

Traduzindo Libras para Português usando visão computacional

...

TDC Innovation - Florianópolis - 1 a 3 de Junho de 2022





Vitor Casadei

Data Scientist Consultant at CESAR | CTO at Stepps - Masters in Computer Science specialized on Data Science and Machine Learning applied on computer vision.

Sorocaba, São Paulo, Brasil · + de 500 conexões



Stepps



Universidade Federal de São Carlos



Sites

Experiência



CTO

Stepps

Recife e Região



Consultor Cientista de Dados

CESAR

Recife, Pernambuco, Brasil



Deficiência Auditiva

Recorte do Brasil em 2010 (IBGE)

2,1 milhões

deficiência auditiva
severa

9,7 milhões

5% da população
brasileira

141 mill

crianças (0 a 14 anos)

Deficiência Auditiva

Recorte do Brasil em 2019
(Instituto Locomotiva)



Deficiência Auditiva

Recorte mundial em 2019
(Federação Mundial dos Surdos)

50 milhões

pessoas com deficiência
auditiva



1 bilhão

estima-se até o ano de
2050



80%

dos surdos de todo o
mundo possuem baixa
escolaridade e problemas
de alfabetização

Sobre Libras

- **Não** é uma representação da língua falada ou escrita!



- **Não** é iconicidade!



- **Não** se limita ao alfabeto manual!



- **Não** é uma língua universal



- É uma **língua reconhecida** do Brasil



- **Não** é a segunda língua oficial do Brasil



- Possui sua **própria** estrutura gramatical

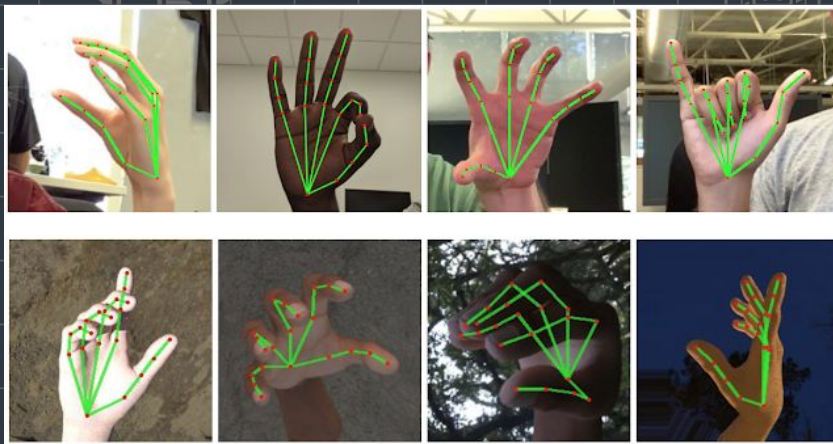


Libras envolve movimentos manuais e expressões faciais





- MediaPipe Hands
 - 21 landmarks
 - multi-hand support
 - 3D model detection



Biblioteca



MediaPipe

- MediaPipe Hands

```
import cv2
import mediapipe as mp
mp_drawing = mp.solutions.drawing_utils
mp_drawing_styles = mp.solutions.drawing_styles
mp_hands = mp.solutions.hands
```

Biblioteca



MediaPipe



- MediaPipe Hands



```
with mp_hands.Hands(static_image_mode=True, max_num_hands=2) as hands:
    image = cv2.flip(cv2.imread(file), 1)
    results = hands.process(cv2.cvtColor(image, cv2.COLOR_BGR2RGB))
    image_height, image_width, _ = image.shape
    annotated_image = image.copy()
    for hand_landmarks in results.multi_hand_landmarks:
        mp_drawing.draw_landmarks(
            annotated_image,
            hand_landmarks,
            mp_hands.HAND_CONNECTIONS,
            mp_drawing_styles.get_default_hand_landmarks_style(),
            mp_drawing_styles.get_default_hand_connections_style())
    cv2.imwrite('image.png', cv2.flip(annotated_image, 1))
```





