Table Recognition and Accessibility in RAVI with Layout Parser and MTL-TabNet

K Laxman 2018CS50408

COD893 Major Presentation Advisors: Prof.M.Balakrishnan, Prof.Volker Sorge, Abhishek Baghel Supervisor: Prof. Vireshwar Kumar

July 10, 2023



About RAVI

- RAVI (Reading Assistant for Visually Impaired) aims to make digital content in PDFs accessible to readers with blindness and low vision.
- **The overall objective** of this project is to develop a web application that automatically converts PDF files to editable HTML.
- The output HTML can further be edited on the platform.

Overview of MTP1 work

- The generated HTML was not containing image descriptions. The challenges of lacking image descriptions was addressed.
- The aim Enhance text extraction and improve NLP-based descriptions in the output HTML.
- Crowdsourcing feature, allowing users to add Alt text to images in the HTML.
- Improve code by integrating detection modules, using Azure Computer Vision API, and optimizing for generalization.

MTP2: Table recognition and Integration in RAVI

- The main objective of this project is to make tables accessible in a document like PDF.
- Analysis of table recognition, content extraction, and placing them in output HTML.
- In previous work, integration of the Bordered Table Algorithm was completed

Previous work on Table Ananlysis

- This work was done by Amar Agnihotri (MTP -2022)
- Two types of tables considered : Bordered tables and Borderless tables
- Amar worked on the bordered tables
- In his work he proposed to use CascadeTabNet for table recognition for borderless tables.
- However, Content extraction and integration of borderless tables into RAVI were not done.

Examples:

Product ID	Product Name	Product Quality	Product Quantity
1	Wheat	Good	200 Bags
2	Rice	Good	250 Bags
3	Sugar	Good	200 Bags

Figure: Example of Bordered table

Topics	Days for Class	Timing
Design For Ux	Monday	8:30 - 10:30 am
Design Thinking		12:30 - 2:30 pm
Empathy Map	Tuesday	9:30 - 11:30 pm
Emotional Intelligence	Wednesday	
Jsability	Thursday	8:30 - 10:30 am
Utility	Friday	8:30 - 10:30 am
Accessibility	Friday	8:30 - 10:30 am

Figure: Example of Borderless table

Bordered table recognition and content extraction by Amar Agnihotri.

- Amar Agnihotri's work focused on detecting tables and extracting their data as tabular cell structures.
- Various techniques and algorithms were discussed, including the Bordered Table Algorithm and Borderless Table Algorithm.

Bordered Table Algorithm

- Detect vertical lines in the table.
- Detect horizontal lines in the table.
- Identify the start and end points of each vertical and horizontal line.
- Map the cell data based on the intersection points of the vertical and horizontal lines.

3.2 Bordered Table Algorithm Explanation with an Example

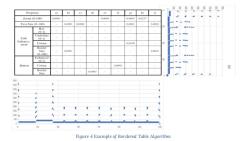
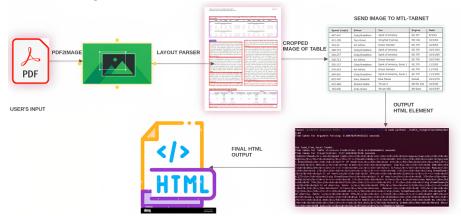


Figure: Example of Borderded table working

Problem Statement

- Oborderless tables are not identified: There is a high chance that tables without borders will be present in PDFs, causing screen readers to miss the table content. This results in visually impaired individuals having partial knowledge and difficulty understanding STEM material, as they are unable to comprehend the borderless tables that describe the data statistics.
- To identify the location of tables in document images and to extract the table data
- Debugging of bordered table algorithm Different individuals contributed code for table analysis and content extraction, but the output HTML displayed extra paragraphs below the table. This bug was been reported and resolved in this work.

