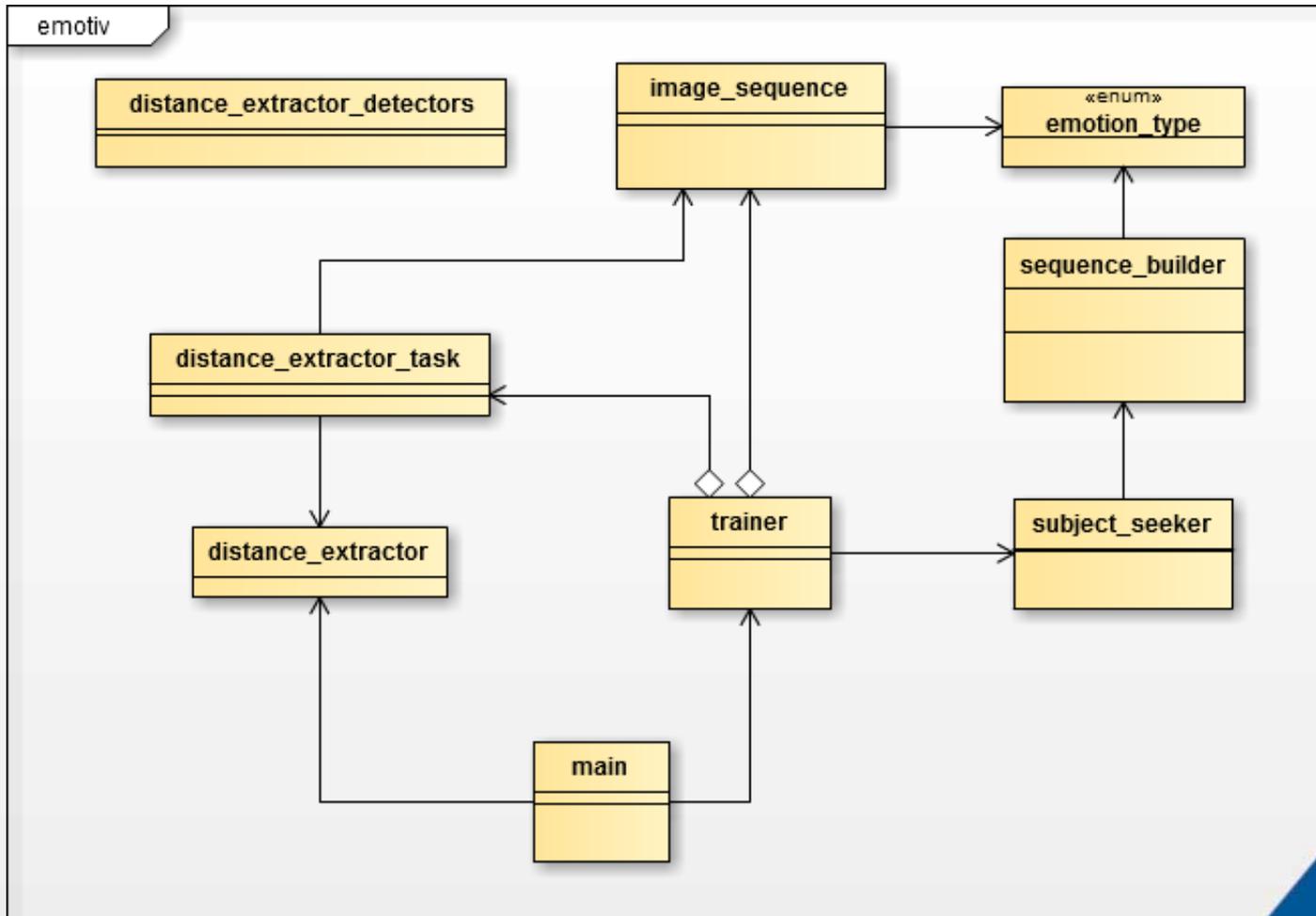
Especificação

Casos de Uso:



