

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

Mateus Lana

DECOM - Instituto de Ciências Exatas e Biológicas
Universidade Federal de Ouro Preto
Ouro Preto, Brasil
mateus_lana7@hotmail.com

Thiago Santana

DECOM - Instituto de Ciências Exatas e Biológicas
Universidade Federal de Ouro Preto
Ouro Preto, Brasil
thiago63.guitarrista@gmail.com

Abstract—Automatic recognition of signal languages promises to facilitate communication between hearing impaired and the rest of the population. Although there are current studies for sign language recognition techniques, this is still a very complex subject. The inclusion of the deaf in society has faced the lack of knowledge of most listeners on sign languages, making communication between them more complicated. Given this, one can think about whether image recognition techniques can aid in the communication between deaf and non-deaf individuals. The purpose of this work is to elaborate a prototype of a tool that can recognize some symbols of the alphabet of the American Sign Language (ASL), seeking to facilitate communication between listeners who do not have knowledge of the language and the hearing impaired besides seeking to contribute to research in this scenario.

Resumo—O reconhecimento automático de linguagens de sinais promete facilitar a comunicação entre deficientes auditivos e o restante da população. Apesar de existirem estudos atuais para técnicas de reconhecimento de linguagens de sinais, este ainda é um assunto bem complexo. 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. Dado isso, pode-se pensar se técnicas de reconhecimento de imagens podem auxiliar na comunicação entre surdos e não surdos. A finalidade deste trabalho é elaborar um protótipo de uma ferramenta que consiga reconhecer alguns símbolos do alfabeto da American Sign Language (ASL), buscando facilitar a comunicação entre os ouvintes que não possuem conhecimento da língua e os deficientes auditivos além de procurar contribuir para pesquisas nesse cenário.

I. INTRODUÇÃO

É sabido que a comunidade das pessoas deficientes auditivas enfrenta, dentre outros empecilhos, a dificuldade de se comunicar com a sociedade majoritariamente não surda. Afazeres considerados habituais para a população ouvinte, como ir a um estabelecimento comercial ou assistir aulas, para os surdos podem se tornar algo bem mais complicado.

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, um número “maior que o total da população dos Estados Unidos”.

A maioria dos deficientes auditivos utilizam alguma linguagem de sinais para se comunicarem. Assim como as diversas línguas naturais e humanas existentes, elas são compostas

por níveis linguísticos como: fonologia, morfologia, sintaxe e semântica. Sendo assim, para se comunicar nelas, não basta apenas conhecer sinais, é necessário conhecer a sua gramática para combinar as frases, estabelecendo a comunicação de forma correta. Tendo isso em vista é possível compreender a dificuldade de comunicação entre surdos e não surdos.

Apesar de nos últimos tempos existirem alguns esforços para simplificar a comunicação entre a população ouvinte e os deficientes auditivos, ainda é evidente a necessidade de avanços nesse âmbito. 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.

Com pesquisas recentes e abertas a estudo, a tradução e reconhecimento de línguas de sinais por computadores são esferas bem complexas. Isso se deve ao fato dessas línguas apresentarem não somente sinais estáticos para indicar algumas palavras e letras mas também sinais dinâmicos, que envolvem além das configurações das mãos, os movimentos corporais. Portanto, para uma computação automática mais significativa seria necessário o processamento de vídeos ou a utilização luvas com sensores, por exemplo. No entanto, isso poderia demandar um custo bem maior.

É possível encontrar algumas diferentes abordagens computacionais para reconhecimento de linguagens de sinais, utilizando diversas metodologias. Algumas delas utilizando redes neurais; técnicas de reconhecimento da mão através da cor da pele; técnicas para tratar os movimentos das mãos como HMM (Hidden Markov Model) e DTW (Dynamic Time Warping); classificação dos sinais usando k-Nearest Neighbors (k-NN), dentre outras.

As técnicas de reconhecimento por computador das línguas de sinais em geral, tiveram uma substancial evolução. Visão computacional se estabeleceu como a abordagem mais utilizada, por ser mais acessível aos usuários. Os algoritmos de detecção de imagens, são mais robustos às variações do ambiente, prescinde de ambientes controlados e é de fácil uso pelos usuários em geral.

O propósito deste trabalho é elaborar um protótipo de uma ferramenta em MATLAB que consiga reconhecer alguns dos sinais do alfabeto da ASL, buscando contribuir nos estudos de reconhecimento de linguagem de sinais. Dessa maneira,

participar da evolução na facilitação da comunicação entre os ouvintes que não possuem conhecimento da língua e os deficientes auditivos. A ideia é utilizar um conjunto de técnicas de processamento digital de imagens aplicadas em sinais estáticos procurando identificar e traduzir símbolos do alfabeto da língua e averiguar a efetividade do protótipo.

