

Simpósio de Integração Acadêmica

“Ciências Básicas para o Desenvolvimento Sustentável”

SIA UFV 2023



Coleção e Anotação de Imagens Centradas no Apresentador: Aprimorando a Acessibilidade para Pessoas com Deficiência Visual

Luísa Ferreira*†, Daniel Fernandes‡, Fabio Cerqueira‡, Marcos Ribeiro‡, e Michel Silva‡

†Departamento de Informática, Universidade Federal de Viçosa - UFV, Viçosa, Brasil.

‡Departamento de Engenharia de Produção, Universidade Federal Fluminense - UFF, Petrópolis, Brasil
{luisa.ferreira, daniel.louzada, marcosh.ribeiro, michel.m.silva}@ufv.br, frcerqueira@id.uff.br

Palavras chaves: Inteligência artificial, Conjuntos de dados, Tecnologias assistivas

Categoria: Pesquisa; Grande Área: Ciências Exatas e Tecnológicas; Área temática: Inteligência Artificial

Introdução

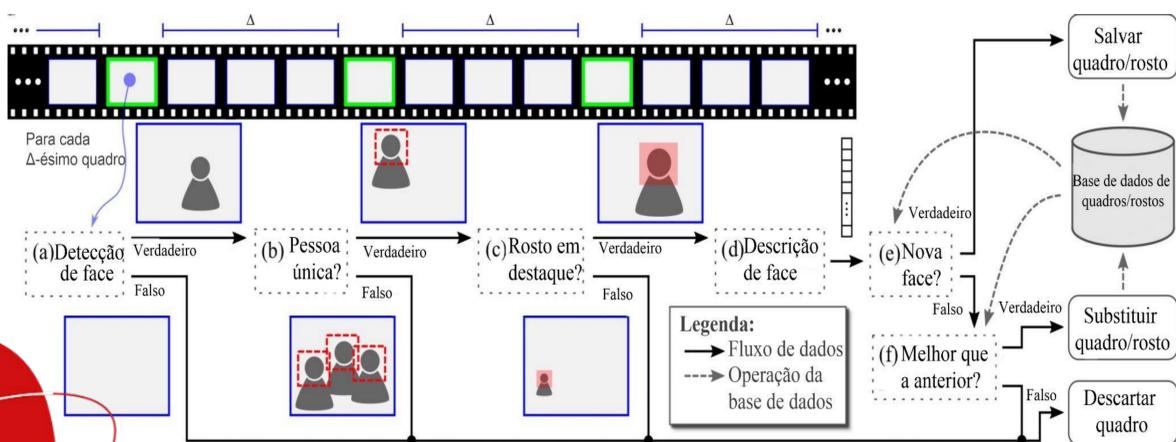
Com a pandemia de COVID-19, nossa dependência da tecnologia aumentou significativamente. Isso levou ao crescimento das conferências virtuais como substitutas dos encontros presenciais, resultando em maior exclusão dos deficientes visuais devido à perda de informações não visuais para contextualização das cenas. Embora haja interesse em utilizar IA para automatizar a criação de descrições contextuais, ainda há carência de dados suficientes para treinamento.

Objetivos

O principal objetivo deste estudo é formar um conjunto de dados com imagens focadas no apresentador. Para alcançar isso, propomos uma abordagem de coleta automática de dados e um protocolo de anotação específico para esse público, visando apoiar o desenvolvimento de sistemas de Tecnologia Assistiva.

Material e Método

A coleta envolveu vídeos públicos do YouTube, abrangendo jornais, podcasts, aulas e conferências virtuais. Utilizamos técnicas de detecção e descrição facial, analisando cada Δ -ésimo quadro para identificar rostos e verificar sua relevância. A figura abaixo ilustra o processo detalhado.



Resultados e Discussão

No final do estudo, foram coletadas 10.939 imagens focadas no apresentador. O dataset original foi dividido em dois conjuntos: um com 5.689 imagens de indivíduos únicos, levando em conta repetições de apresentadores fixos, e outro com 967 imagens rotuladas conforme um protocolo específico para atender deficientes visuais. Uma análise de diversidade foi conduzida no dataset, abrangendo gênero, raça, emoção, vestimenta, cenário e acessórios. Utilizamos pelo menos dois modelos de classificação existentes para evitar viés e os resultados indicaram uma satisfatória diversidade no conjunto de dados proposto.



Conclusões

Este trabalho propôs um método automático de coleta de imagens centradas no apresentador, com finalidade de auxiliar em sistemas de acessibilidade para deficientes visuais. Além disso, foram apresentadas três versões do dataset, destacando sua qualidade e diversidade.

Agradecimentos

Os autores gostariam de agradecer à CAPES, FAPEMIG, PIBIC-UFV/CNPq, CNPq e BICJr-UFV/FAPEMIG pelo financiamento de diferentes partes deste trabalho; e aos nossos incansáveis anotadores Allan Lopes, Júlia Vieira, Júlia Lopes e Sophia Jorge.