Identification of Formal Features of Comic Book page

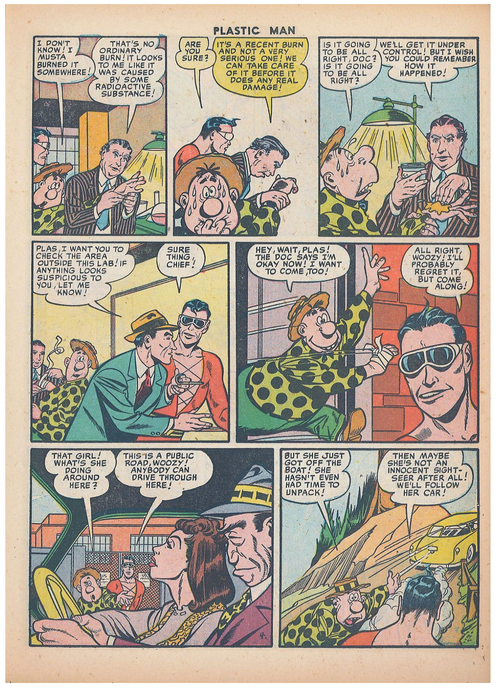
Siddhanth U Hegde, Nilesh Rathi, Abhishek V Dhar, School of Informatics and Computing, Indiana University

1. Motivation

- Scholarly attention towards comic books and graphic novels is increasing.
- Identifying features such as number of panels, avg size of panels, and number of speech balloons per page can aid in the classification of comic books and graphic novels.
- Analyzing these features can provide insight into the creative processes of comic book artists as well as the evolution of the medium over time.

2. Data Description

- We utilized a dataset provided by the University of Maryland which consists of 500 images and annotations of panels and text boxes. The annotations are provided as text files consisting of edge coordinates for each of the bounding boxes.



Panel annotations

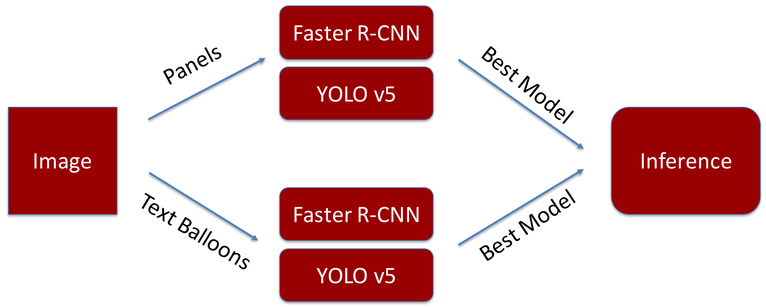
```
1 58 78 427 627
1 440 84 813 627
1 822 86 1257 633
1 70 636 593 1175
1 608 644 1251 1183
1 60 1194 703 1735
1 718 1190 1253 1741
```

(xmin,ymin) (xmax,ymax)

```
0 0.18 0.19 0.28 0.30
0 0.48 0.19 0.29 0.29
0 0.81 0.19 0.33 0.29
0 0.25 0.49 0.40 0.29
0 0.72 0.50 0.50 0.29
0 0.29 0.80 0.50 0.29
0 0.76 0.80 0.41 0.30
```

Normalized points with class

3. Pipeline

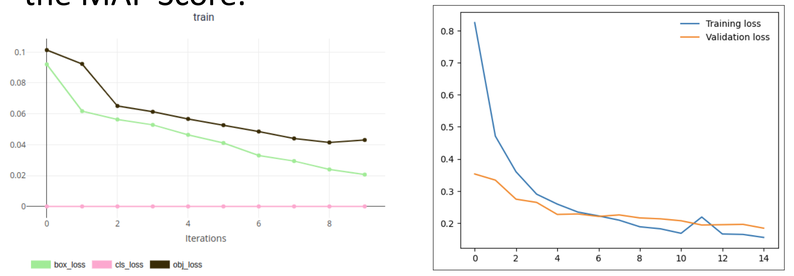


- The proposed pipeline would take a single page as an input and use the trained models to predict the panels and text balloons.
- These predictions will be passed on to the inference step that is a script that summarizes the results and outputs informative features.
- Examples of informative features are:

1. Number of panels/text boxes detected per page
2. Average panel size
3. Maximum and minimum panel size relative to the page size.

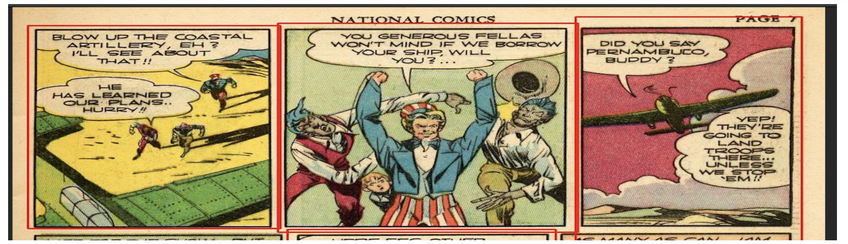
4. Model Training

- We performed image resizing and data augmentations on the training dataset.
- We trained YOLO v5 and Faster R-CNN models for 15 epochs across the dataset and evaluated them using the MAP Score.



Training Loss for Panel Detection Training for YOLO v5 and Faster R-CNN

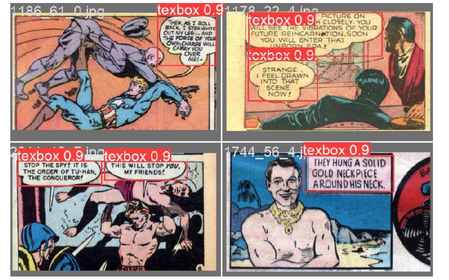
5. Results



Panel detection using Faster R-CNN

Text Box detection using YOLO v5

Text Box detection - Faster RCNN



Text Box detection – Yolo v5

Model	Panel IoU	Text Balloon IoU
Faster RCNN	92	88
Yolo V5	97	95

6. Future Work

- Extend the model to process multiple pages and detect page layout changes.
- Further segment into text and image regions for text recognition and character detection.
- Apply transfer learning techniques to adapt the model to other comic book datasets.