

Project 5

Box/Point Detector using CenterNet



Team Members

Maruf Shaikh

Parth Rajput

Tyler Perison



Overview

- Visual representation of data in form of graphs/charts conveys more information.
- Lot of research papers/articles/journals portray graphical information in form of images.
- Extracting this information from just these images is therefore critical.
- To encourage research in this area, ICDAR organizes competitions.



Competition on Harvesting Raw Tables from Infographics

Tasks:

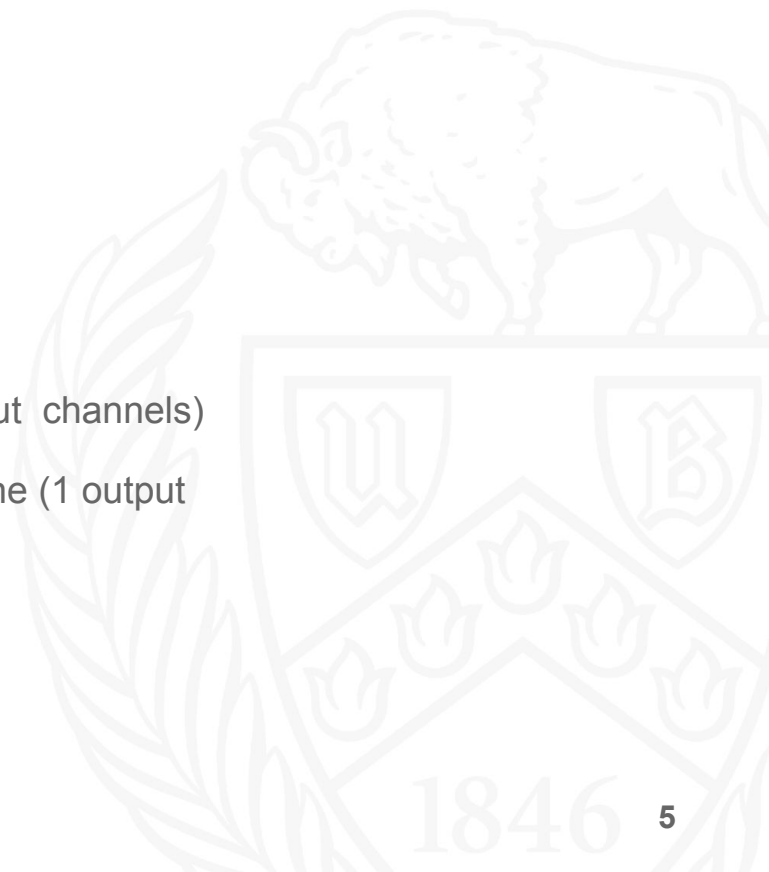
1. Chart Classification
2. Text Recognition/Classification
3. Text Role Classification
4. Axis Analysis
5. Legend Analysis
6. A. Plot element Detection/Classification
B. Raw Text Extraction
7. End-to-end Classification

Datasets:

1. UB PMC
2. Adobe Synth

Scope

- Explore Solution Proposed by Deep Blue AI.
- Task 6 to be reproduced.
 - Bar Charts- Box Detections using CenterNet
 - Box Plots - CenterNet using DLA-34 background (5 output channels)
 - Line and Scatter Plots- CenterNet using DLA-34 backbone (1 output channel)
 - Data Extraction.
- Explore possible improvements.



Progress

- Reviewing papers provided. (Completed)
- Exploring Additional resources (Completed):
 - Relevant Papers.
 - Github repositories for reference models.
- Implementation-
 - Setup Github repository (Done).
 - Setting up Data Loaders (Done).
 - Replication of the approach in PyTorch (WIP).