- MediaPipe Holistic
 - 21 per-hand landmarks
 - 33 pose landmarks
 - 468 facial landmarks
 - 3D model detection



Biblioteca



MediaPipe

- MediaPipe Hands

```
import cv2
import mediapipe as mp
mp_drawing = mp.solutions.drawing_utils
mp_drawing_styles = mp.solutions.drawing_styles
mp_holistic = mp.solutions.holistic
```



- MediaPipe Holistic



```
with mp_holistic.Holistic(static_image_mode=True,model_complexity=2) as holistic:
    image = cv2.imread(file)
    results = holistic.process(cv2.cvtColor(image, cv2.COLOR_BGR2RGB))
    annotated_image = image.copy()

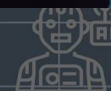
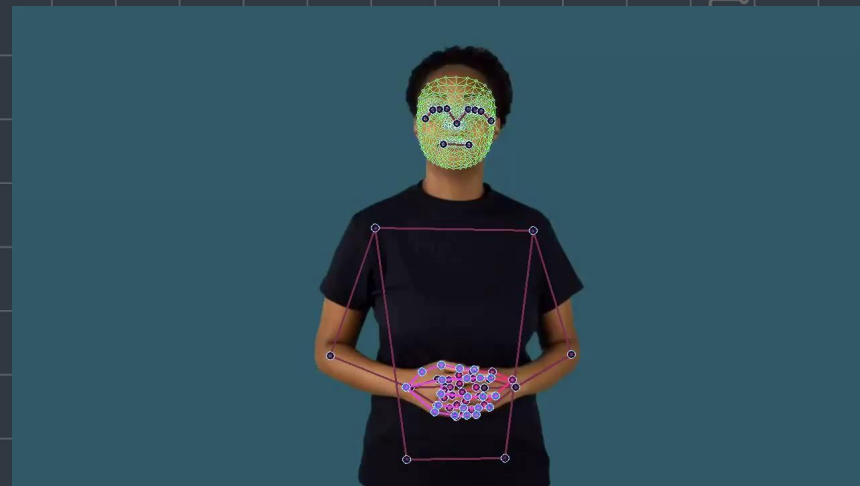
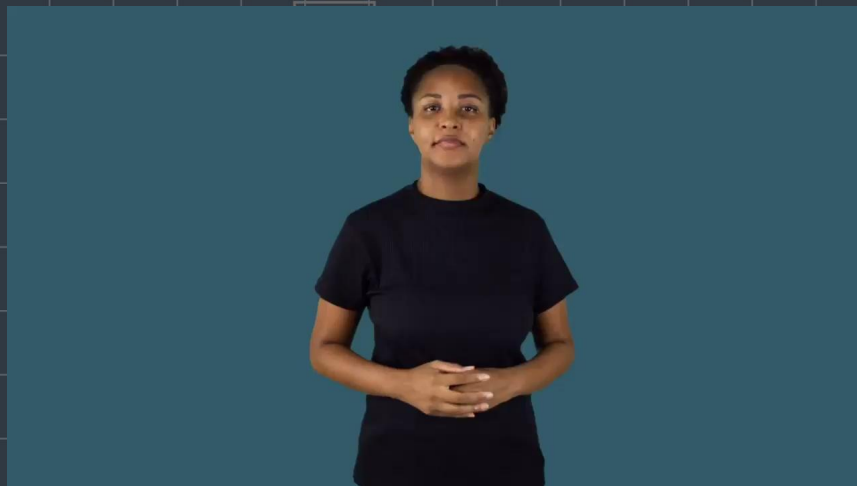
    ...

    annotated_image = np.where(condition, annotated_image, bg_image)
    mp_drawing.draw_landmarks(annotated_image, results.face_landmarks,
                              mp_holistic.FACEMESH_TESSELATION, landmark_drawing_spec=None,
                              connection_drawing_spec=mp_drawing_styles
                              .get_default_face_mesh_tesselation_style())
    mp_drawing.draw_landmarks(annotated_image, results.pose_landmarks,
                              mp_holistic.POSE_CONNECTIONS, landmark_drawing_spec=mp_drawing_styles
                              .get_default_pose_landmarks_style())
    cv2.imwrite('image.png', annotated_image)
```



