

Equipe 2 - Visão Computacional em Linguagem de Sinais

Alunos:

Carlos Matheus De Lima Ferreira

Jonas Leite De Medeiros Junior

Juan Paulo Marques Neto

Juliana Mozaga Neves

Lucas Leandro Dos Santos Almeida

Matheus Vieira Barros

Rafael Ferretto Londero

Rebeca Carneiro Cordeiro

Sumário

- Introdução
- Metodologia
- Resultados
- Conclusões

Introdução

Assistentes de voz capacitados abrem as portas para uma série de recursos úteis, muitas vezes transformadores, em dispositivos modernos.

No entanto, essas tecnologias frequentemente não são acessíveis para pessoas surdas. Embora a língua de sinais seja muito mais abrangente do que apenas digitar, a digitação é comumente usada para inserir informações como nomes, endereços e números de telefone em dispositivos móveis.

A tecnologia de reconhecimento de linguagem de sinais para entrada de texto ainda está bastante atrasada em comparação com o reconhecimento de voz para texto.

Introdução

Essa nova tecnologia ofereceria aos usuários surdos e com deficiência auditiva a opção de soletrar palavras com gestos em vez de utilizar um teclado tradicional.

Além de possibilitar a entrada conveniente de texto para buscas na web, direções em mapas e mensagens de texto.

Introdução

Objetivo geral:

Realizar uma tradução de linguagem de sinais ASL para texto, usando Machine Learning para classificação.

Metodologia

INFORMAÇÕES ESPERADAS

- Base de dados;
 - Alfabeto ASL, contém 3.000 imagens de 200x200 pixels por classe = 29 classes