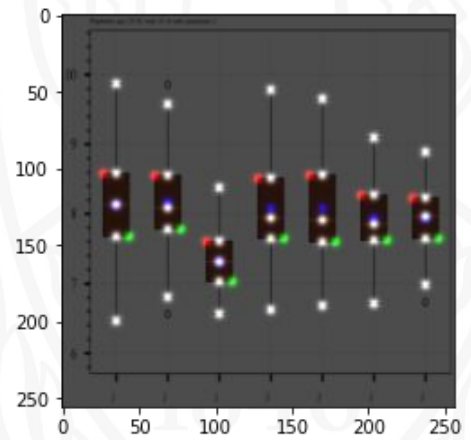
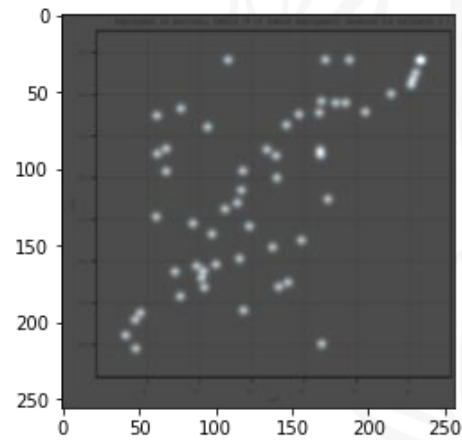
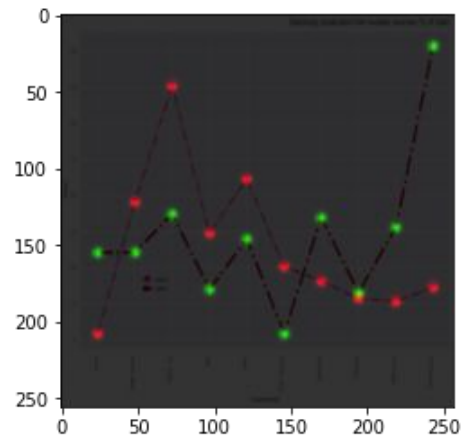
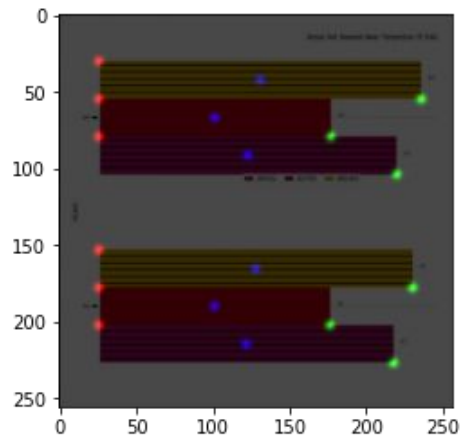
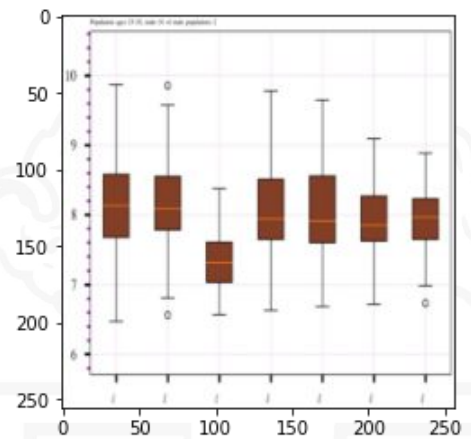
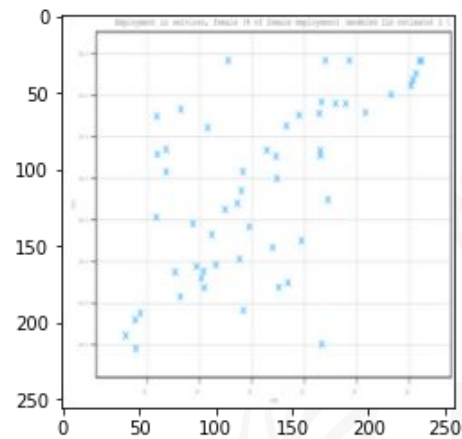
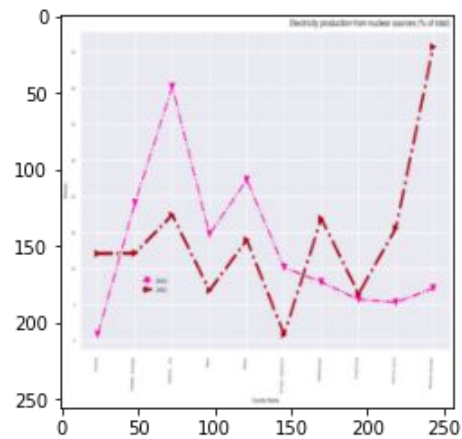
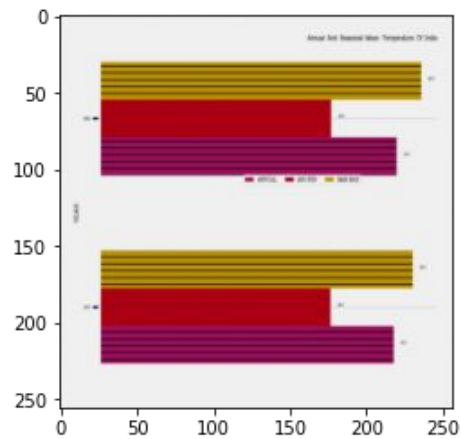


DataSets

- UB PMC - Limited but generated from actual charts from PubMedCentral
- Adobe Synth - Synthesized Dataset
- Both similarly annotated per task
- Dataset Annotations required cleaning (Task wise data missing)
- Annotations contained details such as coordinates of chart type, bounding box position, etc.

