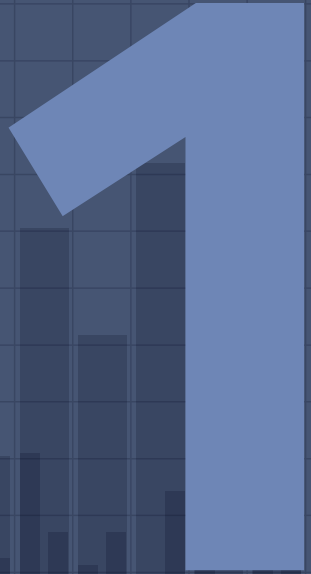


Reconhecimento de Sinais do Alfabeto ASL Utilizando o Descritor HOG e Classificador SVM



INTRODUÇÃO



Contexto

- A comunidade das pessoas deficientes auditivas enfrenta, dentre outros empecilhos, a dificuldade de se comunicar com a sociedade majoritariamente não surda.

Contexto

- De acordo com informações apuradas em 2016 pela OMS, as pessoas com deficiência auditiva representam 5% da população mundial;
- Esse número equivale a aproximadamente 360 milhões de pessoas.



Contexto

- A maioria dos deficientes auditivos utilizam alguma linguagem de sinais para se comunicarem.



Descrição do problema

- É possível compreender a dificuldade de comunicação entre surdos e não surdos;
- A inclusão dos surdos na sociedade tem enfrentado a falta de conhecimento de grande parte dos ouvintes sobre línguas de sinais, tornando consideravelmente complicada a comunicação entre eles.



Justificativa

- Facilitar a comunicação entre os ouvintes que não possuem conhecimento da língua e os deficientes auditivos seria algo muito pertinente;
- Além disso, pesquisas nesse âmbito podem contribuir para uma maior inclusão dos surdos na sociedade.

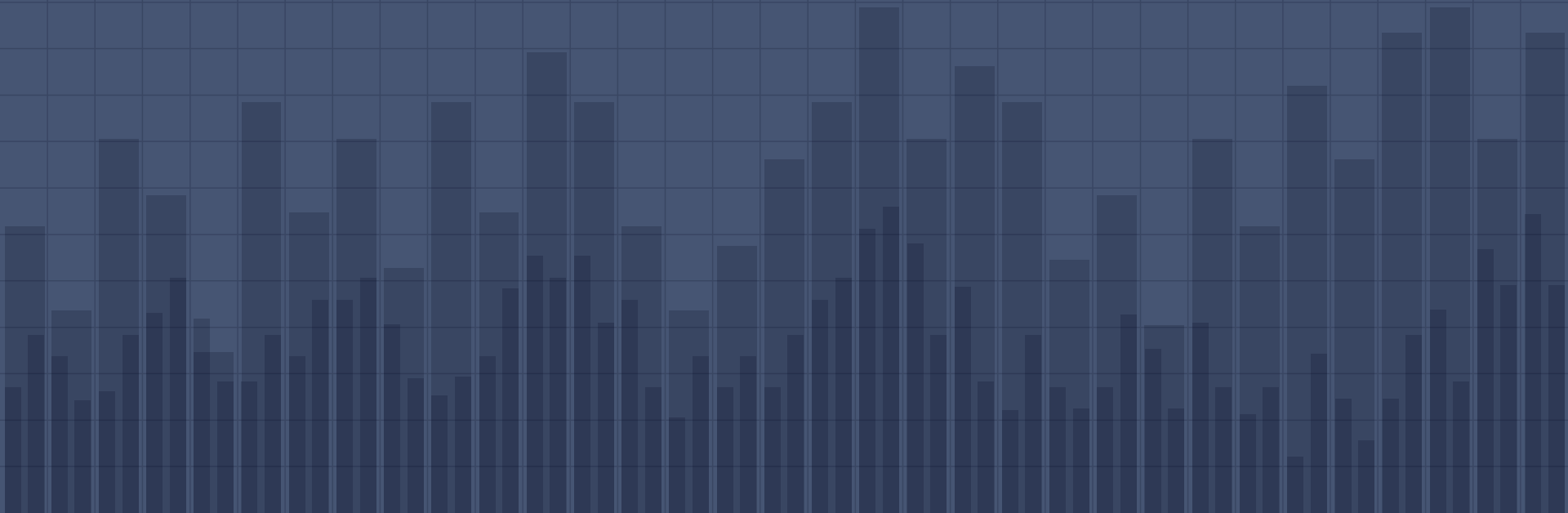
Objetivos

- ▣ A finalidade deste trabalho é elaborar um protótipo de uma ferramenta na linguagem MATLAB, que consiga reconhecer alguns sinais do alfabeto da ASL(American Sign Language) utilizando técnicas de processamento imagens;
- ▣ Identificar e classificar símbolos estáticos da língua e averiguar a efetividade do protótipo.

