A Novel Pipeline for Table Extraction using Deep Learning

Student: Lee Seng Cheong          Supervisor: Dr Loke Yuan Ren

Tables and tabular formats are structured manners of presenting information. Much of the information within tables cannot be recovered without preserving the structure of the table, and manual extraction of tables is tedious. However, automated systems for table extracting must deal with tables from different domains, font, styles, structures and mediums with such systems mainly oriented towards digital documents only. We propose an image-based deep learning pipeline for extracting tables that can accommodate the variety in formats and styles from both images and digital documents.

1. To perform table detection, a modified Single Shot Detector (SSD) model is utilized with deformable convolutions for enhanced table detection performance
2. To perform table cell segmentation, Mask RCNN is used to identify and locate table cell instances
3. Using Tesseract OCR, the text content in each table cell is recovered
4. Using the position coordinates of each table cell instance, each table cell’s row and column position is determined to reconstruct the table

The pipeline thus presents an automated end-to-end solution for table extraction, allowing automated extraction of tables from both digital documents and images. In addition, a modified version of the SSD model is proposed, incorporating deformable convolutions that reports improved table detection performance over the original version. It is portable as it can accept tables from different domains and mediums via conversion of documents to an image-based format. Thus, tables from sources such as scanned documents and photos can be automatically converted to a digital format. In addition, being a computer vision approach, it is robust to the various table styles and fonts present in documents of different domains.

Table cell segmentation using instance segmentation approaches

<table>
<thead>
<tr>
<th>F1-measure @ IoU Threshold</th>
<th>Table Detection Performance on ICDAR2013 table dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.6</td>
<td>79.9%</td>
</tr>
<tr>
<td>0.7</td>
<td>78.5%</td>
</tr>
<tr>
<td>0.8</td>
<td>77.2%</td>
</tr>
</tbody>
</table>

Table Cell Segmentation

Pascal-VOC mAP@ IoU=0.5 on UNLV table dataset

Mask RCNN + Resnet-101: 63.5%