Tesseract:

<https://guiem.medium.com/how-to-train-tesseract-4-ebe5881ff3b7>

<https://www.endpoint.com/blog/2018/07/09/training-tesseract-models-from-scratch>

<http://crblpocr.blogspot.com/2013/04/banglaocr-bangla-ocr-bengali-ocr.html#:~:text=However%2C%20it%20appears%20that%20Tesseract,Bangla%20%2F%20Bengali%20for%20version%203.&text=However%2C%20yet%20people%20are%20interested,characters%20or%20combined%20character%20images>.

<http://palashray.com/training-tesseract-4-x-lstm-for-bengali/>

<http://dspace.bracu.ac.bd/xmlui/bitstream/handle/10361/166/Integration%20of%20Bangla%20script%20recognition%20support%20in%20OCRopus.pdf?sequence=1&isAllowed=y>

Training Tesseract for Bengali: <https://github.com/paawak/blog/tree/master/docs/tesseract/training-bengali>

Training Documentation: <https://tesseract-ocr.github.io/tessdoc/tess4/TrainingTesseract-4.00.html>

<https://github.com/tesseract-ocr/tesstrain>

<https://github.com/tesseract-ocr/langdata>

Dataset:

<https://www.kaggle.com/c/bengaliai-cv19/discussion/122604>

<https://www.kaggle.com/ipythonx/137x236>

<https://www.kaggle.com/ipythonx/keras-grapheme-gridmask-augmix-ensemble>

Bangla OCR paper:

<https://www.researchgate.net/publication/329048432_BornoNet_Bangla_Handwritten_Characters_Recognition_Using_Convolutional_Neural_Network>

<https://www.researchgate.net/publication/329606551_Bangla_Handwritten_Digit_Recognition_Using_Convolutional_Neural_Network>

<https://www.researchgate.net/publication/326989744_Unconventional_Wisdom_A_New_Transfer_Learning_Approach_Applied_to_Bengali_Numeral_Classification>

<https://www.researchgate.net/publication/329391937_Bangla_Handwritten_Digit_Recognition_Using_Deep_CNN_for_Large_and_Unbiased_Dataset>

<https://www.researchgate.net/publication/333618625_Bangla_Handwritten_Character_Recognition_using_Convolutional_Neural_Network_with_Data_Augmentation>

<https://www.researchgate.net/publication/336287494_Basic_to_Compound_A_Novel_Transfer_Learning_Approach_for_Bengali_Handwritten_Character_Recognition>

Issues:

1. Official documentation for windows training not available
2. Have to install tesseract engine separately to be able to run on machine
3. Requires .tiff images
4. Requires box files containing annotations and labels\*\*\*
5. Requires .lstmf files that combines each image with its box file
6. Requires charset file
7. Does not have GUI annotation tool