DESENVOLVIMENTO

2

BASE DE DADOS



Base de dados

- ▣ O conjunto de dados utilizado foi coletado para os 2 trabalhos a seguir:
 - "Hand gesture recognition with Leap Motion and Kinect devices", IEEE, 2014;
 - E a extensão do mesmo "Hand Gesture Recognition with Jointly Calibrated Leap Motion and Depth Sensor", Multimedia Tools and Applications, 2015.

Base de dados

- O conjunto de dados foi adquirido com os dispositivos Leap Motion e Kinect, permitindo a construção e avaliação de sistemas híbridos de reconhecimento de gestos;
- Este conjunto de dados contém gestos realizados por 14 pessoas diferentes, cada uma executando 10 gestos diferentes, repetidos 10 vezes cada.



PRÉ-PROCESSAMENTO



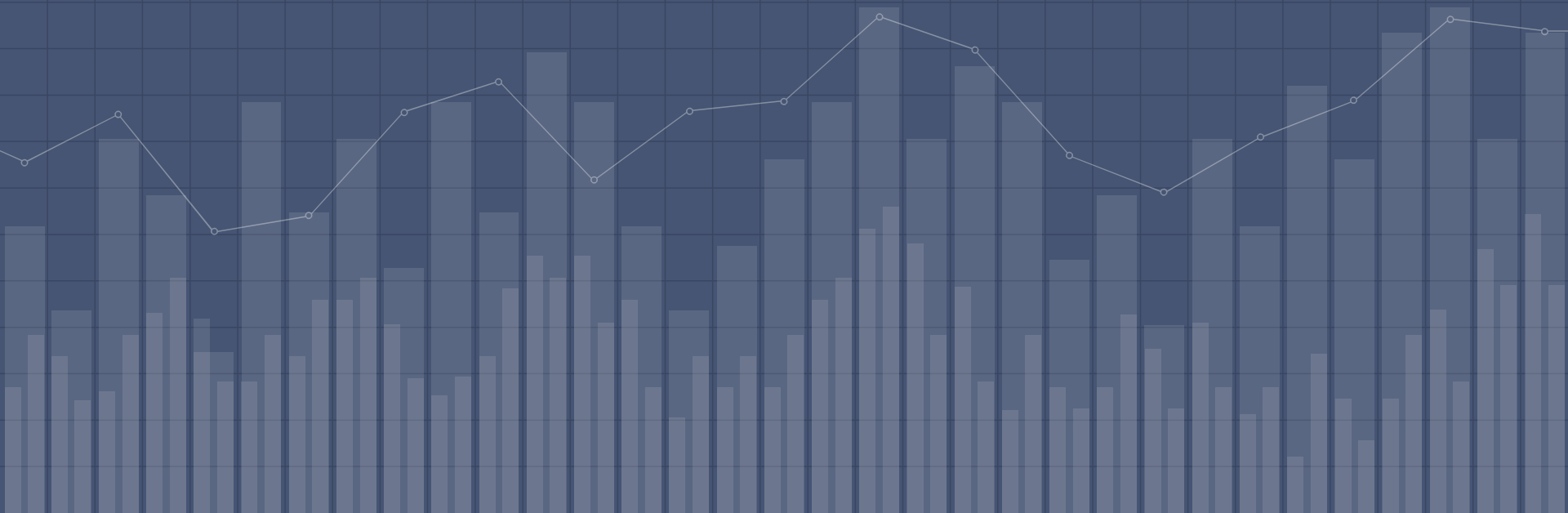
Pré-processamento

- O único pré-processamento realizado se trata da conversão das imagens em escala de cinza, algo simples e comum em alguns casos de reconhecimento;
- Para isso foi utilizada a função *rgb2gray*.

Pré-processamento



SEGMENTAÇÃO



Segmentação

- A parte de segmentação da área da mão não foi realizada utilizando algoritmos de rastreamento, devido ao fato do interesse maior do trabalho ser classificar os sinais abordados ao invés de encontrar objetos em uma imagem;
- Sendo assim, a segmentação foi feita manualmente selecionando parte das imagens de cada sinal e recortando a área de interesse utilizando um software de edição de imagens.