Especificação

Classes:





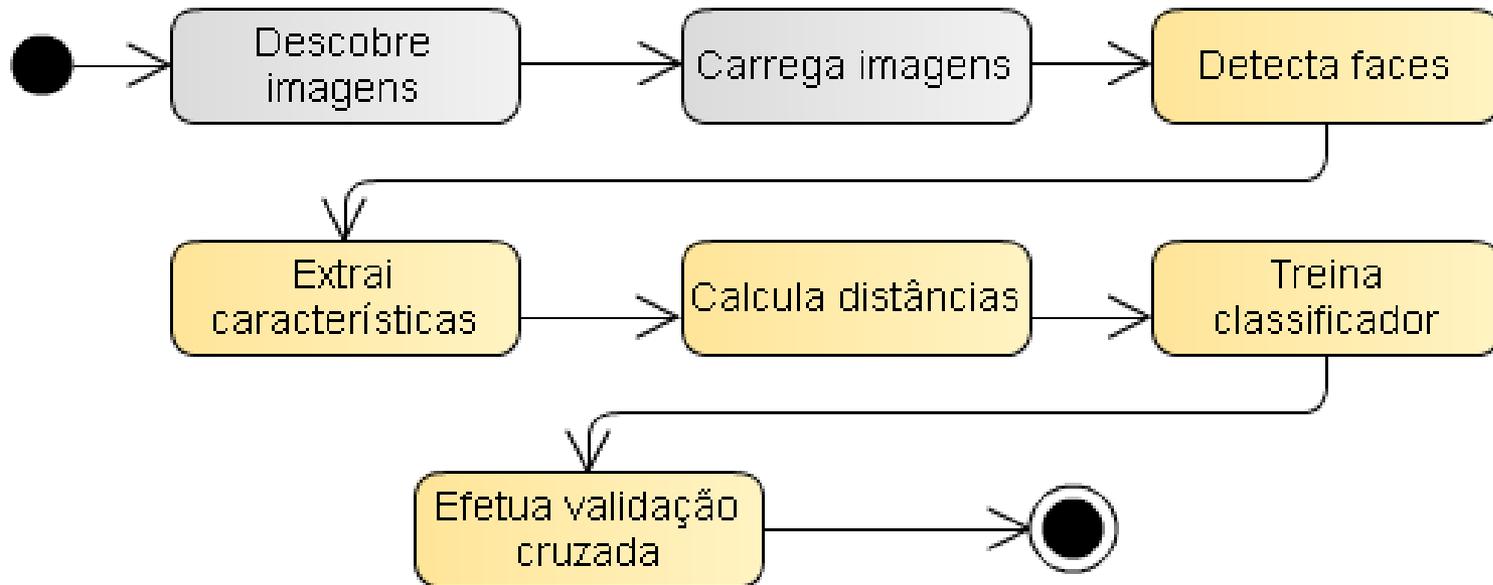
Implementação

Implementação

- Biblioteca *dlib*
 - Processamento de imagens
 - Aprendizagem de máquina
 - Utilitários
 - C++
- Base de testes *Extended Cohn-Kanade Dataset (CK+)*
 - 7 emoções (inclui satisfação) + neutra
 - 123 modelos
 - 327 imagens rotuladas