II. TRABALHOS RELACIONADOS

O reconhecimento e a tradução de línguas de sinais por computadores são áreas bastante complexas, com estudos ainda recentes e abertos à pesquisa (NETO; OQUENDO, 2013). É possível encontrar algumas diferentes abordagens computacionais para reconhecimento de linguagens de sinais, cada uma delas utilizando metodologias diferentes. Serão apresentadas a seguir algumas metodologias que foram estudadas e abordadas em diferentes artigos revisados. Foram avaliados um total de 10 artigos, dentre eles há também alguns somente teóricos.

Além dos métodos que serão apresentados existem outras diferentes abordagens utilizando técnicas de reconhecimento da mão através da cor da pele e técnicas para tratar os movimentos das mãos como HMM (Hidden Markov Model) e DTW (Dynamic Time Warping), por exemplo, dentre outras técnicas.

Primeiramente, em [5] temos uma metodologia que foi utilizada no reconhecimento da Língua brasileira de sinais (LIBRAS) utilizando-se técnicas de processamento de imagens digitais binárias. A ideia abordada no artigo é, após capturar uma imagem, primeiramente transformá-la em binária. Em seguida os devidos ajustes de reescalonamento e centralização devem ser efetuados. Para finalizar, as diferenças entre a imagem processada e as imagens arquivadas no programa devem ser calculadas, obtendo a menor diferença que consequentemente corresponderá à imagem do arquivo que traduz o símbolo introduzido pelo usuário. Apesar do bom resultado de 100% no reconhecimento esse trabalho se restringiu muito ao reconhecer apenas os algarismos de 0 a 4.

Outro artigo [6] apresenta um conjunto de técnicas da área de Visão Computacional aplicado em um modelo de sistema de comunicação mediado por computador, são elas: pré-processamento, segmentação de imagens, extração de características e reconhecimento e interpretação. Trata-se de um classificador de imagens(Classlib) capaz de reconhecer determinados sinais da Libras e apresentar ao usuário o respectivo significado em Português. O reconhecimento é realizado em 2 fases: uma para encontrar qual o grupo a que pertence o sinal da Libras e outra uma busca refinada dentro de um subconjunto restrito de sinais. Na segunda fase é aplicado o conceito de árvore de decisão da área de Inteligência Artificial.

Uma abordagem diferente que foi utilizada no reconhecimento da American Sign Language(ASL) em [9], utiliza uma técnica que elabora um conjunto de testes através do treinamento de várias imagens de cada letra do alfabeto da ASL. Os pontos de referência de cada letra produzem dois descritores (sendo um de 72 pontos e outro de 180 pontos) e eles são armazenados no banco de dados. O descritor da

imagem de teste é então comparado com os do banco de dados. Finalmente, a distância euclidiana classifica as imagens de teste na letra reconhecida. A taxa de acerto é notável, com um total de 75% com descritor de 72 pontos e 79.9% com um descritor de 180 pontos.

Foi analisada também uma implementação em [7] que realiza a captura de imagens gestuais via Kinect e aplica treino da Rede Neural Artificial (RNA), entrelaçados em uma gama de conceitos e métodos, possibilitando distinguir os gestos obtidos, que define as letras do alfabeto da Libras. Para o reconhecimento dos gestos, foram utilizadas técnicas de rastreamento de movimento, processamento de imagens e inteligência artificial. A ideia central do projeto é, aplicar-se técnicas de rastreamento de movimento para rastrear a mão do usuário e assim capturar a imagem via Kinect que terá background removido e transformado em pixels pretos. Em seguida a imagem é convertida para escala de cinza e aplicado o filtro laplaciano para destacar as bordas contidas na imagem. Para finalizar, é gerado uma matriz de características da imagem resultante e utilizada uma rede neural artificial treinada para reconhecer o gesto.

Para finalizar temos um trabalho que propõe um método simples para reconhecer gestos em LIBRAS em [8]. Este método é composto por três etapas: geração de uma sequência de resíduo, subtraindo quadros adjacentes; obtenção da matriz de características através da acumulação do movimento em células acumuladoras; classificação usando k-Nearest Neighbors (k-NN). Os autores realizaram testes com diversos tamanhos k da matriz de características, começando de 3x3 até 39x39; assim, de acordo com os resultados, concluíram que o tamanho ideal foi k=7, visto que o tamanho desta matriz está diretamente ligada a complexidade do sistema.