Segmentação



EXTRAÇÃO DE CARACTERÍSTICAS



Extração de Características

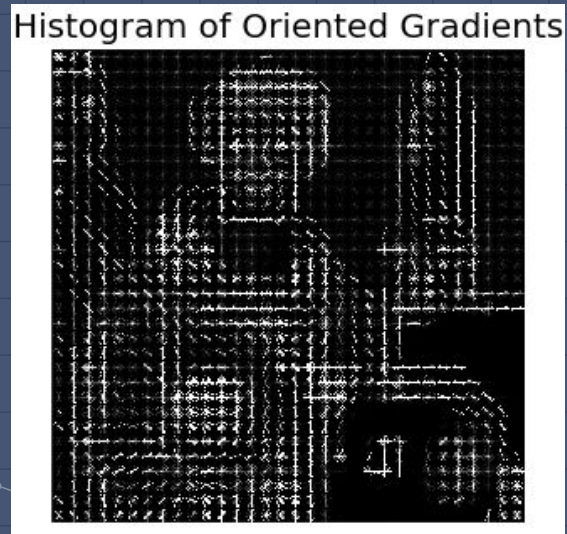
- ▣ O descritor de forma utilizado na etapa de extração de características foi o HOG(Histogram of Oriented Gradient).

HOG(Histograma de Gradientes Orientados)

- O algoritmo HOG (Dalal and Triggs 2005) é um descritor que calcula o histograma da orientação dos gradientes na imagem;
- O descritor final é um vetor de histogramas extraídos da imagem.

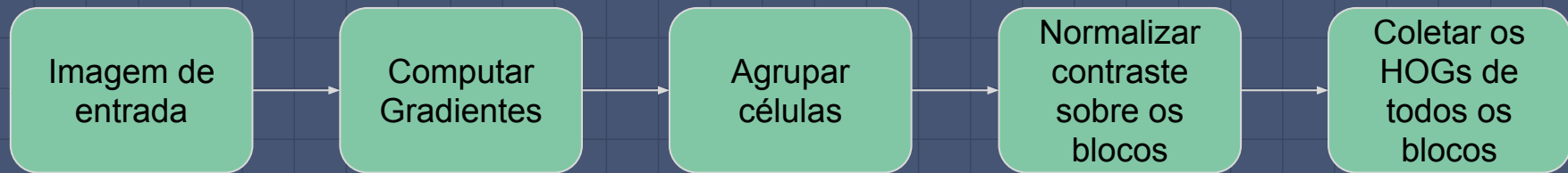


HOG(Histograma de Gradientes Orientados)



HOG(Histograma de Gradientes Orientados)

O HOG segue os seguintes passos:



CLASSIFICAÇÃO₅



Classificação

- ▣ No processo de classificação criamos um template SVM(Support Vector Machine) e são utilizados dois métodos, um para treinar e outro para classificar propriamente:
 - templateSVM;
 - fitcecoc;
 - predict.

SVM (Support Vector Machine)

- Uma máquina de vetores de suporte é um conjunto de métodos do aprendizado supervisionado que analisam os dados e reconhecem padrões, usado para classificação e análise de regressão.



Treinamento (*fitcecoc*)

- Os parâmetros fornecidos são um descritor a ser treinado e um vetor de etiquetas que está relacionado ao descritor de treino, além de parâmetros para otimizar o treinamento.



Classificação (*predict*)

- Os parâmetros fornecidos são o modelo de treino criado pela função *fitcecoc* e o descritor de teste a ser classificado.

Classificação

Chamada das funções:

```
parametrosSVM = templateSVM('KernelFunction','polynomial','KernelScale',  
    'auto','Standardize',1);  
modelo = fitcecoc(desc_treinoHOG, etiqueta_treinoHOG,'Learners',parametrosSVM,  
    'Coding','onevsall');  
classificador = predict(modelo, desc_testeHOG);  
matConf = confusionmat(etiqueta_testeHOG,classificador)
```

