

Transcrição de Imagens para Áudios

Arthur Francisco Ramos & Renato Leite Camilo
Aprendizado Profundo | Programa de Pós-Graduação em Ciência da Computação

Introdução

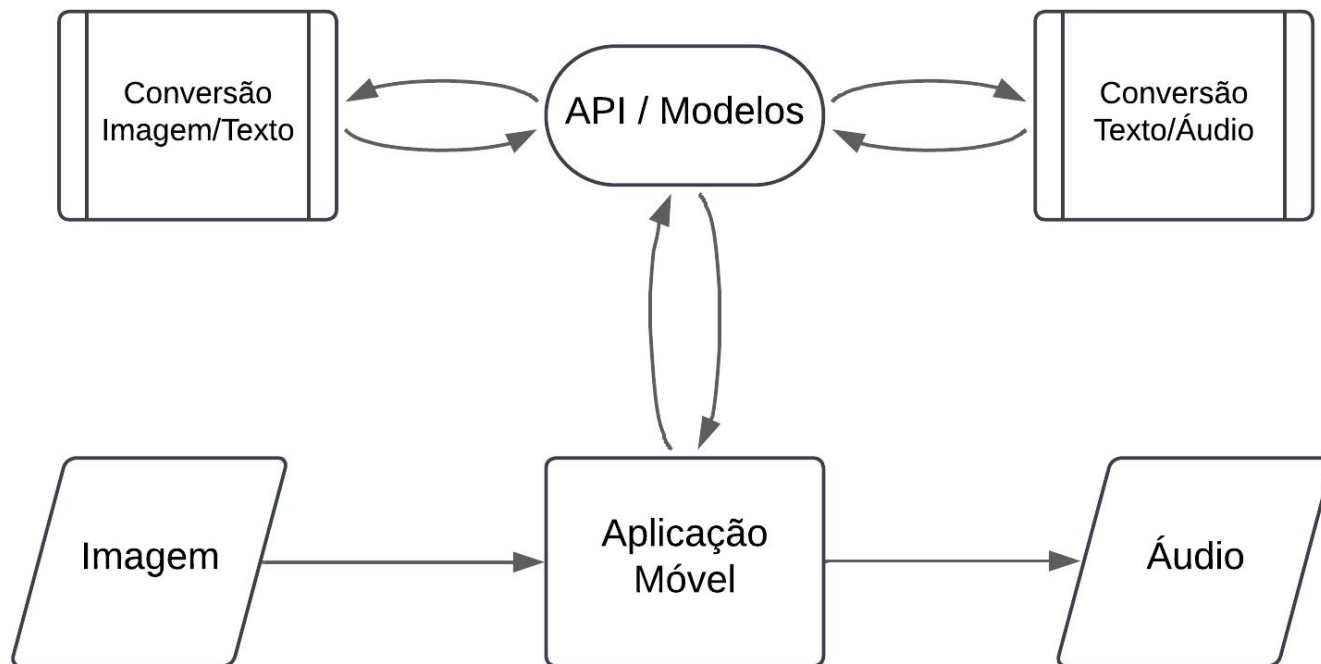
Objetivos

— — —

- Transcrição de **imagens** para **áudio**.
- Permitir uma **maior acessibilidade** para pessoas com alguma dificuldade visual.
- **Ampliar o acesso** de elementos visuais.
- **Avaliar e combinar técnicas/modelos** de *Deep Learning*.

Metodologia

— — —



Ferramentas Utilizadas

— — —

SOFTWARES:

- *Visual Studio Code*

LINGUAGENS:

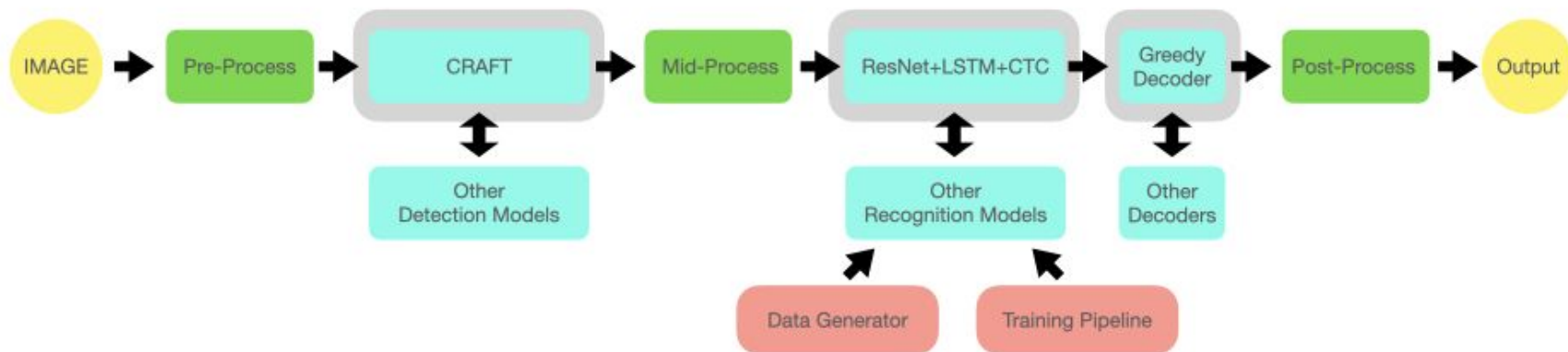
- *Dart/Flutter*
- *Python/Jupyter/Flask*