As técnicas de reconhecimento por computador das línguas de sinais em geral, e LIBRAS em particular, tiveram uma substancial evolução[11]. Visão computacional se estabeleceu como a abordagem mais utilizada, por ser mais acessível aos usuários. Os algoritmos de detecção de imagens, são mais robustos às variações do ambiente, prescinde de ambientes controlados e é de fácil uso pelos usuários em geral (J.Netom e L.Oquendo,2015).

Após a análise de diversas metodologias, foi possível compreender melhor o problema e como podemos resolvê-lo de diferentes maneiras. Sendo assim, a seguir é detalhada a metodologia escolhida na implementação deste trabalho.

III. METODOLOGIA

Como o foco deste trabalho é algo mais restrito no que se refere aos sinais que serão reconhecidos, a metodologia foi baseada em técnicas mais simples para serem implementadas. A ideia é aplicar os conceitos de Visão Computacional.

Sendo assim, são realizadas as etapas de pré-processamento, segmentação, extração de características e classificação. Portanto, utilizaremos parte das imagens do banco de dados para treino e uma proporção menor para testes.

Existem alguns fatores que diferenciam a implementação adotada neste projeto para os trabalhos analisados para estudo.

Primeiramente não foi explorado em nenhum dos artigos revisados algum que utilizasse o descritor HOG, aqui empregado. Além disso, não há relatos também de trabalhos que tenham aplicado um classificador SVM. Estes conceitos ainda serão destrinchados nesta seção.

A. Base de Dados

Para a execução deste trabalho, o banco utilizado possui imagens estáticas de alguns dos sinais que representam letras do alfabeto da ASL (American Sign Language). Esse conjunto de dados utilizado foi coletado para o trabalho [1] em 2014 e uma extensão do mesmo [2], em 2015.

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. Na figura 1 é ilustrado a forma como são capturadas as imagens.



Figura 1. Captura das imagens para os trabalhos [1] e [2]

Para o projeto que está sendo realizado neste trabalho, o banco é minimizado e construído após as etapas de Pré-Processamento e Segmentação. Utiliza-se sete sinais executados por nove pessoas no reconhecimento. Os sinais são A, D, I, L, V, W e Y conforme estão representados na figura 2 os gestos de 1 a 7, respectivamente. Foi obtido um total de 397 imagens.

B. Pré-Processamento

A fase de pré-processamento tem o objetivo de realizar algum tratamento nas imagens com o intuito de facilitar as etapas adjacentes. Nesse estágio, os pré-processamentos realizados foram a conversão das imagens em escala de cinza e o redimensionamento de todas as imagens para as dimensões da maior imagem do banco (317x317). Após testar o redimensionamento para as dimensões da maior imagem e para um valor médio das dimensões foi percebido que era melhor optar pela primeira opção. Aplicamos a função *imresize* utilizando interpolação bilinear para o redimensionamento. As figuras 3 e 4 ilustram o processo de transformação para escala de cinza.

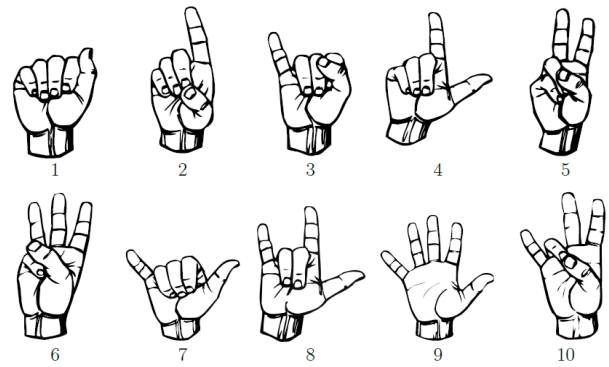


Figura 2. Sinais presentes no banco original

Em nossa implementação essa abordagem de transformação para de escala de cinza é válida, visto que as imagens utilizadas possuem um fundo divergente da área de interesse, no caso a mão. No entanto, em outras situações, como nos casos em que o fundo da imagem é claro e semelhante as intensidades da região da mão, talvez não seja adequado. Além disso, a parte de redimensionamento foi de extrema importância para os resultados que irão ser apresentados posteriormente.



