



# Reconhecimento de Sinais de Libras

**Autores: Matheus Campos | mrpcampos@hotmail.com**  
**Matheus da Hora França | mattheushora@gmail.com**

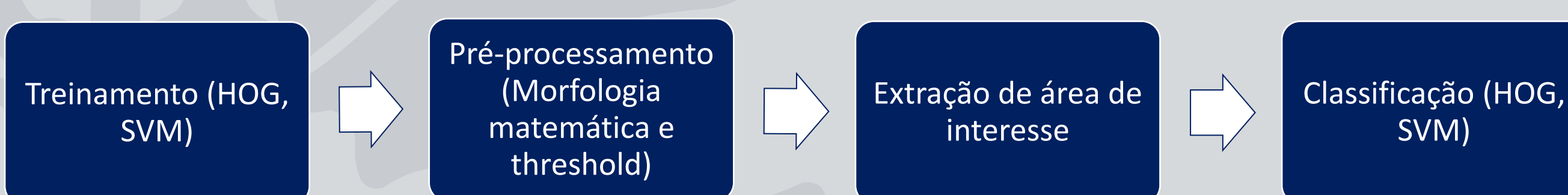
## 1) Introdução

A Língua Brasileira de Sinais (Libras) é a língua de sinais usada por surdos dos centros urbanos brasileiros e legalmente reconhecida como meio de comunicação e expressão [1]. Na língua de sinais são usados como articuladores, além de mãos, outras partes do corpo, como a cabeça, face e tronco [2]. Neste contexto, este trabalho tem como objetivo utilizar técnicas de visão computacional para reconhecer configurações da mão que representam letras e números em libras.

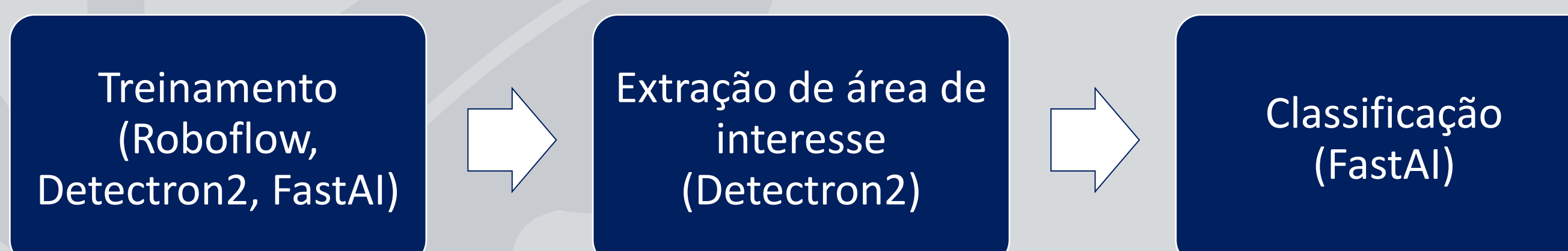
## 2) Metodologia

Este trabalho foi realizado em 2 etapas: implementação da solução com visão computacional clássica e implementação da solução com redes neurais convolucionais (CNN). As etapas seguiram os processos abaixo.

