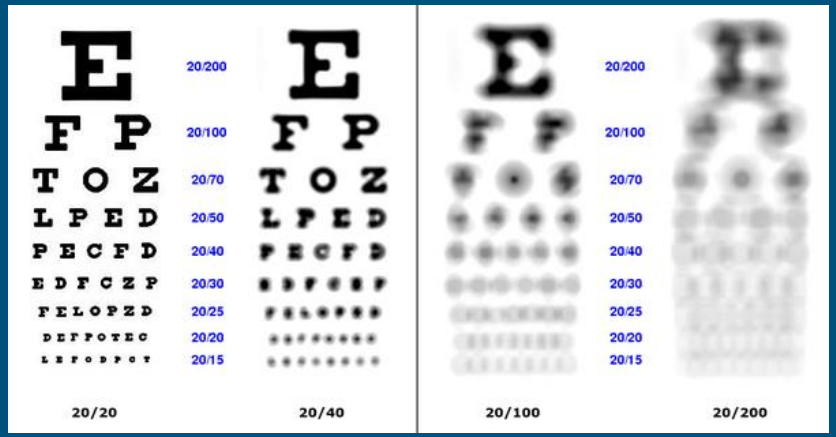
EZRead

Application of Deep Learning Text Recognition

Problem Statement

- The difficulty of reading text for the visually-impaired



Existing Solutions

Low Vision Devices

- Device enlarges text to improve readability
- Still requires the user to read the text

Optical Character Recognition

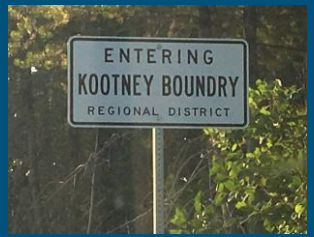
- Converts documents into machine legible text
- Greater difficulty or impossible to recognize text from “noisy” documentation

Project Objective

- Initially to explore the application of deep learning text recognition
- Utilizing Deep Learning, make a device that would be capable of reading text in regular day circumstances

Solution

- Embed our Program on a Raspberry Pi
- Program utilizing two neural networks to detect (localize) and recognize text from a given picture
- Recognized text is converted into machine legible text and passes it through text-to-speech (TTS)