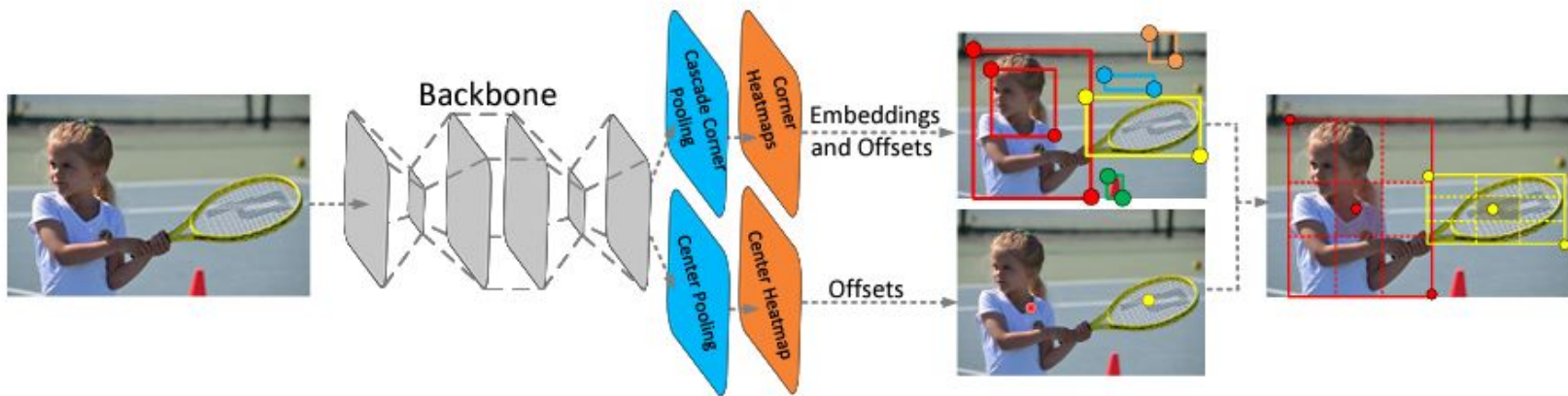
DataLoaders

- Needs to handle Chart type based inputs.
- Initial approach was to use a DataLoader for each model.
- Now there is a single PyTorch Dataset which customizes outputs based on chart type and dataset.
- Standard outputs of the Loader:
 - Input Image (resized to 1024x1024)
 - Corresponding HeatMap (blurred wrt 256x256 size then resized to 256x256)
 - 3 channels for boxes (bar and box), and 0, N, N, or 5 channels for points (bar, line, scatter, and box)
 - Bounding box details (important x,y coordinates for each box)
 - Points (lists of x,y coordinates remapped to new HeatMap size)



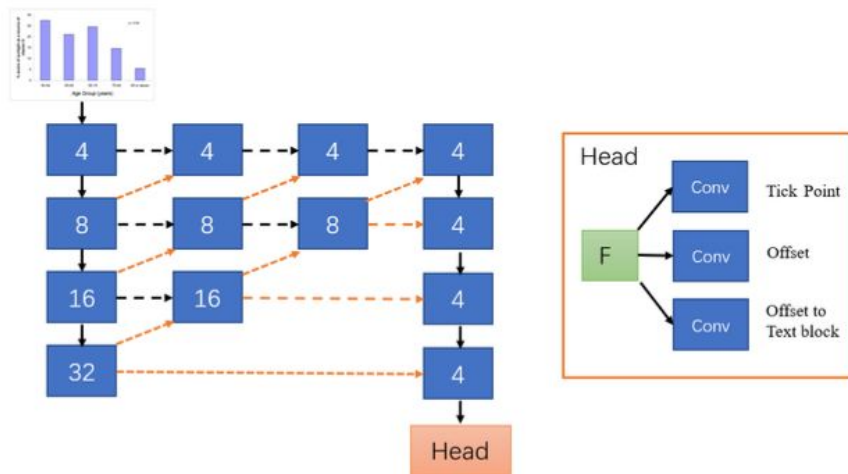
Model Implementation

For CenterNet, we use a simplified implementation that was based off the original model.