Coding

<https://github.com/vcasadei/tdcinnovation2022>

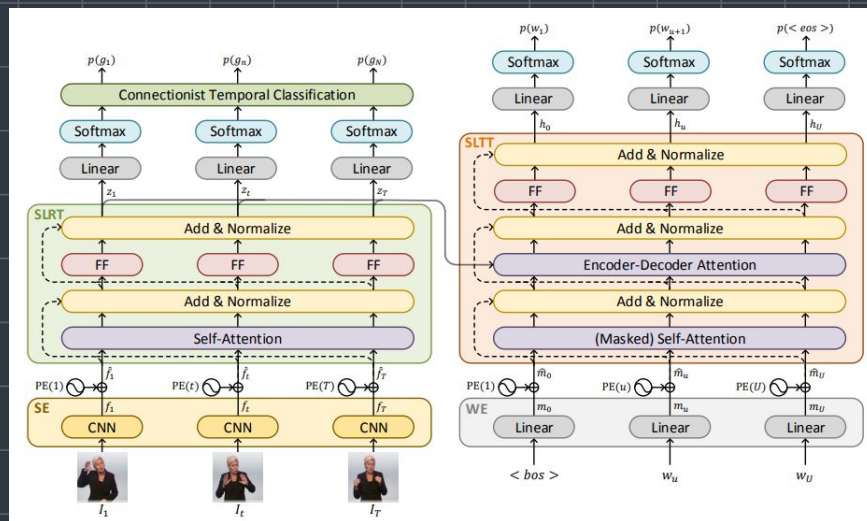




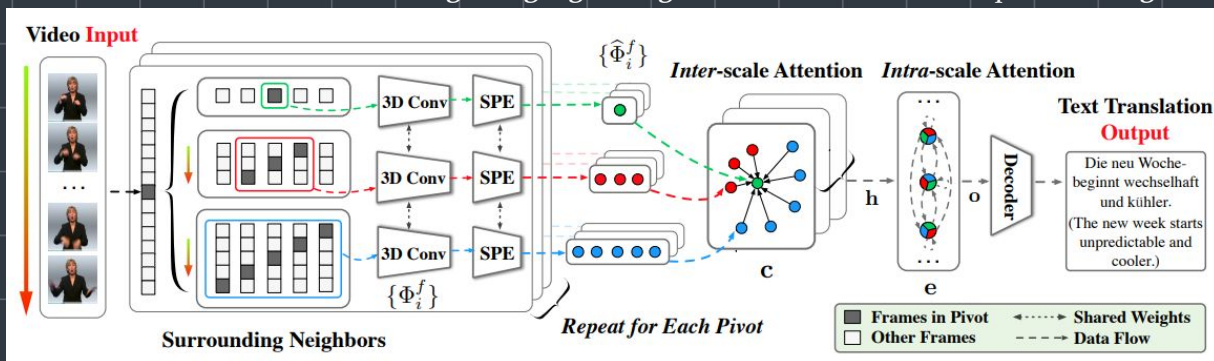
Limitações



Outras possibilidades



Camgoz, N. C., Koller, O., Hadfield, S., & Bowden, R. (2020). Sign language transformers: Joint end-to-end sign language recognition and translation. <https://arxiv.org/abs/2003.13830>



Li, D., Xu, C., Yu, X., Zhang, K., Swift, B., Suominen, H., & Li, H. (2020). TSPNet: Hierarchical feature learning via temporal semantic pyramid for sign language translation. <https://arxiv.org/abs/2010.05468>



Vitor Casadei

Email: vitor.casadei@gmail.com / vc@cesar.org.br

LinkedIn: [linkedin.com/in/vitorcasadei/](https://www.linkedin.com/in/vitorcasadei/)