Initial Image

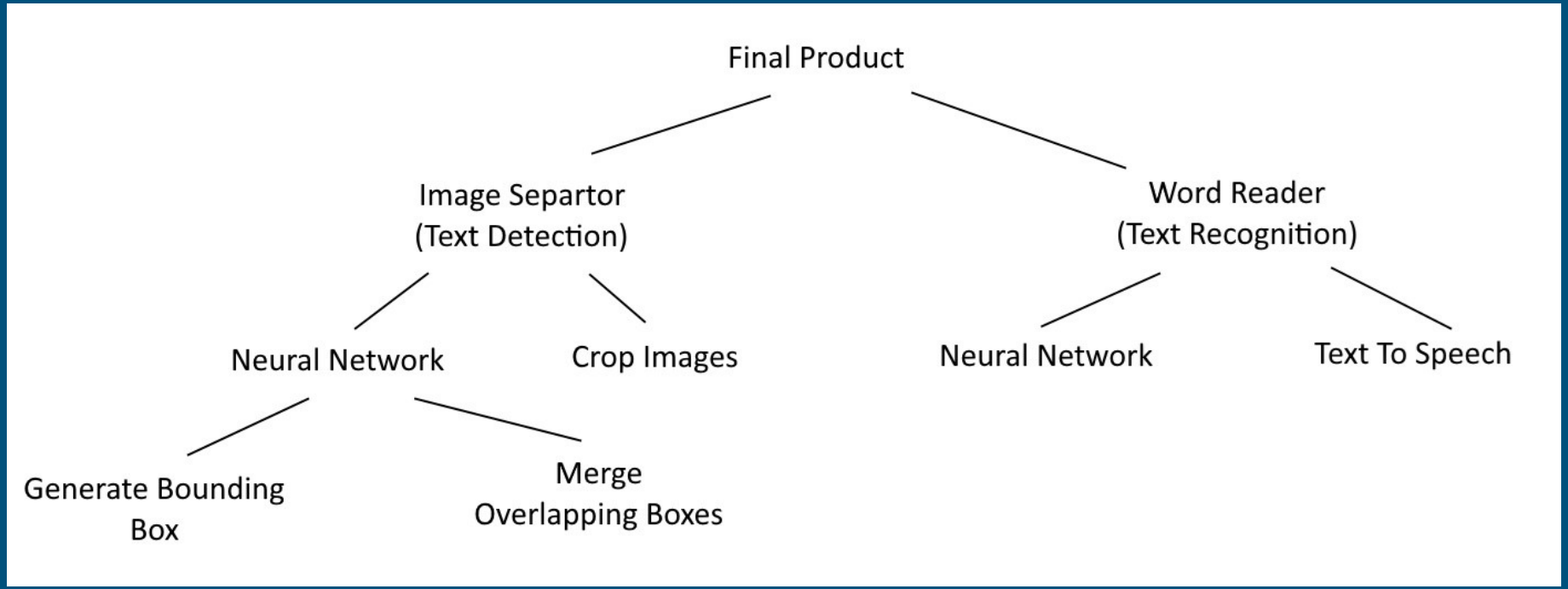
Text Detection

ENTERING
KOOTNEY
BOUNDRY
REGIONAL
DISTRICT

Text Recognition

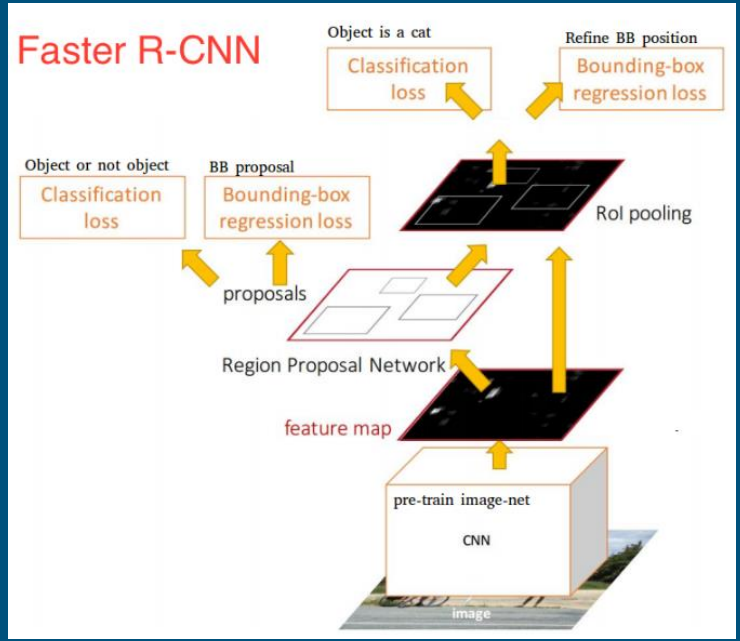
Text-to-Speech

Product Breakdown



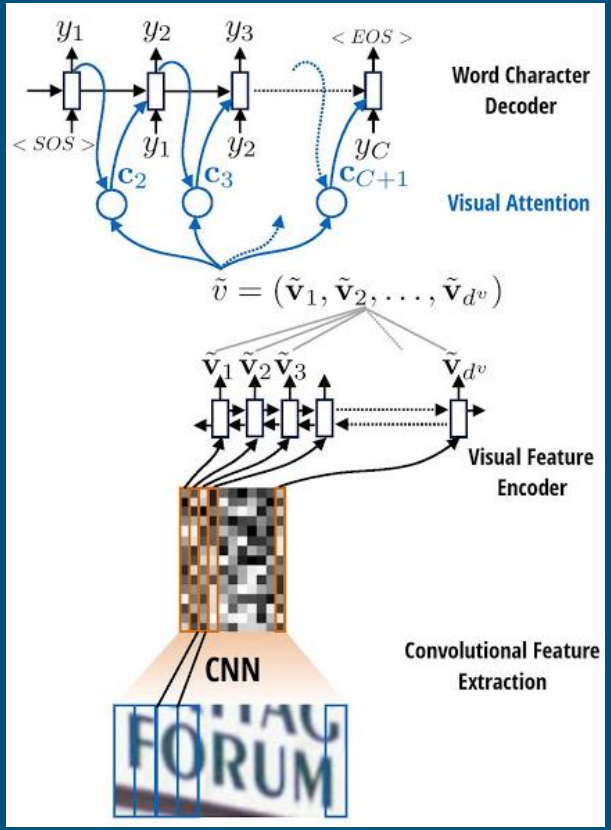
Text Detection Model

- The first step is to run a CNN over the whole image
- It uses a region proposal network on the features that were extracted from the CNN
- After receiving region proposals, it is fed into a pooling layer, some fully connected layers, and a softmax classification layer.
- *Used pre-existing model from Tensorflow [1]*
- *Tensorflow Object Detection API for training [2]*