Das quais 26 são para as letras A a Z e 3 classes para SPACE, DELETE e NOTHING



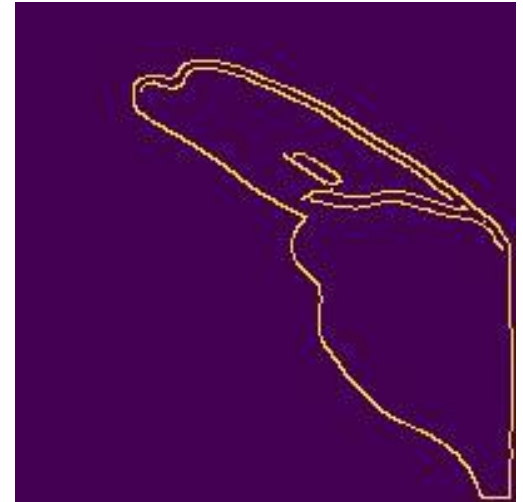
Metodologia

INFORMAÇÕES ESPERADAS

- Pré processamento – Exemplo letra H
 - Imagem original



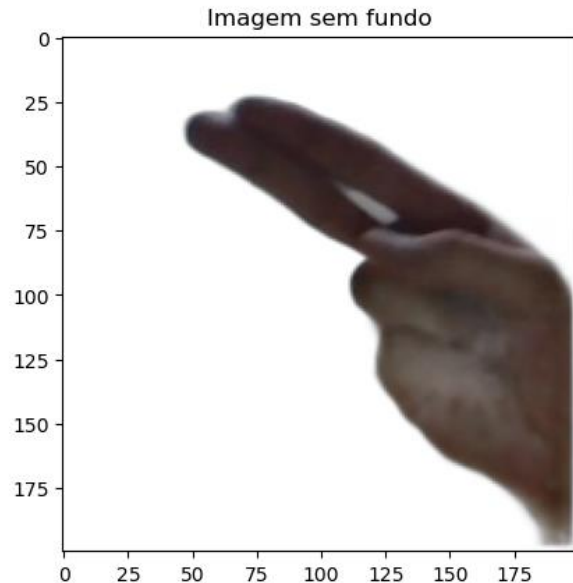
Imagem processada



Resultados

INFORMAÇÕES ESPERADAS

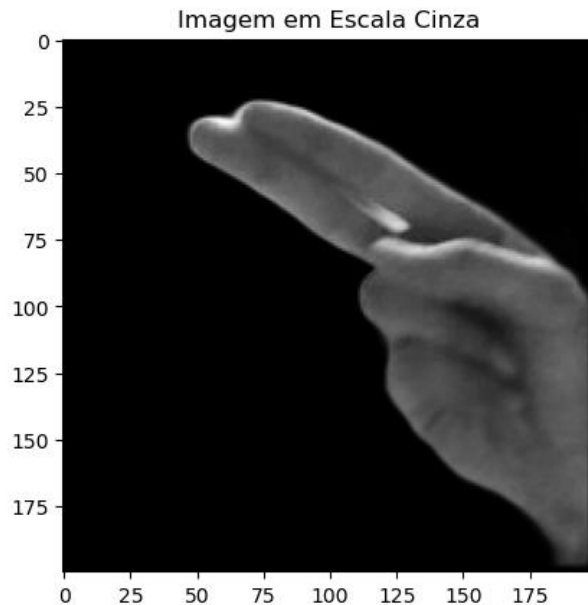
- Apresentar tabelas/figuras para ilustrar métricas de avaliação por classe
- Métricas
 - Remoção de fundo: Remover o fundo da imagem para reduzir a complexidade do aprendizado pelo modelo, dada a variabilidade de cores, contraste e iluminação entre as imagens.



Resultados

INFORMAÇÕES ESPERADAS

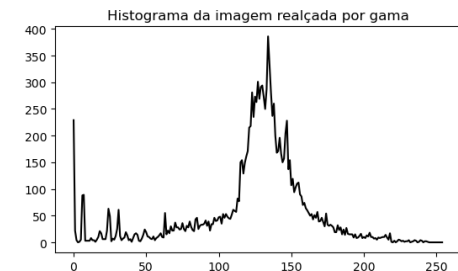
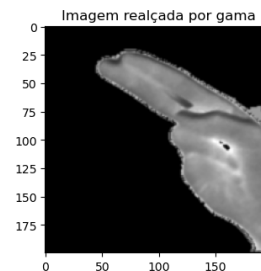
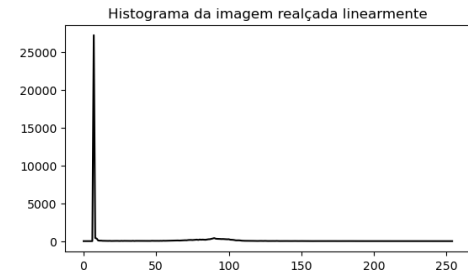
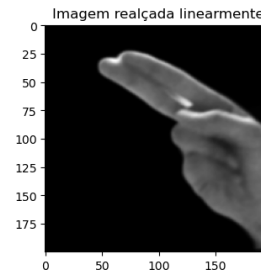
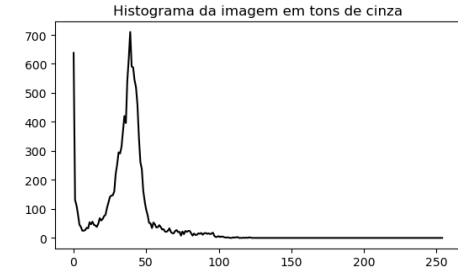
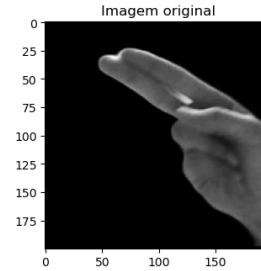
- Apresentar tabelas/figuras para ilustrar métricas de avaliação por classe
- Métricas
- Conversão para escala de cinza: Como a cor não é uma informação relevante para detecção do gesto, as imagens serão convertidas para escala de cinza.