Imagem para Texto

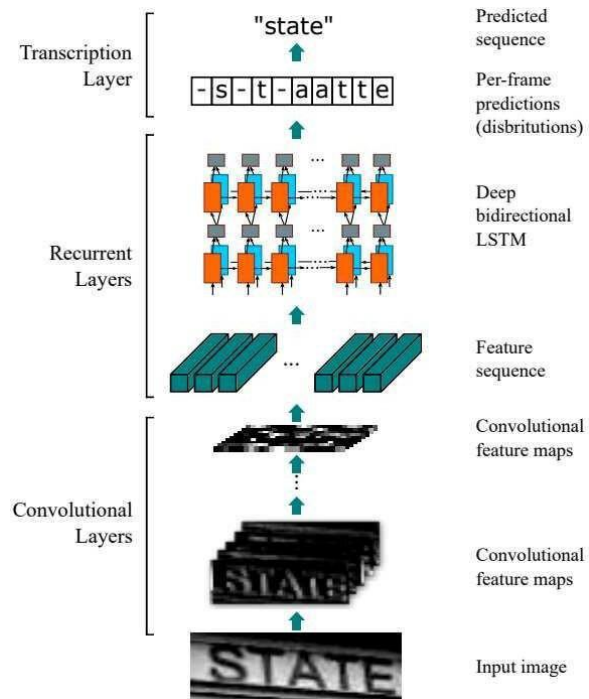
EasyOCR

— — —

EasyOCR Framework

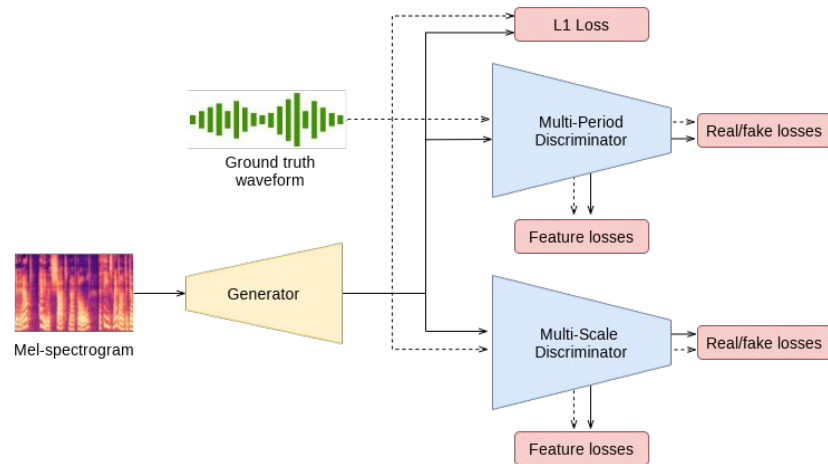
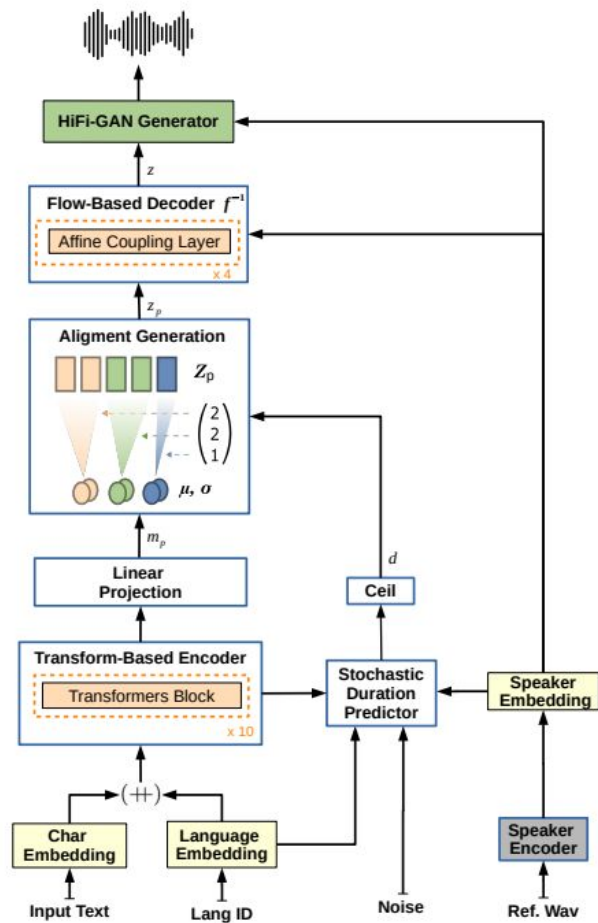


Tesseract



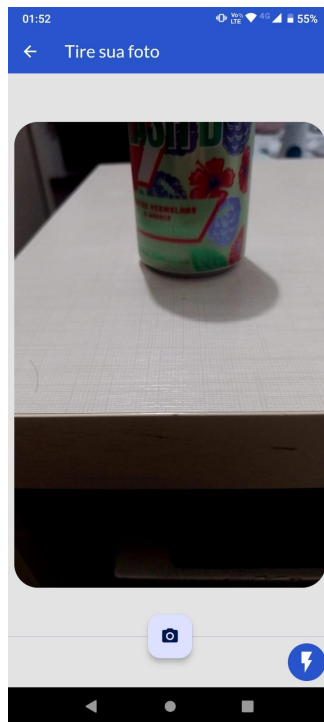
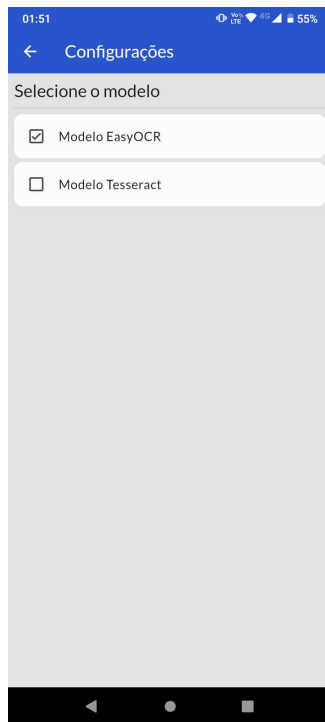
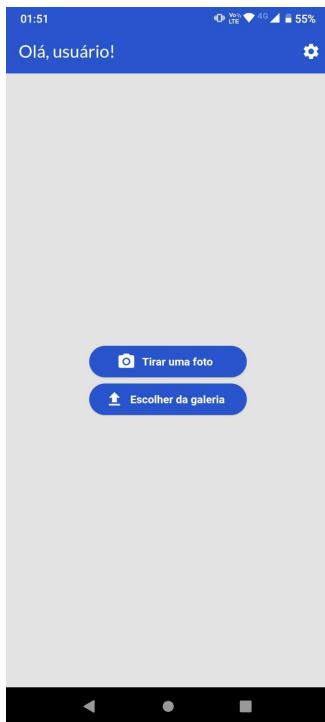
Texto para Áudio

