Graph2Text: Get Textual data from Bar charts

Prateek Mehta
School of Information
Rochester Institute of Technology
Email: **pm7372@rit.edu**

Motivation

This project assists non-visual students by converting bar charts into tabular data for seamless integration with text-to-speech systems, enhancing their comprehension of chart content.

Research Questions

RQ1: System performance comparison between generated and extracted graphs.
RQ2: System performance comparison between vertical and horizontal bar graphs.

Dataset

Subset of dataset featured in the Kaggle competition titled "Benetech - Making Graphs Accessible" (available at https://www.kaggle.com/competitions/benetech-making-graphs-accessible/)

Methodology

Pipeline

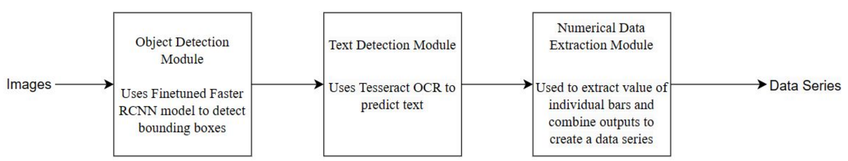

Figure 1: Pipeline Overview

Object Detection Module

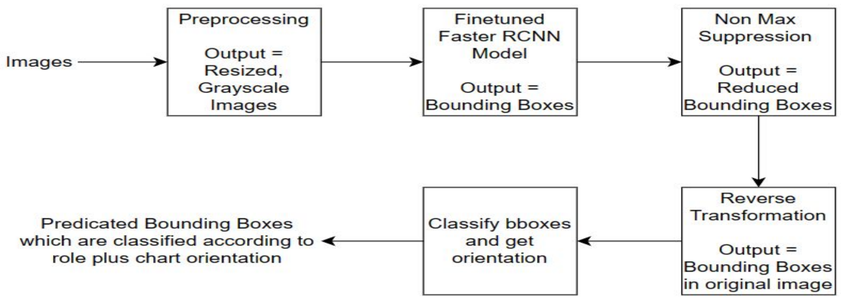

Figure 2: Object Detection module overview

Methodology (cont'd)

Text Detection Module

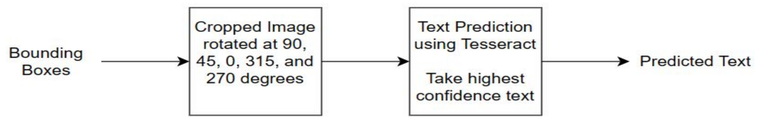

Figure 3: Text Detection module overview

Numerical Data Extraction Module

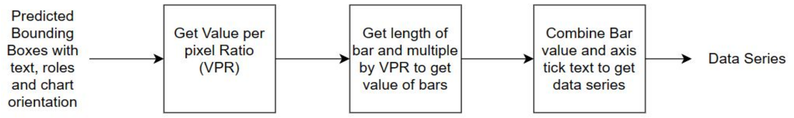

Figure 4: Numerical Data Extraction module overview

Results

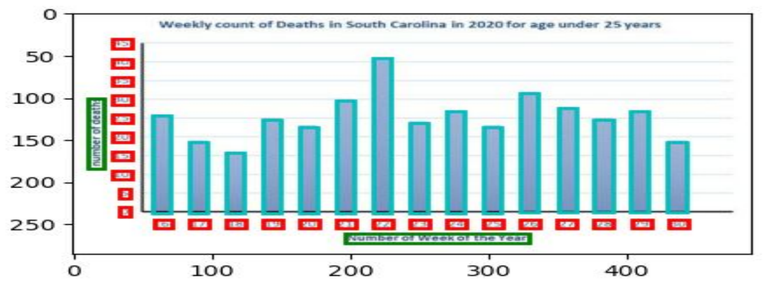

Figure 5: Output of Object Detection Module

```
Y Axis Title =  number of deaths
X Axis Title =  Number of Week of the Year
Data Series
        16.0       17.0       18.0       19.0       20.0       21.0       22.0  \
0  26.275593  19.057676  16.165941  25.146349  23.133963  30.195679  41.60091

        23.0       24.0       25.0       26.0       27.0       28.0  \
0  24.120855  27.386711  23.079058  32.059498  28.119787  24.941734

        29.0       30.0
0  27.164059  18.939537
```

Figure 6: Combined Output of Text Detection and Numerical Data Extraction Modules

Results (cont'd)

Dataset	Loss
Source = Extracted	0.4152
Source = Generated	0.2340
Chart Type = Vertical	0.2368
Chart Type = Horizontal	0.7846
Overall	0.2376

Table 1: Validation Loss for object detection module on each testing dataset subset

Discussions

- The Object Detection Module performs well in identifying axis titles, axis ticks, and bars. However, it may face challenges when confronted with oversized text or the presence of grid lines. Additionally, it consistently struggles to recognize chart titles.
- Tesseract OCR occasionally assigns higher confidence to incorrect text while assigning lower confidence to accurate text, leading to a compromise in the overall output quality. Moreover, it sometime adds strings in float values or vice versa complicating the process of data series generation and evaluation.
- Text and Number detection modules have to perform a lot of action on each individual image therefore are slow processes and cannot be vectorized.
- The above two points makes it difficult to evaluation system output.

Conclusions & Future Work

- **RQ1:** The system performs better on extracted graphs.
- **RQ2:** The system performs better on Vertical bar graphs.
- The above two results are partly due to the fact that the dataset had a huge majority of generated vertical bar graphs as opposed to other subsets. To increase performance we can try retraining after balancing the dataset.
- Future work includes tweaking the object detection module to start detecting chart titles, exploring alternate ways for text detection to avoid issues raised by Tesseract OCR, exploring post-processing techniques to sanitize output, and finally increasing performance on horizontal bar graphs and extracted graphs.

References

1. Dai, Wenjing, et al. "Chart decoder: Generating textual and numeric information from chart images automatically." Journal of Visual Languages & Computing 48 (2018): 101-109.
2. Zhou, Fangfang, et al. "Reverse-engineering bar charts using neural networks." Journal of Visualization 24 (2021): 419-435.
3. Choi, Jinho, et al. "Visualizing for the non-visual: Enabling the visually impaired to use visualization." Computer Graphics Forum. Vol. 38. No. 3. 2019.