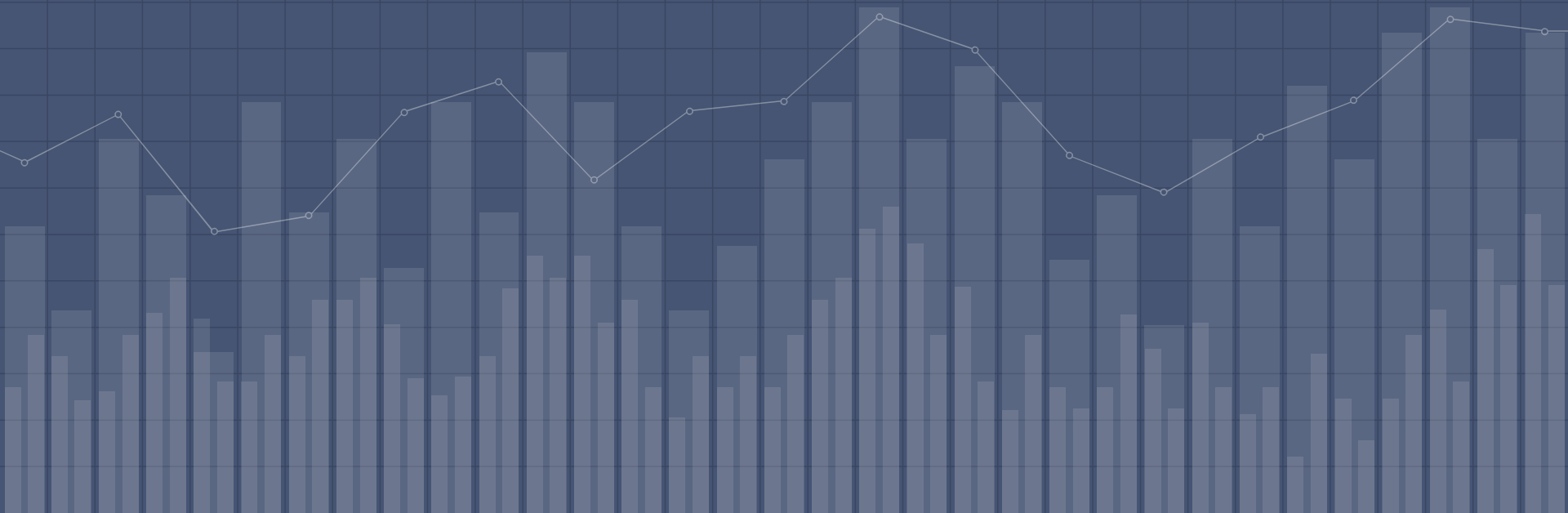
RESULTADOS

3

Resultados

- ▣ Serão apresentados os resultados antes e após fazermos uma melhoria na implementação;
- ▣ Primeiramente os testes foram realizados com todas as imagens com suas dimensões originais;
- ▣ Em seguida utilizamos a função *imresize* utilizando interpolação bilinear e redimensionamos todas as imagens para as dimensões da maior imagem (317x317).

ANTES DA MELHORIA



10 TESTES (70-30)

Taxa de acerto médio:

Sinais	A	D	I	L	V	W	Y
Acerto(%)	96.47	45.00	71.18	72.50	25.63	40.59	54.12

Media total (por sinais): 57.93%



10 TESTES (70-30)

Maior taxa de acerto por sinal:

Sinais	A	D	I	L	V	W	Y
Acerto(%)	100.00	68.75	82.35	81.25	43.75	52.94	76.47

Total: (84/116) : 72.40% das imagens

10 TESTES (70-30)

Taxa de acerto médio por teste:

Testes	1	2	3	4	5	6	7	8	9	10
Acerto(%)	48.58	54.88	59.30	62.76	58.46	59.30	52.31	60.82	58.30	56.51

