On-Device Machine Learning

Team Members: Collin Peirce, Cole Riechert
Advisor: Dr. Ryan Kim

Department of Electrical and Computer Engineering

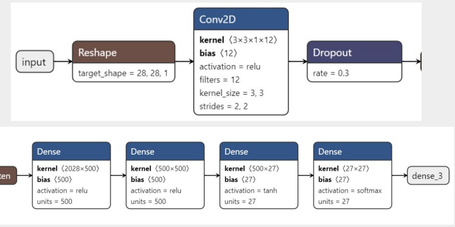
Colorado State University

Fort Collins, CO 80523

Optical Character Recognition

Optical Character Recognition, or OCR, is the process by which handwritten characters are parsed into machine-encoded text.

- This is done using a CNN built in Keras
- Model was trained on EMNIST, a dataset of over 146,000 handwritten letters
- Model's greatest strength is its 98% top-3 classification accuracy

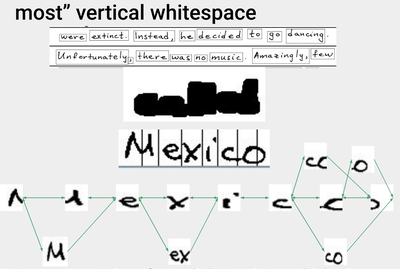


Block diagram for OCR model structure; bottom image continues from the left from the top image.

Segmentation

Segmentation is the step where lines, words, and letters are separated from each other for processing.

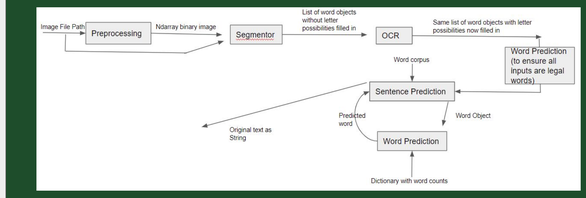
- Lines are separated by looking for horizontal white space
- Words are separated by increasing the thickness of lines and seeing what is touching
- Letters are separated by looking for "the most" vertical whitespace



Letter segmentation often split letters into multiple pieces. This data structure, a graph, is how we kept track of how these pieces could be connected.

Project Overview

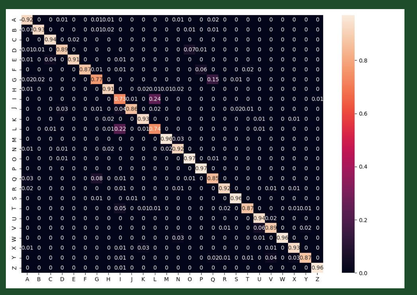
The goal of this project is to implement an accurate handwritten text recognition system, with the eventual goal of implementing the program on a smartphone app.
The project takes a pipeline approach to text recognition, passing through four basic stages: segmentation, OCR, word prediction, and sentence prediction.



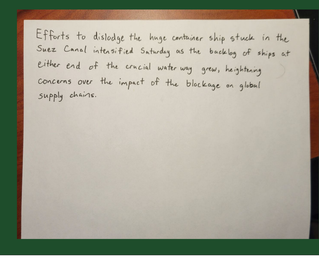
Block diagram for program pipeline

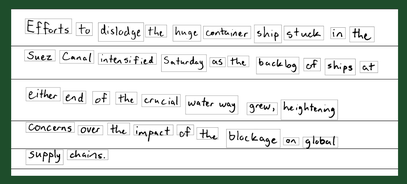
Results

- Currently able to obtain perfect output from a test image with golden segmentation
- Problems encountered earlier in the pipeline get fixed by later stages, demonstrating pipeline's effectiveness
- Current OCR model achieves 98% top-3 character accuracy
- Word predictor can accurately determine the correct word from OCR output using golden segmentation



Confusion matrix displaying top-1 classification accuracy for the OCR model we created





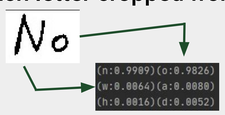
Efforts to dislodge the huge container ship stuck in the Suez canal intensified Saturday as the backlog of ships at either end of the crucial waterway grew heightening concerns over the impact of the blockage on global supply chains

Text output obtained using golden segmentation.

Word Prediction

Word Prediction pulls information from a variety of sources to piece together the most likely words and put them in order.

- OCR - The 3 most likely letters for each letter cropped from text



- Segmentation - The ways in which letter data can be used (The graph from the segmentation box)
- A word corpus - A list of words and how often they were used in the data scraped from Wikipedia

Result for the Mexico Example:

```
['mexico', "mexico's", 'ratiu', 'hoju', 'main', 'adoze', 'mazed', 'mazers', 'moils',
```

Sentence Prediction

Sentence prediction uses a large corpus of text to predict the likelihood that a word comes before or after another word.

- Word context is determined by leading and following words

```
... he decided to ...
```

- Module loops through sentence, determining likelihood of each word
- If likelihood too low, word predictor module is called to generate new possibilities

Next Steps

If more work was to be done on this project in the future, there are a couple things to keep working on.

- App development - The original project idea was to get this working on a phone.
- More segmentation - a lot of progress has been made this year, but we know that more time will lead to much better results.
  - There have been ideas thrown around to add a second neural network to help with this problem.