Text Recognition Model

- First runs a sliding CNN on the image.
- Then an LSTM is stacked on top of the CNN.
- Finally, an attention model is used as a decoder for producing final outputs.
- *Pre-existing model by Qi Guo and Yuntian Deng [3]*
- *Used training application by Ed Medvedev [4]*



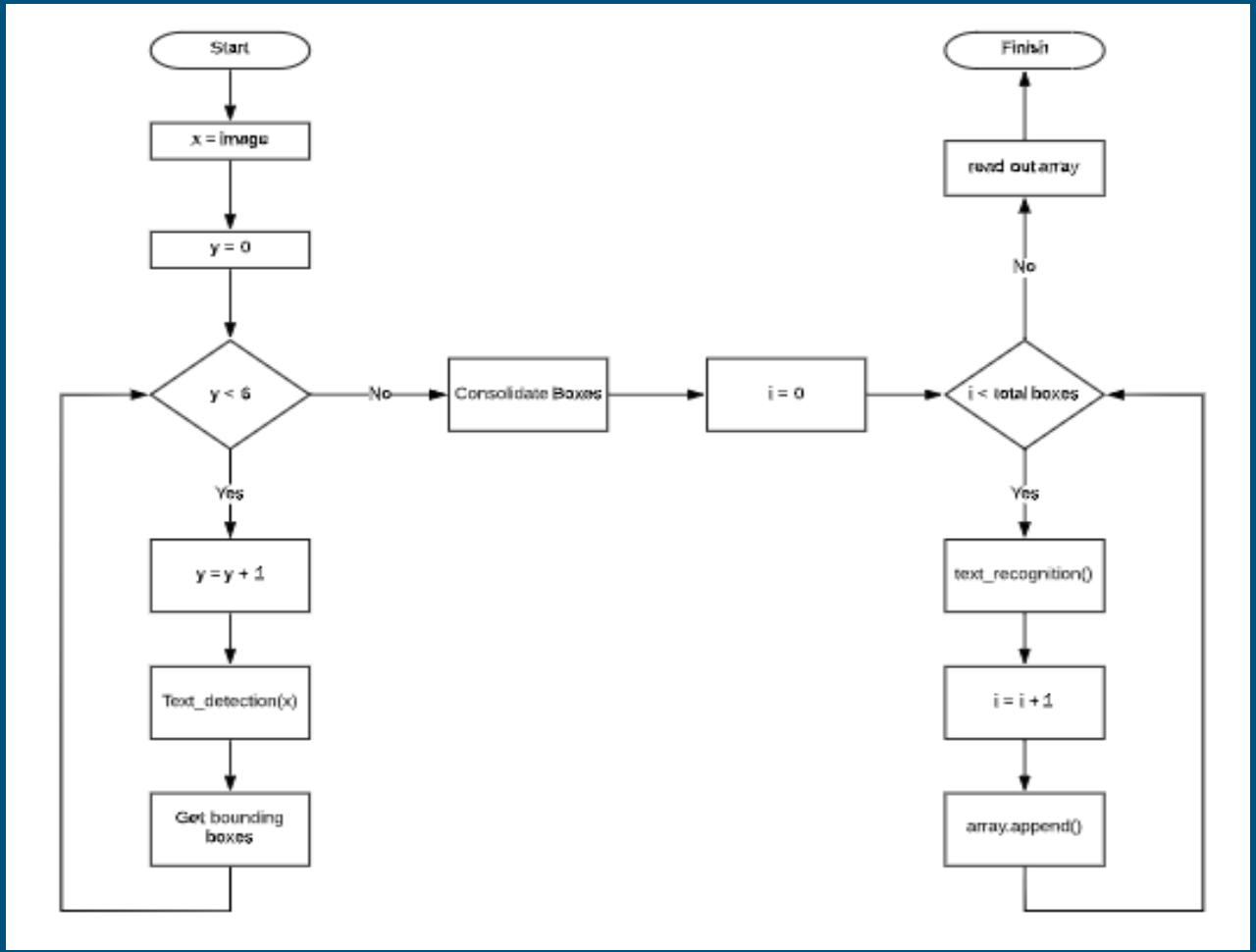


Figure: Flowchart of the Application

Testing

- Initial testing for our product in the fall term

Figure: Test Image with location boxes

```
ANNE COULBY AND INTERNSHIP OFFICER FACULTY OF ENGINEERING AND APPLIED SCIENCE 2000 SIMCOE
STREET NORTH IO UI OSHAWA ONTARIO LIH IKA 01 905 121 8668 5702 INSTITUTE INNECOLITY LOUOIT FACULTY AND
```

