Honours project report - CVIT

Understanding architectures for automatic and semi-automatic layout parsing of historical documents.

Vivek R

December 25, 2021

Abstract

Analysing the outputs of PALMIRA: A Deep Deformable Network for Instance Segmentation of Dense and Uneven Layouts in Handwritten Manuscripts (automatic layout parsing) and BoundaryNet - An Attentive Deep Network with Fast Marching Distance Maps for Semi-automatic Layout Annotation (semi automatic layout parsing).

1 Introduction

Palmira first introduces Indiscapes2, a new large-scale diverse dataset of Indic manuscripts with semantic layout annotations, to address the issue of poor results in complex deformations seen across semantic regions. It also propose a novel deep network Palmira for robust, deformation-aware instance segmentation of regions in handwritten manuscripts.

The fully automatic boundary estimation approaches tend to be data intensive, cannot handle variable-sized images and produce sub-optimal results for irregular and overlapping multi-class region instances with large range in aspect ratio. BoundaryNet is a novel resizing-free approach for high-precision semi-automatic layout annotation. It takes variable size user input as region of interest and provides a better region boundary.

2 Work Done

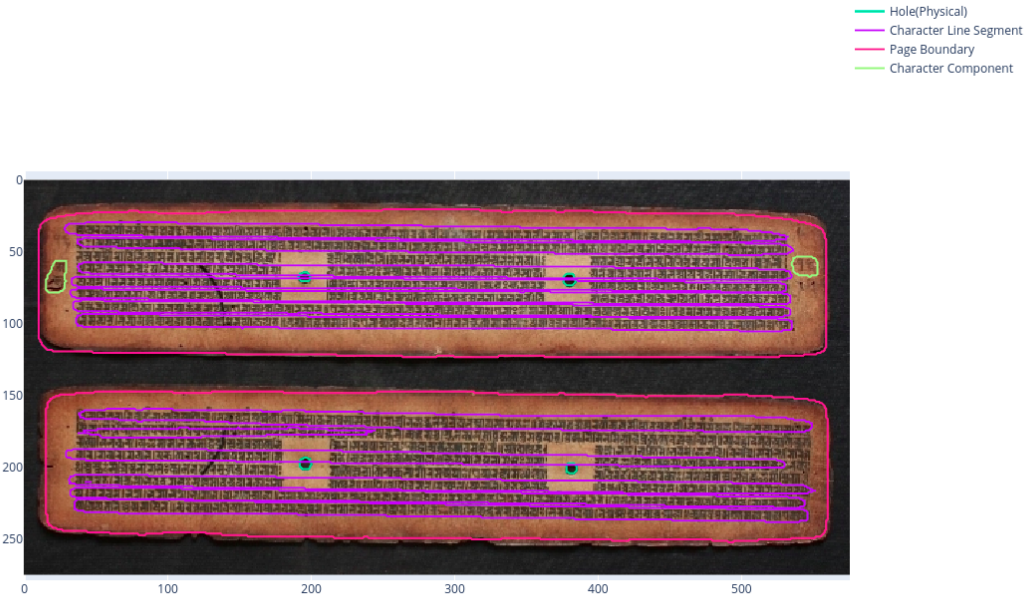
I learned about various methods of region segmentation, getting an overview of how the methods work.(Notes)

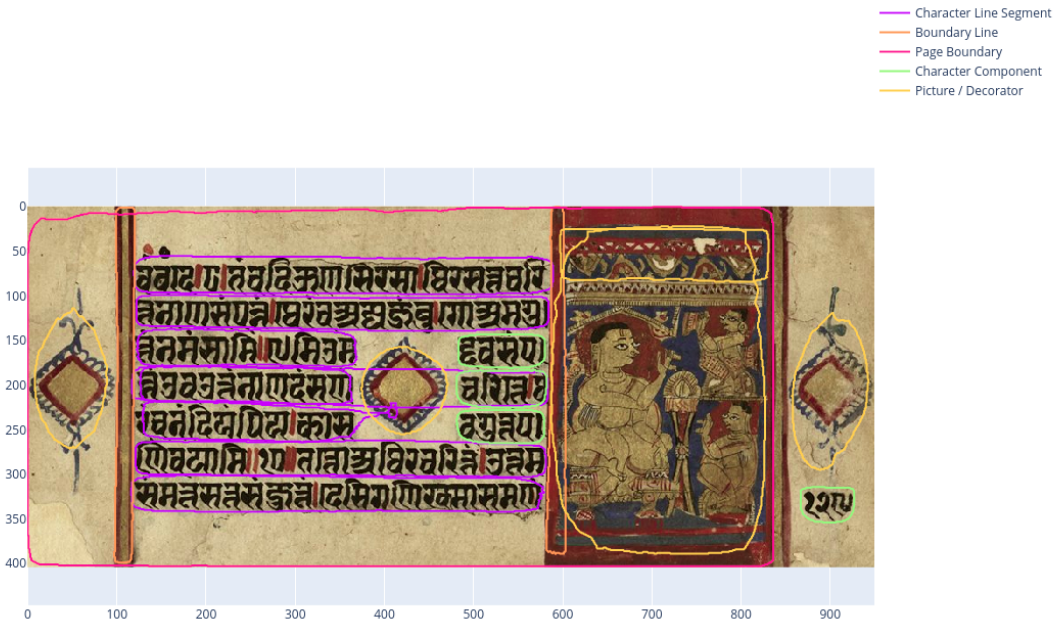
I made google Colab notebooks to visualize the models:

- Palmira(link)
- BoundaryNet(link)

3 Results

Palmira results





4 Conclusion

Outputs of both automatic and semi automatic region segmentation are obtained.

5 Future plans

I am planning to work on word and layout segmentation for OCR. Currently I am testing the existing models of word and line detections familiarizing with their outputs and issues.