### Visão Computacional Clássica

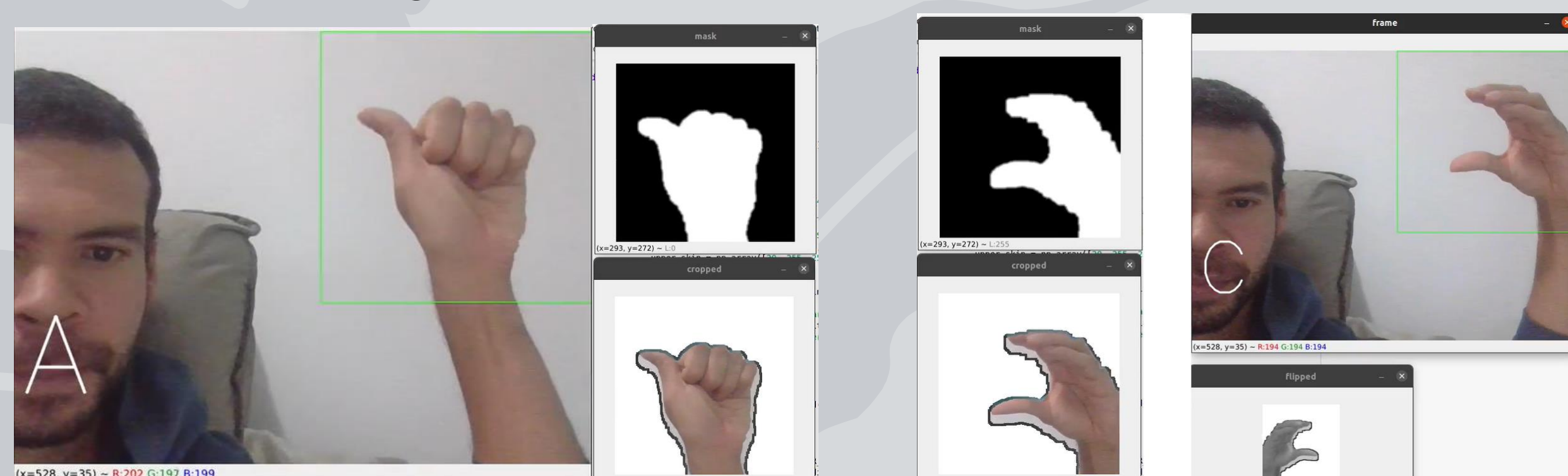


### Redes Neurais Convolucionais



## 3) Visão Computacional Clássica

Nesta solução utilizamos a biblioteca OpenCV [3]. O dataset usado [4] é composto por letras do alfabeto Libras e números, e, para cada um dos sinais, foram usadas 240 imagens com resolução de 50x50 pixels [4]. Na etapa de treinamento utilizamos HOG (*Histogram of Oriented Gradients*) para descrever as características das imagens do dataset e utilizamos estes dados em um SVM (*Support Vector Machine*) para realizar o treinamento. Após o treinamento o SVM é utilizado para classificar as imagens capturadas em tempo real. Neste processo, foi preciso tratar a imagem para extrair a área de interesse, sendo necessário gerar uma máscara que usa como base a faixa de cor da pele em HSV, além de aplicar operações morfológicas para eliminação de ruídos e fechamento de buracos na máscara. Por fim, utilizamos HOG para descrever as características da imagem e SVM para classificar a imagem.

