**các phương pháp làm đã có :**

**OCR engines ABBYY Finereader, OCRopus, Tesseract, and Calamari,**

**Calamari**

[**https://github.com/Calamari-OCR/calamari**](https://github.com/Calamari-OCR/calamari)

**implements a deep CNN-LSTM network structure**

**Tesseract**

[**https://github.com/tesseract-ocr**](https://github.com/tesseract-ocr)

**OCR engine based on LSTM neural networks and a wide variety of trained mixed models.**

[**https://github.com/JaidedAI/EasyOCR?fbclid=IwAR23teXxV2uCIRA37GV7FCvblEo40hN3M1JlP3tfQDUkxojDDQ9DTj6mj8Y**](https://github.com/JaidedAI/EasyOCR?fbclid=IwAR23teXxV2uCIRA37GV7FCvblEo40hN3M1JlP3tfQDUkxojDDQ9DTj6mj8Y)

**OCT sử dụng deeplearning ( CNN LSTM CTC loss) hỗ trợ cho 40 ngôn ngữ có cả tiếng việt . Nó có 3 option decoder greedy, beam-search và word beam-search. Repo này đã được so sánh với Tessaract thì độ chính xác tốt hơn nhiều nhưng time chậm hơn chút**

[**https://colab.research.google.com/github/vistec-AI/colab/blob/master/easyocr.ipynb#scrollTo=d\_OUXPmfvNkh**](https://colab.research.google.com/github/vistec-AI/colab/blob/master/easyocr.ipynb#scrollTo=d_OUXPmfvNkh)

**Khó khăn :**

**Dữ liệu có nhiều font chữ, sự giống nhau của các loại font chữ, chất lương hình ảnh quét**

**The self-trained Calamari models achieve the best results, outperforming ABBYY by 70% without voting and even by 78% with voting averaged over all 20 datasets yielding an average CER below 1%**

[**https://github.com/chongyangtao/Awesome-Scene-Text-Recognition**](https://github.com/chongyangtao/Awesome-Scene-Text-Recognition)