

# Reconhecimento do Alfabeto Datilológico da Língua Brasileira de Sinais Utilizando Técnicas de Visão Computacional

Bianka Tallita Passos<sup>1</sup>, Eros Comunello<sup>1</sup>

<sup>1</sup>Escola do Mar, Ciência e Tecnologia – Universidade do Vale do Itajaí (UNIVALI)  
São José – SC – Brasil

biankatpas@edu.univali.br, eros.com@univali.br

**Abstract.** *Communication is an essential factor for the global social interaction of individuals. There are several forms of communication, among them we have the Brazilian Sign Language, a gestural language, whose main objective is to enable the inclusion of deaf people in society. Sign Language is visual and gestural, so its possible to affirm that the use of technology, together with Computer Vision techniques, can contribute to the inclusion of deaf people in society. This paper describes the development of a Computer Vision software that, through the use of digital image processing techniques, performs the recognition of gestures of LIBRAS alphabet, allowing its comprehension by people who do not have knowledge of that language.*

## 1. Introdução

De acordo com o último Censo do IBGE - Instituto Brasileiro de Geografia e Estatística, realizado em 2010, existem mais de 10 milhões de deficientes auditivos no país (IBGE, 2010). Grande parte desta população utiliza a LIBRAS como sua língua principal. A oficialização da LIBRAS como língua utilizada pelos surdos no Brasil ocorreu em 2002, através da Lei n.º 10.436, art. 1º (BRASIL, 2002).

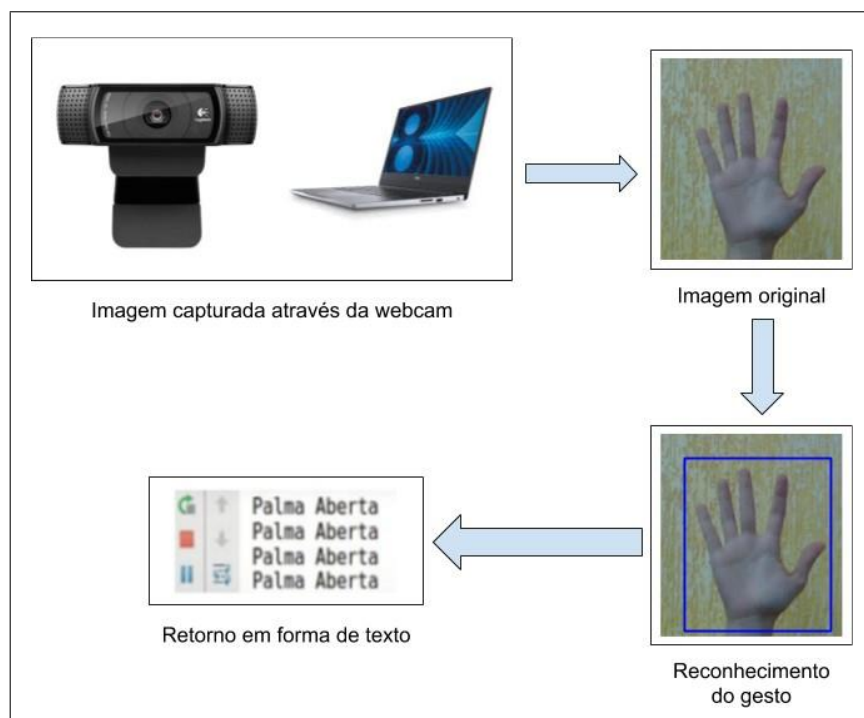
Segundo Pinheiro (2010), a LIBRAS possui inúmeros sinais representativos que correspondem às palavras da língua oral, no entanto algumas palavras ainda não possuem uma representação dentro da LIBRAS e são representadas utilizando os sinais do alfabeto - símbolos de datilologia. Além disso, nomes próprios também são representados utilizando-se a datilologia. Nos últimos anos, foram desenvolvidos trabalhos com esta temática abordando inúmeras técnicas para efetuar o reconhecimento de gestos em imagens, com aplicação no alfabeto das línguas de sinais.

Como exemplo, Pérez et al. (2017), através de uma solução baseada na extração dos sete momentos invariantes de Hu e no treinamento de uma Rede Neural Artificial, obtiveram 91% de acerto no reconhecimento de gestos do alfabeto da Língua Mexicana de Sinais (LSM). Islam et al. (2017) também utilizaram uma Rede Neural Artificial, no entanto, formaram um vetor de características - como fecho convexo, curvatura-K, centróide, entre outros - e chegaram a um resultado de 99,5% de acerto no reconhecimento de gestos do alfabeto da Língua Americana de Sinais (ASL).

## 2. Solução Proposta

A solução proposta neste trabalho é baseada na utilização de aprendizagem de máquina para efetuar o reconhecimento dos gestos do alfabeto da LIBRAS. Além disso, a solução

proposta parte do princípio que o gesto da LIBRAS é um objeto presente na cena, portanto é possível afirmar que a utilização de técnicas de Visão Computacional, comumente aplicadas para efetuar o reconhecimento de objetos em imagens, podem ser úteis para efetuar o reconhecimento dos gestos da LIBRAS. O Fluxograma da solução é apresentado na Figura 1.