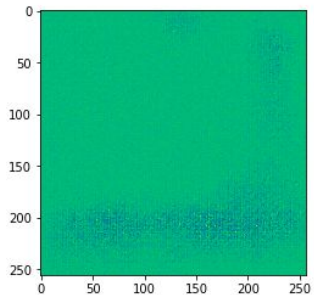
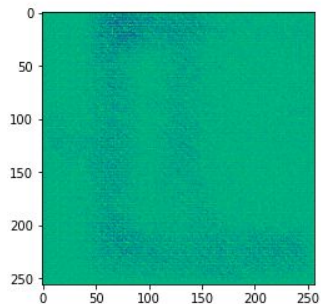
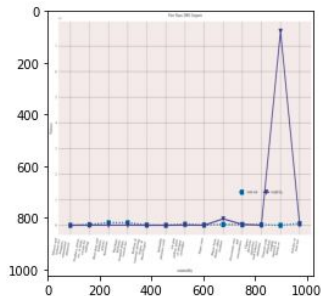
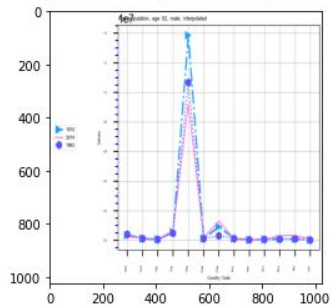
Model Implementation

For DLA backbone, we used the official DLA implementation and used that as a backbone to CenterNet modules. That makes 2 of our models(just changing the output channels).

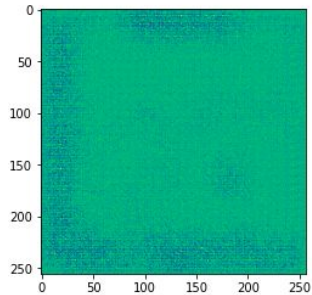
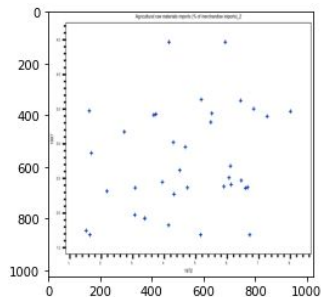


Output Heatmap Visualizations (Initial Forward Pass)

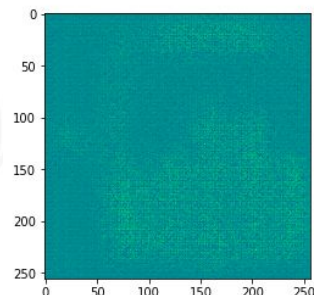
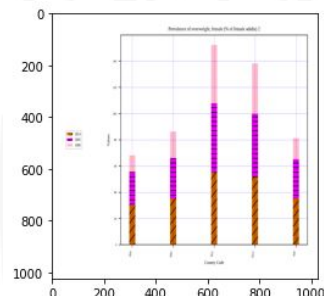
Line Plots



Scatter Plots



Bar Charts



Challenges/Issues

- Dataloader takes time to load data (Loading images/annotations from google drive/ generating heat maps)
- Difficulty in understanding/implementation provided for original model.
- Alternative implementations explored.
- Need to figure out the inputs for the loss function/corresponding DataLoader modifications,

Plan for next two weeks

7th December(Today)- Presentation

By 8th December- Start with Training the models

8th December to 12th December- Training + Element grouping with legend (HoG)

12th December to 14th December- Reproducing the results achieved by the benchmark paper +
Looking into ICP and GM code

15th December to 17th December- Implement improvement techniques

17th December to 19th December- Hyper parameter tuning + Modularize the code

References

[CenterNet Model Original Repository](#)

[CenterNet Model Reference Repository](#)

[ICPR 2020 - Competition on Harvesting Raw Tables from Infographics](#)

[Towards an Efficient Framework for Data Extraction from Chart Images](#)

[Improving Machine Understanding of Human Intent in Charts](#)

[CenterNet: Keypoint Triplets for Object Detection](#)

[A Benchmark For Analyzing Chart images](#)