Implementação

Treinamento

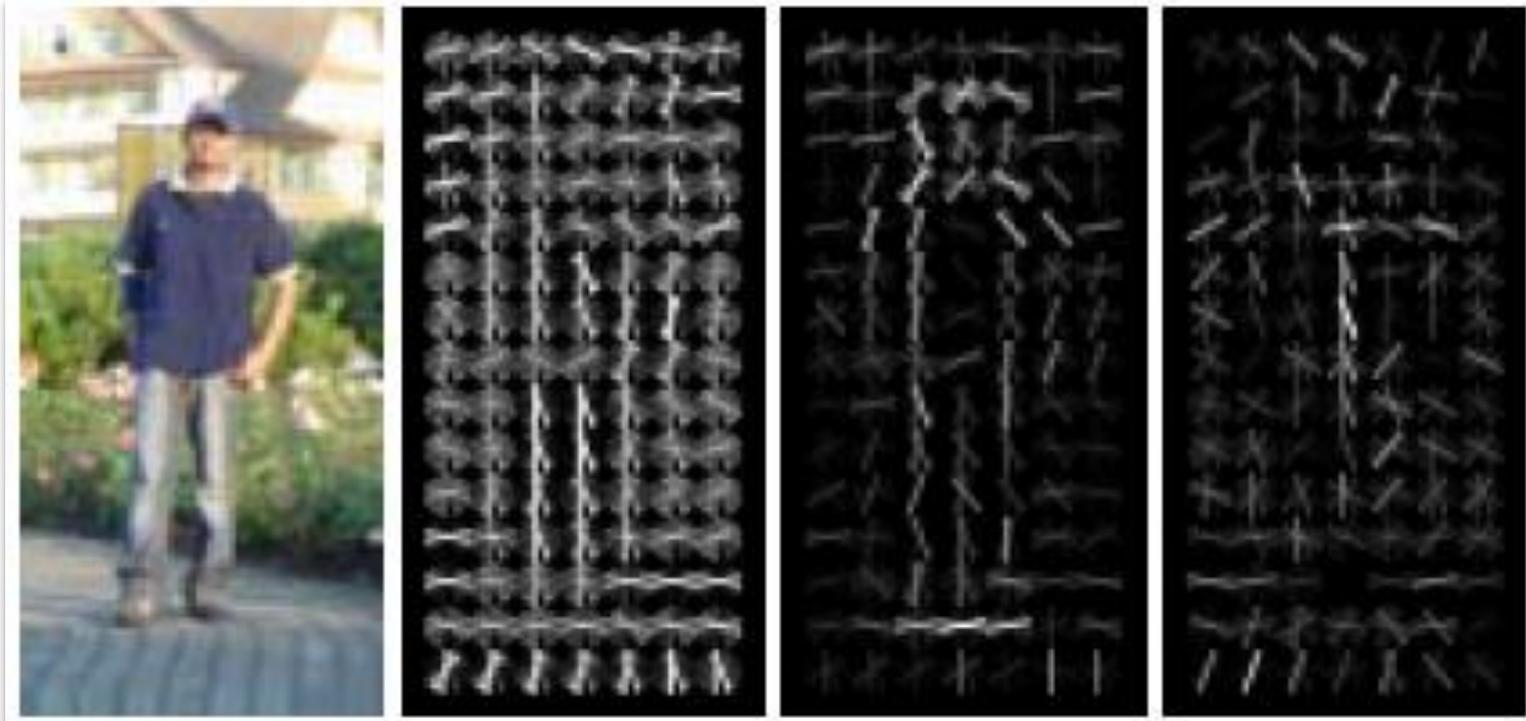


Implementação

- Detecção de faces
 - Histograma de Gradientes Orientados (*dlib*)
 - Apresentada por Dalal & Triggs (2008)
 - Realce bordas
 - Cômputo da intensidade de células espaciais para detecção de partes do objeto
 - Classificação em objeto por SVM
 - Repetição através de pirâmide de imagem e Janela de Detecção Deslizante