**Figura 1. Visão geral da solução proposta**

A técnica de reconhecimento de objetos a ser utilizada é baseada na API de reconhecimento de objetos do TensorFlow - biblioteca aplicável em uma ampla variedade de domínios, como classificação e o reconhecimento de objetos em imagens e amplamente utilizada para aprendizado de máquina (*machine learning*) e pesquisa em redes neurais profundas (*deep learning neural networks*).

Seguindo o exposto em trabalhos correlatos, será criado um *dataset* - contendo os gestos a serem reconhecidos, executados por intérpretes da LIBRAS - que permitirá a extração de características para efetuar o treinamento do modelo de reconhecimento de objetos. Sendo que, 75% das imagens serão utilizadas para treinamento e 25% para testes.

O trabalho está em desenvolvimento, foi criado um algoritmo para efetuar a segmentação da imagem e extrair o objeto de interesse - com base no movimento. Posteriormente, foram realizadas transformações morfológicas com o intuito de remover ruídos e, por fim, foi efetuada a extração do conjunto de características do objeto. Os descritores utilizados foram baseados em atributos geométricos como centróide, momentos de Hu, fecho convexo, a área e o ângulo formado pelo objeto presente na cena.

A partir dos dados extraídos, foi criado um conjunto de regras que permitiu efetuar o reconhecimento de 5 gestos referentes ao alfabeto da LIBRAS e 1 gesto referente a representação dos numerais. A Figura 2 exhibe o resultado do algoritmo criado, demonstrando, respectivamente, o reconhecimento dos seguintes gestos: V, 4, W, J, I e C.

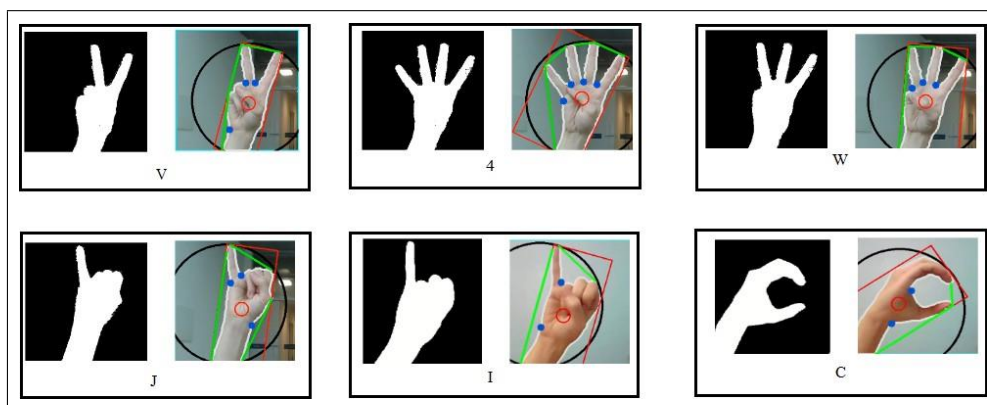


Figura 2. Reconhecimento dos gestos

### 3. Considerações Finais

As soluções propostas pelos trabalhos mais atuais, disponíveis na literatura, são baseadas na extração e combinação de diversas características que, muitas vezes, fornecem pouca capacidade de funcionamento do algoritmo quando exposto a mudanças como rotação, tamanho do objeto e nuances na sua forma.

Sendo assim, buscando garantir que o protótipo tenha maior robustez para se adaptar às possíveis diferenças na execução dos gestos, optou-se por desenvolver uma solução baseada em *deep learning*, utilizando o TensorFlow. Além disso, a literatura carece de estudos relacionados ao uso da biblioteca TensorFlow para efetuar o reconhecimento de gestos em imagens.

Com base nos testes efetuados e no estudo da literatura verificou-se a necessidade de utilizar uma técnica de aprendizado de máquina. A precisão do algoritmo criado será avaliada através da comparação dos resultados obtidos com a utilização do protótipo desenvolvido e os resultados obtidos através da avaliação humana. Os resultados iniciais permitiram comprovar a viabilidade do modelo proposto e a possibilidade de otimização do protótipo desenvolvido.

### Referências

- BRASIL (2002). Lei n.º 10.436, de 24 de abril de 2002. Dispõe sobre a Língua Brasileira de Sinais - LIBRAS - e dá outras providências.
- IBGE (2010). Censo Demográfico de 2010. Fundação Instituto Brasileiro de Geografia e Estatística. <https://censo2010.ibge.gov.br/resultados>. Acesso em 10 mar. 2018.
- Islam, M. M., Siddiqua, S., and Afnan, J. (2017). Real time hand gesture recognition using different algorithms based on american sign language. In: 2017 IEEE International Conference on Imaging, Vision & Pattern Recognition (icIVPR), Dhaka, 2017, pp. 1-6. doi: 10.1109/ICIVPR.2017.7890854.
- Pinheiro, L. M. (2010). Língua de sinais brasileira: libras i. São Paulo: Know How.
- Pérez, L. M., Rosales, A. J., Gallegos, F. J., and Barba, A. V. (2017). Lsm static signs recognition using image processing. In: 14th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE), Mexico City, 2017, pp. 1-5. doi: 10.1109/ICEEE.2017.8108885.