Figura 3. Imagem antes do processamento



Figura 4. Imagem em escala de cinza

C. Segmentação

A parte de segmentação da área da mão neste trabalho não é 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. Além disso, a segmentação da mão nas imagens do banco utilizado poderia se tornar algo complicado, principalmente por causa da semelhança na intensidade de outras partes do corpo com as mãos.

Sendo assim, a segmentação é feita manualmente selecionando parte das imagens de cada sinal e recortando a área de interesse utilizando um software de edição de imagens. As imagens são recortadas em taxas de proporção quadradas, porém com tamanhos diferentes. Esse procedimento está ilustrado na figura 5 a seguir.



Figura 5. Exemplo da segmentação manual

D. Extração de características

Como sabemos objetos em imagens possuem forma, cor e textura. Estas propriedades podem ser mensuradas e tais medidas são denominadas características da imagem. Para extrairmos características específicas de uma imagem utilizamos o que denominamos descritores, que normalmente são representados por um vetor de escalares. É interessante que os descritores sejam invariantes a transformações afins. Utilizasse o descritor HOG(Histogram of Oriented Gradient) neste trabalho. O HOG é um dos descritores de forma mais utilizados e que possui uma boa avaliação, o tamanho do vetor descritor gerado por ele depende do tamanho da imagem.

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. A figura 6 demonstra a imagem de entrada e o descritor.

O algoritmo se baseia na ideia de que a forma e a aparência de um objeto podem ser descritas muitas vezes pela intensidade dos gradientes ou a direção das bordas, sem um conhecimento prévio da posição de tais bordas. O processo para gerar o descritor pode ser dividido em quatro etapas: cálculo do gradiente em cada pixel, agrupamento dos pixels em células, agrupamento das células em blocos e obtenção do descritor.

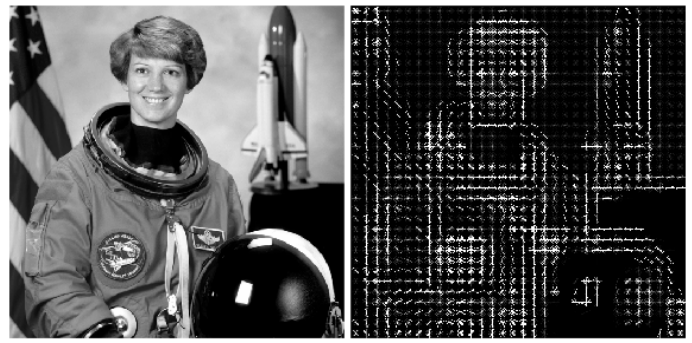


Figura 6. Imagem de entrada e descritor HOG

O primeiro passo é a computação dos gradientes, um estágio crítico na formação do descritor. Dada uma imagem I são calculados os gradientes G_x e G_y em cada pixel de acordo com máscaras derivada discreta pontual tanto no eixo vertical como horizontal.

O passo seguinte é responsável por agrupar os pixels em regiões pequenas conectadas, criando-se o que se chama de células (por exemplo, de tamanho 8 pixels). Após a segunda etapa, os blocos são criados através do agrupamento de células de uma certa região.

Na etapa final, cria-se o descritor que nada mais é do que uma lista dos histogramas de todas as células de todos os blocos. A atenuação do problema das variações locais de iluminação ou de contraste entre o primeiro plano e o plano de fundo, se dá através da normalização de cada histograma de acordo com seus próprios valores. A figura 7 ilustra o passo a passo da obtenção do descritor HOG.

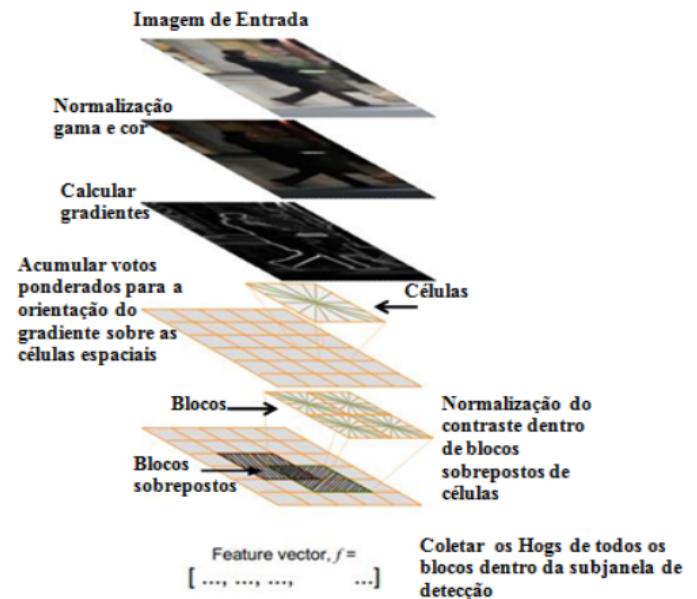


Figura 7. Etapas do HOG

Os descritores de cada imagem gerados neste trabalho apresentam um tamanho fixo, devido ao fato de redimensionarmos

todas as imagens utilizadas. Isso é de suma importância para a etapa de classificação, visto que a matriz de descritores que é passada para o classificador deve ter um tamanho exato.