Workflow image:



Solution

- The Layout parser is utilized to process the input PDF file with multiple pages and accurately identify the regions containing tables throughout the document.
- OpenCV is employed to extract the specific table regions from the identified areas and save them as separate images. The corresponding coordinates (x1, x2, y1, y2) of each table region are recorded and stored in a JSON file for reference.
- The Layout parser's output images are then passed to the MTL-TabNet algorithm. MTL-TabNet processes each image and generates the corresponding HTML element for the table. In case of multiple images, they are processed in sequential order.
- The JSON file is updated with the relevant image and its corresponding HTML table element. This updated JSON file serves as a reference for integrating the table information into the final HTML output.
- The extracted table data, along with the corresponding HTML table elements, are seamlessly integrated into the main original HTML output, ensuring accurate representation and accessibility for visually impaired users.

Layout Parser Introduction

- Layout Parser is a Python library that facilitates document layout understanding and information extraction
- It starts with document preprocessing, enhancing readability and extracting key features. Next, it analyzes the layout structure, identifying elements like text blocks, images, headers, footers, and tables
- Model: Layout Parser uses Detectron2, a pre-trained transformer-based model that incorporates both text and layout features
- Training Data: LAYOUTLM is initially trained on PubLayNet, a large-scale synthetic dataset with over one million document page

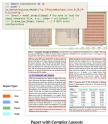


Table Recognition and Accessibility in RAVI with Layor

MTL-TabNet Introduction

- MTL-TabNet is a multi-task learning-based model for image based table recognition
- The model is designed to recognize tables in document images and extract tabular data from them
- The model is trained on various datasets such as PubTabNet, FinTabNet and Marmot.

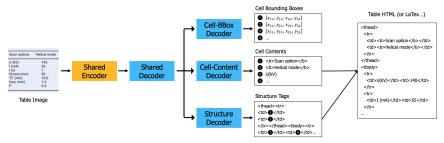


Figure: Working of MTL-TabNet

MTL-TabNet

Why MTL-Tabnet?:

- In this presentation, we will compare two popular methods for table recognition and extraction: MTL-Tabnet and TableMaster
- These methods, developed in the field of document analysis and optical character recognition (OCR), have demonstrated their effectiveness in automatically detecting and understanding tables within documents.
- We will explore their key features, advantages, and provide a comparative analysis.

Table: Comparison between MTL-TabNet, Tablemaster, and Cascade TabNet

Approach	MTL-TabNet	Tablemaster	Cascade TabNet	
Year	2023 (Latest Work)	2021 (Released)	2022 (Released)	
Models	Transformer based model	PSENet + MASTER	Cascade RCNN + TabNet	
Accuracy	0.9885	0.9676	0.9823	
TEDS Score	96.67	0.9676	97.21	
Other Differences				
Training Approach	Fraining Approach End-to-End Multi-Task Learning with Tasks		Cascade Training	
Architecture	Transformer-based Model	CNN-based Model	Ensemble Model	
Dependency	PyTorch Framework	TensorFlow Framework	PyTorch Framework	

TEDS(Text Extraction from Document Structures) Score: Evaluation metric for text extraction algorithms, measures accuracy compared to reference text. Accuracy: Overall correctness of a system or algorithm in text extraction.

COD892 (IIT Delhi)

Table Recognition on PubTabNet



MTL-Tabnet has been chosen over other approaches because of Latest, best approach in 2023 and big database(PubTabnet ,FinTabnet)

Implementaion of MTL-Tabnet

Input image:

Disability		Ballots	Ballots	Results		
Category	Participants	Ballots Completed	Incomplete/Terminated	Accuracy	Time to complete	
Blind	5	1	4	134.5% n=1	1199 sec, n=1	
Low Vision	5	2		(97.7%,	1716 sec, n=3 (1934 sec, n=2)	
Dexterity	5	4	1	198.3% n=4	1672.1 sec, n=4	
Mobility	3	3	0	95.4%, n=3	1416 sec, n=3	