Implementação

- Detecção de faces



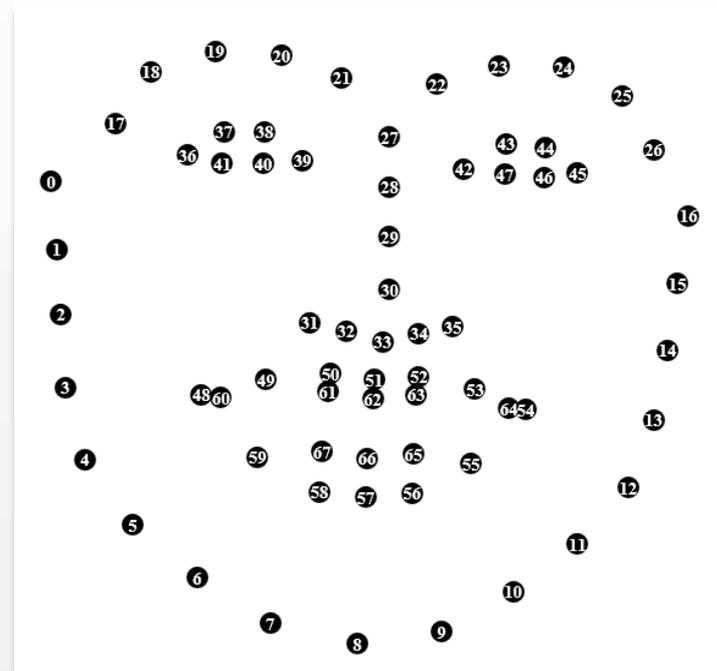
Fonte: Dalal e Triggs (2008)

Implementação

- Extração de características faciais
 - Arranjo de Árvores de Regressão (*dlib*)
 - Apresentado por Kazemi e Sullivan (2014)
 - Aproxima 68 pontos faciais com base em intensidade de *pixels*
 - Apenas uma detecção por imagem
 - Rápido, na ordem de milissegundos