Resultados

INFORMAÇÕES ESPERADAS

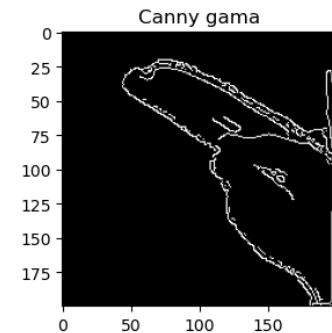
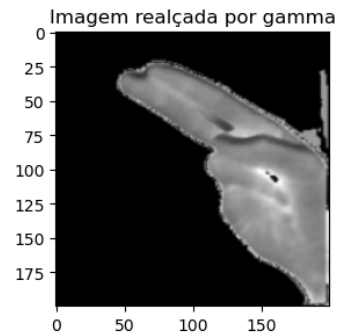
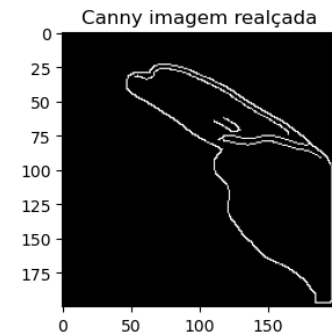
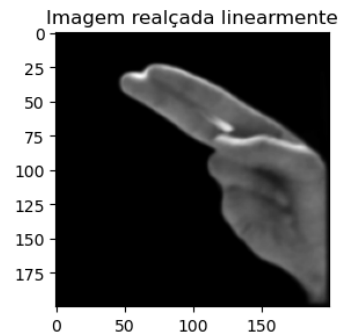
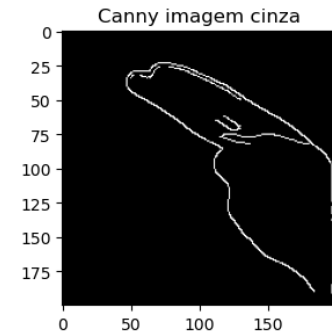
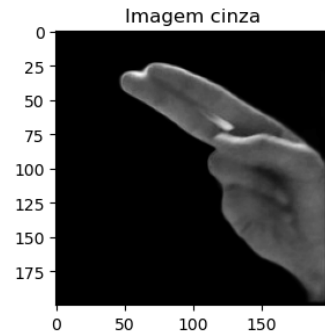
- Apresentar tabelas/figuras para ilustrar métricas de avaliação por classe
- Métricas
 - Realce e ajuste de iluminação: aplicação de maior contraste nas cores da imagem para aprimorar o facilitar a detecção da posição dedos na frente da palma da mão e ajuste da variação de iluminação entre imagens.



Resultados

INFORMAÇÕES ESPERADAS

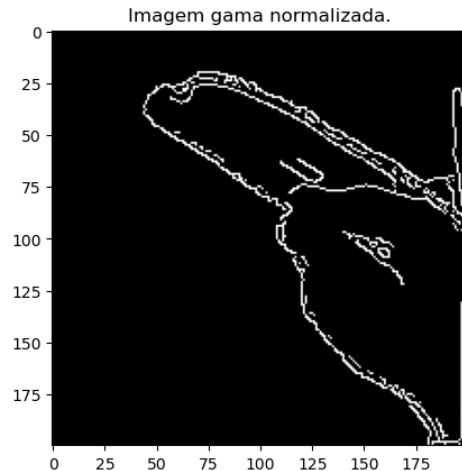
- Apresentar tabelas/figuras para ilustrar métricas de avaliação por classe
- Métricas
 - Suavização para redução de ruídos: Aplicar filtros de suavização para redução de ruídos afim de facilitar a detecção das bordas.



Resultados

INFORMAÇÕES ESPERADAS

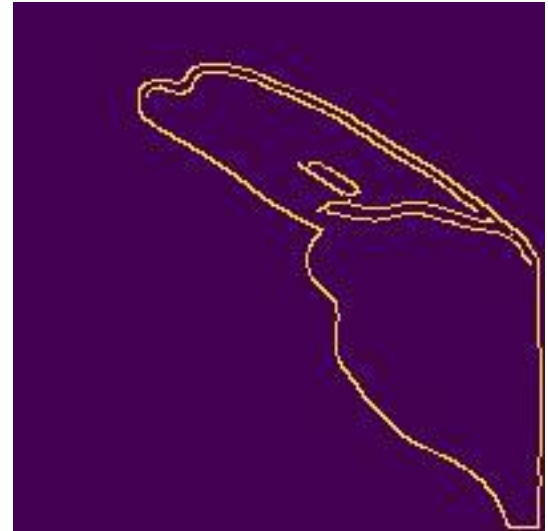
- Apresentar tabelas/figuras para ilustrar métricas de avaliação por classe
- Métricas
 - Normalização: Normalizar os pixels da imagem.



Resultados

INFORMAÇÕES ESPERADAS

- Apresentar tabelas/figuras para ilustrar métricas de avaliação por classe
- Imagem já processada:



Conclusões

- Erros do TensorFlow obtidos:
KeyError u1 e f4, possível incompatibilidade entre os dados do opencv e keras.
- Possibilidades de novos estudos, terminando e padronizando o modelo, e realizando deploy em um web app, ou em um aplicativo, para realizar tradução em tempo real.

Referências Bibliográficas

@misc{https://www.kaggle.com/grassknotted/aslalphabet_akash_nagaraj_2018,
title={ASL Alphabet},
url={https://www.kaggle.com/dsv/29550},
DOI={10.34740/KAGGLE/DSV/29550},
publisher={Kaggle},
author={https://www.kaggle.com/grassknotted/aslalphabet and Akash Nagaraj},
year={2018}}