Figure: Log of the recognized text from the test image

Testing

Issues found during testing:

- Large body of text accuracy
- Long execution time

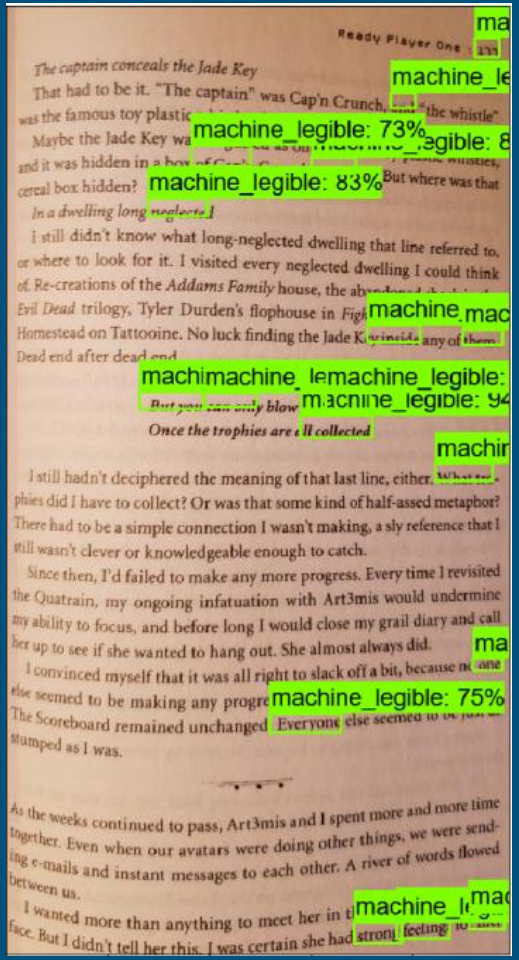


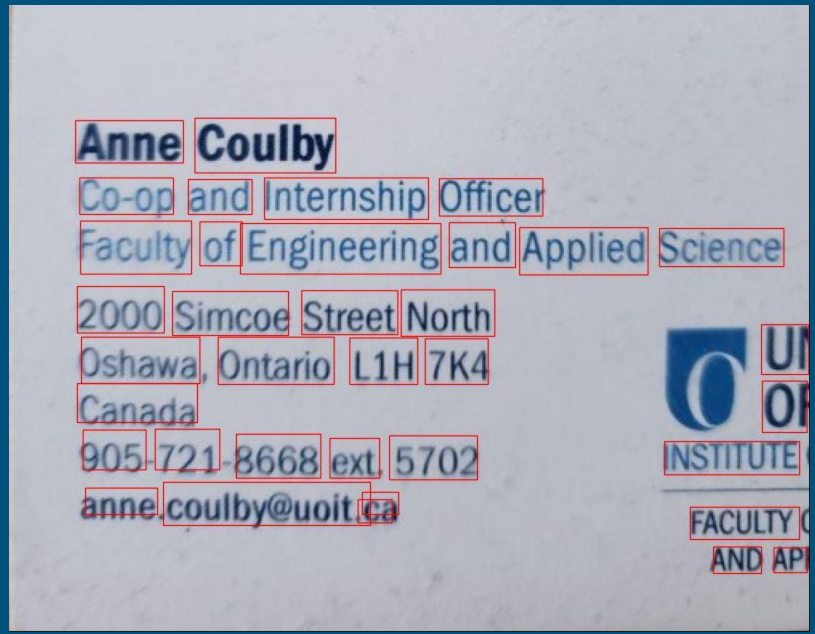
Figure: Text detection of a page in a textbook