(base) ravi@ravi-Inspiron-5570:~/Desktop/MTP2/Approach1/MTL-TabNet\$ <table id= Disabi ck;">Participants<td rowspan="2" colspan="1" style="border: 1px solid bla solid black:">Ballots Incomplete/Terminated<td rowspan="1" colspan="2" st n="1" style="border: 1px solid black;">Accuracy<td rowspan="1" colspan="1 an="1" colspan="1" style="border: 1px solid black;">Blind<td rowspan="1" ="1" style="border: 1px solid black;">1<td rowspan="1" colspan="1" style= : 1px solid black:">34.5%, n=1<td rowspan="1" colspan="1" style="border: tyle="border: 1px solid black;">Low Vision<td rowspan="1" colspan="1" sty der: 1px solid black;">2<td rowspan="1" colspan="1" style="border: 1px so ck:">98.3% n=2 (97.7%, n=3)<td rowspan="1" colspan="1" style="border: 1px colspan="1" style="border: 1px solid black;">Dexterity<td rowspan="1" col " style="border: 1px solid black;">4<td rowspan="1" colspan="1" style="bo px solid black: ">98.3%, n=4<td rowspan="1" colspan="1" style="border: 1px yle="border: 1px solid black;">Mobility<td rowspan="1" colspan="1" style= : 1px solid black:">3<td rowspan="1" colspan="1" style="border: 1px solid ">95.4%, n=3

Examples of tables:

Showcasing the effectiveness of MTL-Tabnet in handling different types of tables:

Speed (mph)	Driver	Car	Engine	Date
407.447	Craig Breedlove	Spirit of America	GE J47	8/5/63
413.199	Tom Green	Wingfoot Express	WE J46	10/2/64
434.22	Art Arfons	Green Monster	GE J79	10/5/64
468.719	Craig Breedlove	Spirit of America	GE J79	10/13/64
526.277	Craig Breedlove	Spirit of America	GE J79	10/15/65
536.712	Art Arfons	Green Monster	GE J79	10/27/65
555.127	Craig Breedlove	Spirit of America, Sonic 1	GE J79	11/2/65
576.553	Art Arfons	Green Monster	GE J79	11/7/65
600.601	Craig Breedlove	Spirit of America, Sonic 1	GE J79	11/15/65
622.407	Gary Gabelich	Blue Flame	Rocket	10/23/70
633.468	Richard Noble	Thrust 2	RR RG 146	10/4/83
763.035	Andy Green	Thrust SSC	RR Spey	10/15/97

Crystal	Туре	Melting temp.	Notes
T	0	17 °C (63 °F)	
	N E	Soft, crumbly, n	nelts too easily
II	T W O	21 °C (70 °F)	Soft, crumbly, melts too easily
Ш	T	26 °C (79 °F)	Firm, poor snap, melts too easily
	H R E	28 °C (82 °F)	
IV	E		Firm, good snap, melts too easily
	O U R		, mm, good onup, mone too cassily
V	F	34 °C (93 °F)	Glossy, firm, best snap, melts near body temperature (37 °C)
	V E		Hard, takes weeks to form
VI	s	36 °C (97 °F)	
	1		

output image after:

(base) ravi@ravi-Inspiron-5570:-/Desktop/NTL_tabnet/NTL-TabNet\$ sudo python3 ./table_recognition/demo/dem o.py Time taken for Argument Parsing: 0.000762939453125 seconds

Use load from local loader

Time taken for Table Structure Prediction: 2136.6154890060425 seconds

Time taken for Visualization: 2137.046593427658 seconds

<thead>Speed (mph)DriverCat^aDriverCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat<sup
aCat^aCat^aCat^aCat^aCat^aCat^aCat<sup>aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat^aCat< EngineDate<ftp><ftp><ftp><fy>447C<sub>Telig</sub > Bree(dloveSpirit of AnnericaGE J478/5/63 tr><b/td>13.19943.1995 6.277Craig BreedloveSpirit of AnericaE J79< b>10/15/65536.712ArfonsGreen Monster</td ><bp>SE J79<bp>Craiq Breedlove</bp> >576.5534b>Art Arfons4b>Greer. Monster4b>GE J794b>4b>B of Anerica. Sonic 1<bGE J79>b>1.1/15/65<52.407

<!DOCTYPE html>
<html>
<style>
table, th, td {
 border:lpx solid black;
}
</style>

<thead><byCrvstal<byCrvstal</by><byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</byCrvstal</br> TypeMelting telmp.< Notes</thead>I 0underline>17</underline> Soft, crulmbly, melts too equsily °F)28 °C (82 °F)Firm, (00H Sizp, melts t00 @SI)<b/IV<b/IV FRFirm, Q0d SinAp, melts
b>GloSsv. firm. best snap. melts near body (emper)dblure (37 °C)
b>
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c
c</td