E. Treinamento e Classificação

No processo de classificação é criado um template SVM(Support Vector Machine) e são utilizados dois métodos, um para treinar e outro para classificar propriamente as imagens dos sinais. São eles: *fitcecoc* e *predict*.

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.

Os parâmetros fornecidos para a função *fitcecoc* 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.

Os parâmetros fornecidos para a função *predict* são o modelo de treino criado pela função *fitcecoc* e o descritor de teste a ser classificado. Na figura 8 a seguir é possível visualizar o código responsável pelo processo descrito.

```
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)
```

Figura 8. Trecho do algoritmo responsável pelo treinamento e classificação

IV. RESULTADOS

Foram realizados testes sem utilizar o pré-processamento de redimensionamento e utilizando-o. Foi possível notar uma melhora significativa nas taxas de reconhecimento ao executar o redimensionamento.

A porcentagem adotada nos testes foi baseada no que foi visto que é mais usual. Dessa forma, 70% das imagens foram selecionadas para treino e 30% para teste a cada execução de maneira aleatória.

Após a realização de 25 testes sem a etapa de redimensionamento, foram selecionados 10 testes para avaliação. O motivo para selecionar apenas 10 foi pelo fato de os testes, apesar de variarem um pouco, ainda apresentaram um comportamento homogêneo. A média total obtida através da taxa de acerto média por sinal foi de apenas 57.93%.

Se considerarmos as melhores taxas de reconhecimento para cada gesto temos apenas 72.40% de acerto, o que ainda podemos considerar um número não muito considerável.

Foi possível perceber uma taxa bem menor de acerto e uma maior confusão por parte do classificador para especificamente os sinais D, V e W. Estes sinais apresentam uma forma um pouco semelhante de serem executados.

No entanto, os testes realizados levando em conta a etapa de redimensionamento das imagens apresentaram bons números. Foram realizados 25 testes e selecionados 10 deles também

nesse caso, pelo mesmo motivo exposto para o caso anterior. Os resultados obtidos podem ser vistos nas tabelas a seguir.

A tabela 1 e gráfico 1 mostram a média de acerto por cada sinal. Foi perceptível a melhora na taxa de acerto média do reconhecimento em relação a anterior, que agora é 85.95%. Ainda há o problema na taxa de classificação dos sinais D, V e W, porém estes agora apresentam um número superior a 70%.

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

Tabela 1. Média de acerto por sinal

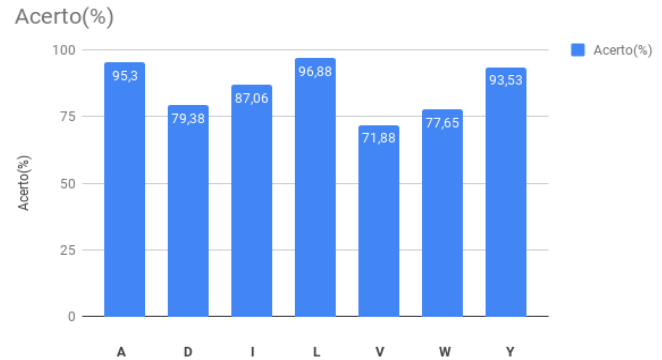


Gráfico 1. Média de acerto por sinal

A tabela 2 e gráfico 2 mostram a média de acerto por cada teste. Os resultados estão na faixa entre 80% e 90% e a taxa média é igual a média por sinais, 85.95%. Ao observar essa tabela não fica tão claro os resultados inferiores que aconteceram nos sinais mencionados anteriormente.