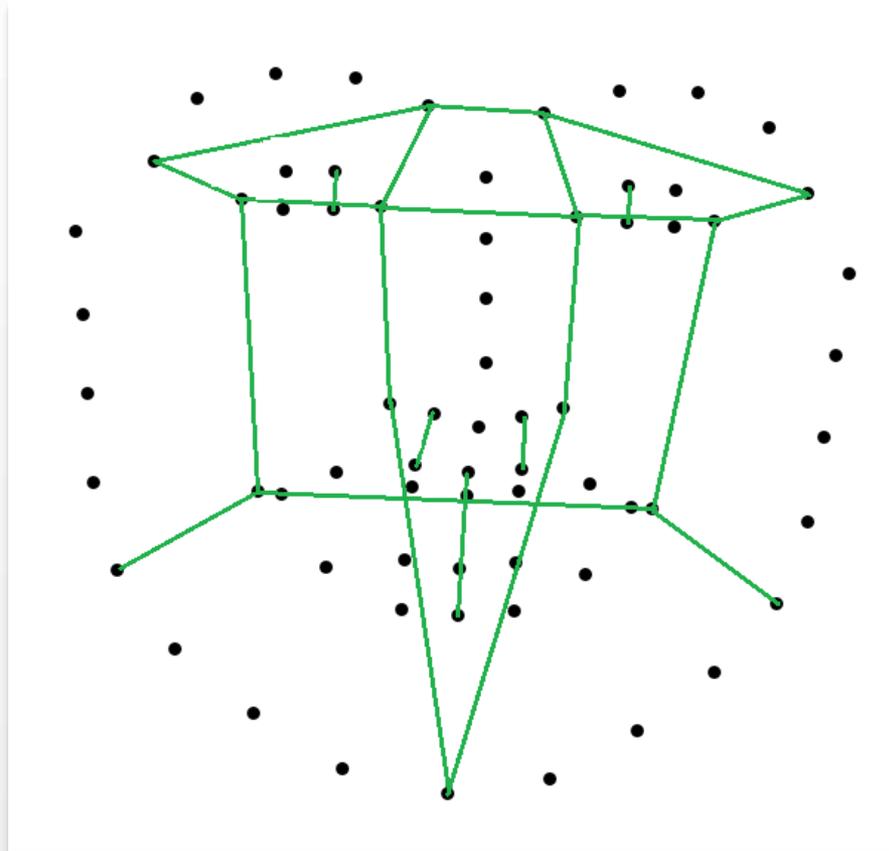
Implementação

- Extração de características faciais



Implementação

- Cálculo de distância entre pontos específicos



Implementação

- Treinamento da Rede Neural
 - Utilizada a implementação de *perceptron* multicamada com retropropagação da biblioteca *dlib*
 - Utilizado apenas 2/3 da base
 - Por padrão:
 - Repete o treino 500 vezes
 - Utiliza 26 nós em cada camada oculta
 - Utiliza 3 nós na camada de saída
- Validação cruzada dos dados
 - Executada sobre o 1/3 restante



Operacionalidade

Operacionalidade

Instruções de uso:

```
E:\dev\git\emotiv\emotiv\x64\Release>emotiv -h
Modo de usar:
  treinamento - emotiv.exe [opcoes gerais] -t --dataset <pasta>
  classificacao - emotiv.exe [opcoes gerais] [-c <arquivo>] <imagens>
execucao:
  -c <arg>          Utiliza o arquivo <arg> como classificador (padrao:
                    'emotiv.dat').

opcoes gerais:
  -h                (help) Exibe esta mensagem de ajuda.
  --sp <arg>        (shape predictor) Utiliza o shape predictor do arquivo
                    <arg> (padrao: 'shape_predictor_68_face_landmarks.dat').
  -t                Treina o detector de emocoes e persiste no disco.
  -v                (verbose) Exibe etapas do treinamento ou classificacao no
                    terminal.

treinamento:
  --dataset <arg>   Utiliza as imagens do dataset definido em <arg> para o
                    treinamento.
  -e <arg>          Define as emocoes que serao treinadas. <arg> deve ser uma
                    concatenacao do codigo numerico de cada emocao, sendo: (1)
                    raiva, (3) desprezo, (4) medo, (5) felicidade, (6)
                    tristeza, (7) surpresa. (padrao: 56)
  -i <arg>          Executa <arg> iteracoes durante o teste. (padrao: 1000)
  --no1 <arg>       Utiliza <arg> nos na primeira camada oculta. (padrao: 26)
  --no2 <arg>       Utiliza <arg> nos na segunda camada oculta. (padrao: 26)
  -r                (randomize) Ordena o dataset de forma aleatoria.
  --threads <arg>   Utiliza <arg> threads para o treinamento (padrao: suportado
                    pelo processador).
```

Operacionalidade

Treinamento:

```
C:\WINDOWS\system32\cmd.exe
E:\dev\git\emotiv\emotiv\x64\Release>emotiv -t --dataset e:\dev\dataset\cohn-kanade-images
194 imagens encontradas
97      neutra
0       raiva
0       satisfacao
0       desprezo
0       medo
69      felicidade
28      tristeza
0       surpresa
Iniciando treinamento

Iniciando testes
1       /S010/006/S010_006_00000001.png
ERRADO: classificou tristeza ao inves de neutra
2       /S011/002/S011_002_00000001.png
ERRADO: classificou tristeza ao inves de neutra
3       /S011/006/S011_006_00000001.png
ERRADO: classificou tristeza ao inves de neutra
4       /S014/002/S014_002_00000001.png
ERRADO: classificou tristeza ao inves de neutra

60      /S076/006/S076_006_00000019.png
CERTO: classificou felicidade
61      /S078/004/S078_004_00000027.png
ERRADO: classificou tristeza ao inves de felicidade
62      /S079/004/S079_004_00000026.png
CERTO: classificou felicidade
63      /S083/003/S083_003_00000019.png
CERTO: classificou felicidade
64      /S085/002/S085_002_00000014.png
CERTO: classificou felicidade

      neu      tri      fel
neu   0        32       0
tri   0         9       0
fel   0         1      22

Acertos: 31
```