Partitioning

Figure: Image with partitioning boxes applied

Testing

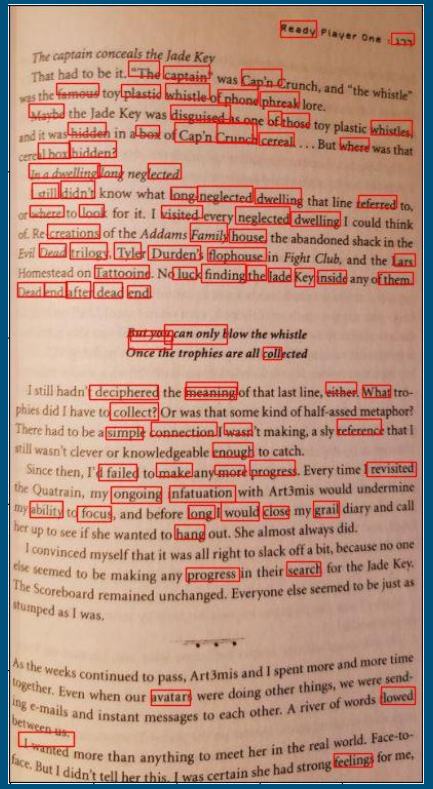
Figure: Image of the updated text detection




```
['ANNE', 'COULBY', 'COOD', 'AND', 'INTERNSHIP', 'OFFICER', 'FACULTY', 'OF', 'ENGINEERING', 'AND', 'APPLIED', 'SCIENCE', '2000', 'SIMCOE',
'STREET', 'NORTH', 'OSHAWA', 'ONTARIO', 'LIH', '7KA', 'UI', 'CANADA', '01', '905', '101', '3668', 'EXT', '5702',
'INSTITUTE', 'ANNE', 'COULTUVISMOUSLY', 'CA', 'NA', 'FACULTY', 'AND', 'AP']
```

Figure: Log output of the updated text recognition

Testing



Ready Player One 133

The captain conceals the Jade Key

That had to be it. "The captain" was Cap'n Crunch, and "the whistle" was the famous toy plastic whistle of phone phreak lore. Maybe the Jade Key was disguised as one of those toy plastic whistles, and it was hidden in a box of Cap'n Crunch cereal . . . But where was that cereal box hidden?

In a dwelling long neglected

I still didn't know what long-neglected dwelling that line referred to, or where to look for it. I visited every neglected dwelling I could think of. Re-creations of the *Addams Family* house, the abandoned shack in the *Evil Dead* trilogy, Tyler Durden's flophouse in *Fight Club*, and the Lars Homestead on Tattooine. No luck finding the Jade Key inside any of them. Dead end after dead end.

But you can only blow the whistle
Once the trophies are all collected

I still hadn't deciphered the meaning of that last line, either. What trophies did I have to collect? Or was that some kind of half-assed metaphor? There had to be a simple connection I wasn't making, a sly reference that I still wasn't clever or knowledgeable enough to catch.

Since then, I'd failed to make any more progress. Every time I revisited the Quatrain, my ongoing infatuation with Art3mis would undermine my ability to focus, and before long I would close my grail diary and call her up to see if she wanted to hang out. She almost always did.

I convinced myself that it was all right to slack off a bit, because no one else seemed to be making any progress in their search for the Jade Key. The Scoreboard remained unchanged. Everyone else seemed to be just as stumped as I was.

• • •

As the weeks continued to pass, Art3mis and I spent more and more time together. Even when our avatars were doing other things, we were sending e-mails and instant messages to each other. A river of words flowed between us.

I wanted more than anything to meet her in the real world. Face-to-face. But I didn't tell her this. I was certain she had strong feelings for me,

Figure: Large body of text with updated text recognition

References

[1] *Tensorflow Detection Model Zoo*, Github. [Online]. Available:https://github.com/tensorflow/models/blob/master/research/object_detection/g3doc/detection_model_zoo.md

[2] *Tensorflow Object Detection API*, Github. [Online]. Available:https://github.com/tensorflow/models/tree/master/research/object_detection

[3] *Attention-OCR*, Github. [Online]. Available:https://github.com/da03/Attention-OCR

[4] *Attention-based OCR*, Github. [Online]. Available:https://github.com/emedvedev/attention-ocr

[5] *Deep Learning for Object Detection: A Comprehensive Review*, Towards Data Science. [Online]. Available:https://towardsdatascience.com/deep-learning-for-object-detection-a-comprehensive-review-73930816d8d9