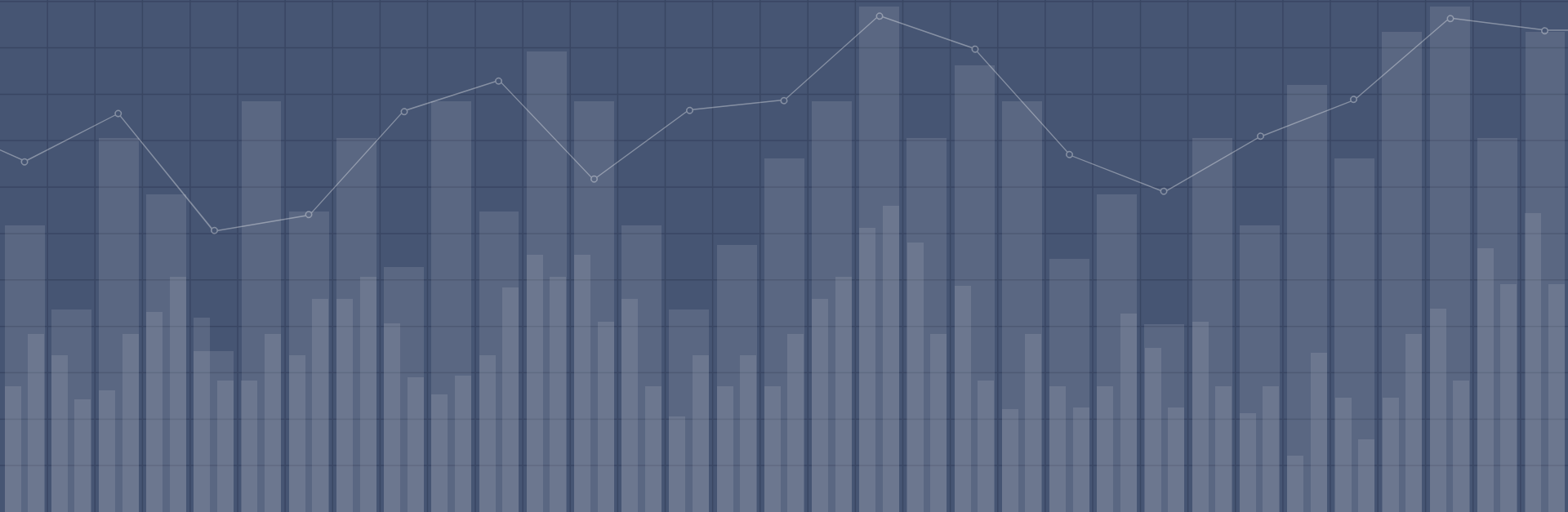
Media total (10 testes): 57.12%



MATRIZ DE CONFUSÃO (Melhor teste)

	A	D	I	L	V	W	Y
A	17	0	0	0	0	0	0
D	4	10	0	0	1	1	0
I	3	2	10	1	1	0	0
L	2	0	1	11	2	0	0
V	0	5	0	1	6	3	1
W	0	2	1	2	3	8	1
Y	2	2	1	1	0	0	11

APÓS A MELHORIA



10 TESTES (70-30)

Taxa de acerto médio:

Sinais	A	D	I	L	V	W	Y
Acerto(%)	95.30	79.38	87.06	96.88	71.88	77.65	93.53

Média total (por sinais): 85.95%

A decorative background graphic at the bottom of the slide. It features a line chart with white circular markers connected by a thin white line, showing an overall upward trend. Below the line chart is a bar chart with numerous vertical bars of varying heights, rendered in a semi-transparent blue color.

10 TESTES (70-30)

Maior taxa de acerto por sinal:

Sinais	A	D	I	L	V	W	Y
Acerto(%)	100.00	87.50	100.00	100.00	87.50	88.24	100.00

Total: (110/116) : 94.83% das imagens de teste

10 TESTES (70-30)

Taxa de acerto médio por teste:

Testes	1	2	3	4	5	6	7	8	9	10
Acerto(%)	89.60	86.08	86.92	86.92	83.61	86.19	83.56	84.35	86.19	86.08

Média total (10 testes): 85.95% dos sinais



MATRIZ DE CONFUSÃO (Melhor teste)

	A	D	I	L	V	W	Y
A	17	0	0	0	0	0	0
D	1	14	0	1	0	0	0
I	0	1	16	0	0	0	0
L	0	0	0	16	0	0	0
V	0	2	0	0	12	2	0
W	2	1	0	0	0	13	1
Y	1	0	0	0	0	0	16

CONCLUSÃO E REFERÊNCIAS



Conclusão

- Os resultados obtidos após os testes realizados com a melhoria foram satisfatórios em nossa perspectiva;
- Eles provavelmente iriam melhorar se conseguíssemos eliminar partes da imagem que podem confundir a geração do descritor e se fossem utilizadas imagens de profundidade;
- A implementação deste projeto contribuiu para o conhecimento dos envolvidos.

Conclusão

- ▣ Os avanços nos estudos e pesquisas na área de reconhecimento de linguagens de sinais por processamento de imagens demandam tempo e apresentam uma dificuldade significativa;
- ▣ É possível perceber que uma ferramenta para reconhecimento em tempo real envolve uma complexidade elevada.

Referências

[1] G. Marin, F. Dominio, P. Zanuttigh, "Hand gesture recognition with Leap Motion and Kinect devices", IEEE International Conference on Image Processing (ICIP), Paris, France, 2014

[2] G. Marin, F. Dominio, P. Zanuttigh, "Hand Gesture Recognition with Jointly Calibrated Leap Motion and Depth Sensor", Multimedia Tools and Applications, 2015

[3] L.Cerna, 2015, "Fusao de descritores de histogramas de gradientes para a deteccao de faces baseado em uma cascata de classificadores", Universidade Federal de Ouro Preto,
http://www.repositorio.ufop.br/bitstream/123456789/4227/1/DISSERTA%C3%87%C3%830_Fus%C3%A3oDescritoresHistogramas.pdf

[4] A.Silva, 2014, "Descritores de Imagem", <http://www.joinville.udesc.br/portal/professores/andretavares/materiais/descritores2.pdf>

Obrigado!

Perguntas