Operacionalidade

Classificação:

```
C:\WINDOWS\system32\cmd.exe
E:\dev\git\emotiv\emotiv\x64\Release>emotiv happy-people.jpg ./album
Analizando happy-people.jpg
  Face 0:      felicidade      0.801938
  Face 1:      felicidade      0.801938
  Face 2:      felicidade      1.1403e-311
  Face 3:      felicidade      1.69407e-21
  Face 4:      felicidade      4.94066e-324
Analizando .\album\Andre sorrindo.jpg
  Face 0:      felicidade      5.26354e-315
Analizando .\album\Eli meio feliz, de oculos.jpg
  Face 0:      neutra      0.529538
Analizando .\album\Gustavo sorrindo, metade sombra.jpg
  Face 0:      felicidade      1.97033e+15
Analizando .\album\Gustavo, Nurielly e amigo sorrindo, baixa qualidade.jpg
  Face 0:      felicidade      1.09951e+12
  Face 1:      felicidade      1.09951e+12
  Face 2:      felicidade      1.69407e-21
Analizando .\album\Leander neutro, off-center.jpg
  Face 0:      neutra      0.525699
Analizando .\album\Leander neutro.jpg
  Face 0:      neutra      0.535808
Analizando .\album\Leander sorrindo, monochrome.jpg
  Face 0:      felicidade      1.69407e-21
Analizando .\album\Leander triste.jpg
  Face 0:      tristeza      -4.96308e-24
Analizando .\album\Maicon serio.jpg
  Face 0:      tristeza      6.95161e-310
Analizando .\album\Reinoldo sorrindo.jpg
  Face 0:      felicidade      6.59707e+12
Analizando .\album\Vivian e Bruno sorrindo.jpg
  Face 0:      felicidade      5.26354e-315
  Face 1:      felicidade      1.69407e-21
Analizando .\album\Vivian neutra.jpg
  Face 0:      neutra      0.533706
Analizando .\album\Vivian sorrindo.jpg
  Face 0:      felicidade      1.69407e-21
E:\dev\git\emotiv\emotiv\x64\Release>
```



Resultados & Discussão

Resultados e Discussões

- Detectado possível problema na implementação de RNA da biblioteca *dlib*
 - Execuções foram repetidas com as mesmas entradas e diferente estratégias de cálculo de distância
 - Sem qualquer padrão evidente, em várias execuções a saída da rede se mostrou “fixa”
- Apenas as expressões Neutra, Alegria e Tristeza puderam ser classificadas corretamente

Resultados e Discussões

- Matriz de confusão para um caso de sucesso (89% de acerto)

	Neu.	Fel.	Tri.
Neu.	28	0	4
Fel.	1	22	0
Tri.	2	0	7

Resultados e Discussões

Características / Trabalhos	Oliveira e Jaques (2013)	Filko e Martinović (2013)	Tang e Huang (2008)	Protótipo
Nº de emoções	6	6	6	2
Classificador	Redes neurais	Redes neurais	SVM	Redes neurais
Dimensão das imagens	2D	2D	3D	2D
Características	Distâncias entre pontos	<i>Eigenfaces</i>	Distâncias entre pontos	Distâncias entre pontos
Taxa de reconhecimento	89,87%	70%	87,1%	89%



Conclusões & Sugestões

Conclusões

- Até 2 expressões faciais reconhecidas com 89% de precisão
- Interface de linha de comando que possibilita testar diferentes arquiteturas de RNA
- Extração de características de imagens de testes em paralelo
- Técnicas e bibliotecas apresentam maturidade para o desenvolvimento de ferramentas de reconhecimento de expressões faciais

Trabalhos Futuros

- Alterar a biblioteca de RNA
- Utilizar outra técnica de aprendizagem de máquina
- Flexibilizar mais a interface de linha de comando de modo que sua integração com outros programas seja facilitada
- Exibir as coordenadas e dimensões das faces detectadas
- Classificar as demais expressões faciais universais
- Disponibilizar o classificador serializado embutido no programa



Demonstração



Perguntas?