YourTTS



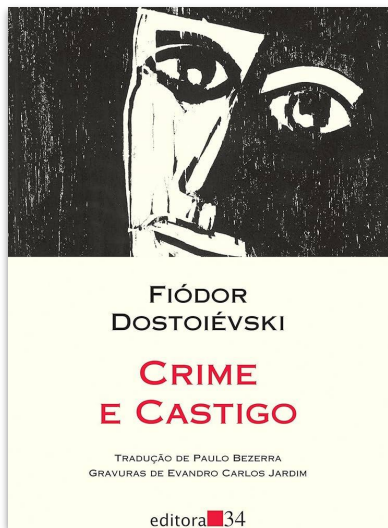
Aplicação e Experimentos

Aplicação



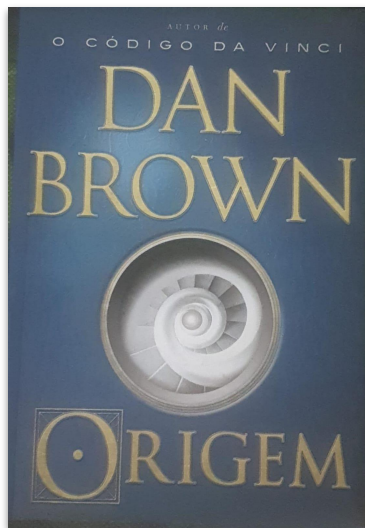
Experimentos

— — —



EasyOCR 

Tesseract 



EasyOCR 

Tesseract 



EasyOCR 

Tesseract 

Conclusão

Resultados

— — —

- **Necessita de recursos**, como *hardware*, para uma melhor qualidade de “transcrição”.
- **Fontes diferentes e imagens de baixa qualidade** interferem no resultado.
- O **EasyOCR** apresentou **resultados melhores** em todos os experimentos, incluindo os três apresentados anteriormente.
- Pode ser aprimorado, como um **trabalho futuro**, para “transcrições” em tempo real de vídeos.

Licença

Copyright (c) 2024 Arthur Francisco Ramos & Renato Leite Camilo

Permissão é concedida, gratuitamente, a qualquer pessoa vinculada à Universidade Estadual Paulista "Júlio de Mesquita Filho" (Unesp), incluindo alunos, funcionários e colaboradores, para usar, copiar, modificar e distribuir o software e sua documentação associada, sem restrição, incluindo, sem limitação, os direitos de usar, copiar, modificar, mesclar, publicar, distribuir, sublicenciar e/ou vender cópias do software, desde que os seguintes termos sejam atendidos:

1. Este software é fornecido "no estado em que se encontra", sem garantia de qualquer tipo, expressa ou implícita, incluindo mas não se limitando às garantias de comercialização, adequação a um propósito específico e não violação. Em nenhuma circunstância os autores ou titulares de direitos serão responsáveis por qualquer reivindicação, dano ou outra responsabilidade, seja em uma ação de contrato, ato ilícito ou de outra forma, decorrente de ou relacionada ao uso ou operação deste software.
2. O uso e modificação do software para fins acadêmicos, administrativos ou de pesquisa dentro do ambiente da Unesp é garantido, sem custos e ilimitado. Para uso fora deste escopo, recomenda-se verificar os termos desta licença.
3. Quaisquer redistribuições do código ou de suas modificações devem incluir este aviso de licença e a atribuição ao(s) autor(es) original(is).

Arthur Francisco Ramos & Renato Leite Camilo, 2024

Referências

EasyOCR

Tesseract

YourTTS

HiFi GAN

— — —