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

Tabela 2. Média de acerto por teste

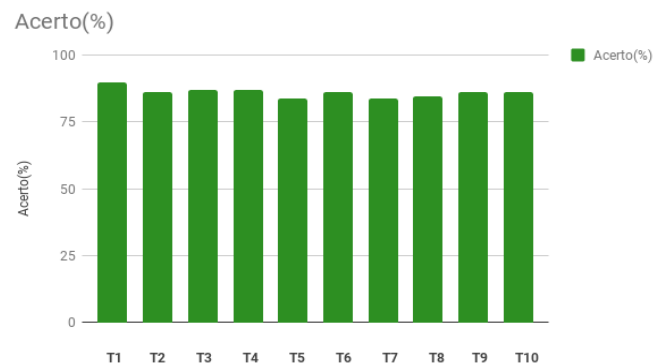


Gráfico 2. Média de acerto por teste

A tabela 3 apresenta o melhor resultado de reconhecimento para cada sinal. Neste caso, os resultados foram todos acima de 87%, sendo que 4 têm 100% e a taxa média é igual a 94.73%. É nítido que houve uma melhora expressiva em relação aos 72.40% obtidos antes do redimensionamento.

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

Tabela 3. Melhores resultados por sinal

A tabela 4 apresenta a matriz de confusão do melhor teste realizado. É possível analisar e reforçar o que já foi afirmado anteriormente em relação aos sinais D, V e W; que em vários testes apresentam confusão entre si.

	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

Tabela 4. Matriz de confusão

Apesar que não ser o maior foco deste projeto, em relação a rapidez o sistema apresentou um tempo de execução razoável. Ainda que não seja um tempo que possa ser aplicado em uma situação em tempo real, ele que gira em torno de 40 segundos a 1 minuto. É importante ressaltar que temos um total de 397 imagens.

CONCLUSÃO

Os resultados obtidos após os testes realizados com o redimensionamento das imagens foram satisfatórios em nossa perspectiva. Além disso, eles poderiam 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 tanto na linguagem de programação utilizada quanto em alguns conceitos de Processamento Digital de Imagens.

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. Além disso, pode-se compreender que uma ferramenta para reconhecimento em tempo real envolveria uma complexidade elevada.

Para a realização de trabalhos futuros um possível incremento a ser realizado neste trabalho seria implementar algum algoritmo para a fase de segmentação. Isso seria interessante para que não houvesse uma necessidade de restringir o banco de dados a imagens que só apresentam a área de interesse.

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, "Fusão de descritores de histogramas de gradientes para a detecção de faces baseado em uma cascata de classificadores", Universidade Federal de Ouro Preto
- [4] A.Silva, 2014, "Descritores de Imagem", 2014.
- [5] P.Filho, S.Furuie, T.Yoneyama, "RECONHECIMENTO AUTOMÁTICO DE SINAIS GESTUAIS (LIBRAS) EMPREGANDO PROCESSAMENTO DE IMAGENS", 2011.
- [6] P.Marcotti, L.Abiuzi, P.Rizol, C.Forst, "Interface para Reconhecimento da Língua Brasileira de Sinais", 2007.
- [7] L.Gonçalves, E.Saad, R. Andrade, B.Romero, R.Duarte, "REDES NEURAIS ARTIFICIAIS E PROCESSAMENTO DE IMAGEM NO RECONHECIMENTO DE LIBRAS, USANDO O KINECT", 2016.
- [8] C.Monteiro, Luiz.Pecoraro, A.Lacerda, A.Corbo, G.Araujo, "Um sistema de baixo custo para reconhecimento de gestos em LIBRAS utilizando visão computacional", 2016.
- [9] A.Gautam, A.Kaushik, "American Sign Language Recognition System Using Image Processing Method", 2017.
- [10] J. Santos, M. Costa, C. Filho, "Reconhecimento das configurações de mão de LIBRAS baseado na análise de discriminante de Fisher bidimensional, utilizando imagens de profundidade", 2015.
- [11] J. Netom, L. Oquendo, "Estudo do estado da arte das técnicas de reconhecimento das línguas de sinais por computador", 2015.
- [12] L. Digiampietri, B. Teodoro, C. Santiago, G. Oliveira, J. Araujo, "Um sistema de informação extensível para o reconhecimento automático de LIBRAS", SBSI 2012.
- [13] E. Flores, J. Barbosa, F. Rigo, "Um estudo de técnicas aplicadas ao reconhecimento da língua de sinais: novas possibilidades de inclusão digital", 2012.
- [14] B.Teodoro, Sistema de Reconhecimento Automático de Língua Brasileira de Sinais, Universidade de São Paulo, 2015.
- [15] M. Zaki, S. Shaheen, "Sign language recognition using a combination of new vision based features", Cairo University, 2010.
- [16] K. Kour, Dr. L. Mathew, "Sign Language Recognition Using Image Processing", Índia, 2017.