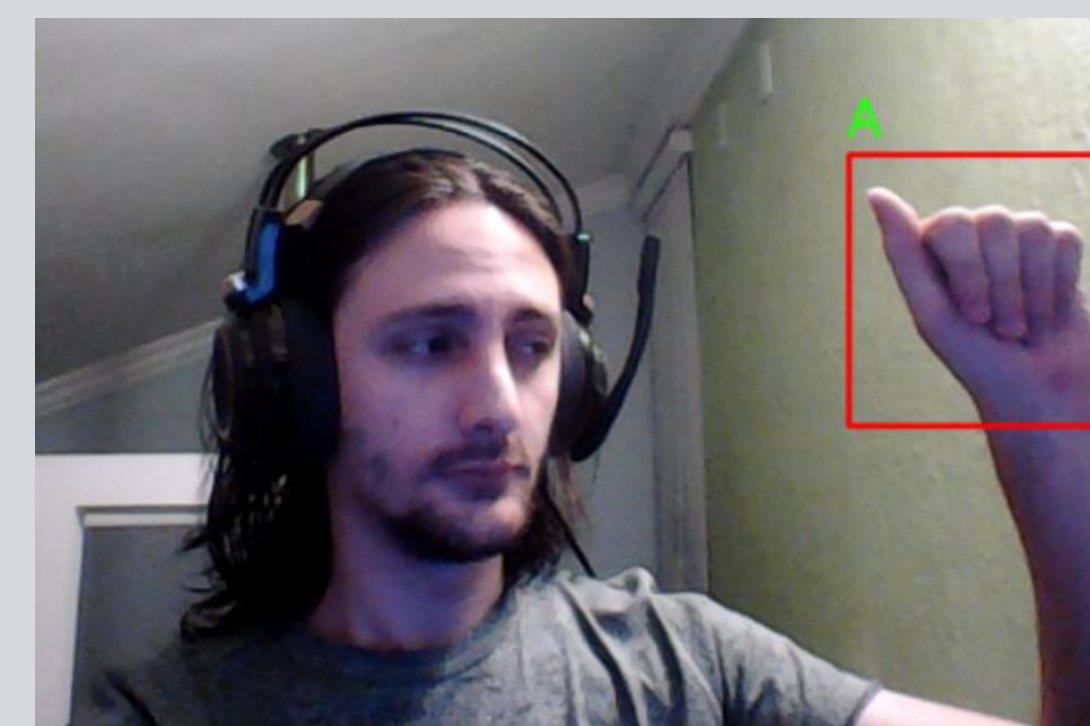


Classificação do sinal A em tempo real

Classificação do sinal C em tempo real

## 4) Redes Neurais Convolucionais (CNN)

Nesta solução, utilizamos um dataset com imagens geradas pelos autores deste trabalho para treinar a detecção da mão do usuário com a biblioteca detectron2 [5]. Nesta etapa, foi preciso utilizar a ferramenta roboflow [6] para descrever as imagens e gerar um COCO dataset compatível com a biblioteca. Assim, utilizamos a detectron2 para detecção e extração da mão na imagem capturada para posterior classificação com a biblioteca FastAI [7]. Na sequência, utilizamos uma ResNet34 para classificação, mas primeiramente foi necessário treinar a rede com o mesmo dataset utilizado na solução clássica acrescido de imagens geradas pelos autores. Após o treinamento, foi possível usar a FastAI para classificar a postura da mão extraída com a detectron2.



Classificação utilizando ResNet34 (FastAI)

epoch	train_loss	valid_loss	error_rate	time
0	2.474249	0.656392	0.168981	00:13
epoch	train_loss	valid_loss	error_rate	time
0	0.514476	0.292767	0.090278	00:17
1	0.310231	0.159780	0.046296	00:16
2	0.199010	0.142393	0.041667	00:16
3	0.128678	0.084821	0.030093	00:16
4	0.079531	0.083310	0.027778	00:16

Resultado de classificação da FastAI

## 5) Resultados e Conclusões

Com o resultado obtido na solução clássica foi possível classificar sinais em tempo real com boa precisão. No entanto, em alguns casos, como por exemplo os sinais "M" e "N" a classificação pode ser confusa. Além disso, essa implementação foi realizada considerando um ambiente controlado, com fundo branco ou escuro, podendo ter a qualidade do resultado impactada pela iluminação. A solução com CNN, por sua vez, apresentou alta taxa de certeza na classificação e é mais robusta que a versão clássica para extrair a mão da imagem, apresentando um melhor resultado em iluminação variável e a fundos ruidosos.

## 6) Referências

- [1] CARVALHO, Paulo Vaz de (2007). Breve História dos Surdos no Mundo. [S.l.]: SurdUniverso.
- [2] Os Cinco Parâmetros da Libras. Disponível em: <<https://www.libras.com.br/os-cinco-parametros-da-libras>>. Acesso em: 23 Maio. 2021.
- [3] OpenCV: OpenCV Tutorials. Disponível em: <[https://docs.opencv.org/master/d9/df8/tutorial\\_root.html](https://docs.opencv.org/master/d9/df8/tutorial_root.html)>. Acesso em: 23 Maio. 2021.
- [4] I. L. O. Bastos, M. F. Angelo and A. C. Loula, "Recognition of Static Gestures Applied to Brazilian Sign Language (Libras)," 2015 28th SIBGRAPI Conference on Graphics, Patterns and Images, Salvador, 2015.
- [5] GitHub - facebookresearch/detectron2: Detectron2. Disponível em: <<https://github.com/facebookresearch/detectron2>>. Acesso em: 23 Maio. 2021.
- [6] Roboflow Dashboard. Disponível em: <<https://app.roboflow.com/>>. Acesso em: 23 Maio. 2021.
- [7] Welcome to fastai | fastai. Disponível em: <<https://docs.fast.ai/>>. Acesso em: 23 Maio. 2021.

Link para solução clássica



Link para solução CNN