</body>

<body>

Crystal	Type	Melting telmp.	Notes
I	O	17 °C (63 °F)	
II	TTQWD	21 °C (70 °F)	Soft, crulmbly, melts too equsily
Ш	T HR	26 °C (79 °F)28 °C (82 °F)	Firm, (OOH Sizp, melts tOO @SI)
IV	FR		Firm, QOd SinAp, melts too ezsily
V	FR	34 °C (93 °F)	GloSsy, firm, best snap, melts near body (emper)dblure (37 °C
VI	SDI	36 °C (97 °F)	



<pre><body> <thead><1 <5>4 <5>4 <5 </thead></body></pre>	gical com by polication by	oi complications (%) oil complications (%) to physicisms (%) to physicisms (%)	227 227) A 227 2	307 209 37A 8 (1-407) 65	040 /t0 000 /b> 000 b>4 000 b>4	(n = 229) :/thead> <t !>40.730.730.7</t 	body> > td>
moratility (%) 11 (2-' 15.1		Q.	1.png			ity (%) </th <th>b></th>	b>
	Speed (mph)	Driver	Car		Engine	Date	1.
<thead:< td=""><td>407.447</td><td>Crain Breedlove</td><td>Spirit of America</td><td></td><td>GE 347</td><td>8/5/63</td><td>/sup></td></thead:<>	407.447	Crain Breedlove	Spirit of America		GE 347	8/5/63	/sup>
407.447<td>413.199</td><td>Tom Green</td><td>Wingfoot Express</td><td></td><td>WE 146</td><td>10/2/64</td><td>1</td>	413.199	Tom Green	Wingfoot Express		WE 146	10/2/64	1
Annerica					GF 179	10/5/64 >	
ch>10/2/64e/h:	434.22	Art Arfons	Green Monster	Green Monster			
donster	468.719	Craig Breedlove	Spirit of America		GE 179	10/13/64	
 <td>526.277</td> <td>Craig Breedlove</td> <td colspan="2">Spirit of America</td> <td>GE 379</td> <td>10/15/65</td> <td>b></td>	526.277	Craig Breedlove	Spirit of America		GE 379	10/15/65	b>
10/13/4 Spirit of	536.712	Art Arfons	Green Monster		GE 179	10/27/65	ta> <ta></ta>
536.712<td>555.127</td><td>Craig Breedlove</td><td colspan="2" rowspan="4">Spirit of America, Sonic 1 Green Monster Spirit of America, Sonic 1 Blue Flame</td><td>GE 179</td><td>11/2/65</td><td>J79</td>	555.127	Craig Breedlove	Spirit of America, Sonic 1 Green Monster Spirit of America, Sonic 1 Blue Flame		GE 179	11/2/65	J79
10-	576.553	Art Arfons			GE 379 11/7/65		
	600.601	Craig Breedlove			GE 379	11/15/65	s/65
<bgi< td=""><td>622.407</td><td>Gary Gabelich</td><td colspan="2">Rocket 10/23/70</td><td>d></td></bgi<>	622.407	Gary Gabelich			Rocket 10/23/70		d>
179	633,468	Richard Noble	Thrust 2		RR RG 146	10/4/83	-
Galoetich <td>763.035</td> <td>Andy Green</td> <td colspan="2">Thrust SSC</td> <td>RR Spey</td> <td>10/15/97</td> <td>></td>	763.035	Andy Green	Thrust SSC		RR Spey	10/15/97	>
RR RG 146					0,707		ntly
	pics	Days for	Class	Timing		c/b><	/td>
De	sign For Ux	Monday		8:30 - 10		[
	sign Thinking			12:30 - 3	2:30 pm	iming	
Em</td <td>pathy Map</td> <td>Tuesday</td> <td></td> <td>9:30 - 1:</td> <td>1:30 pm</td> <td>oncla</td> <td>y 12:30 ±</td>	pathy Map	Tuesday		9:30 - 1:	1:30 pm	oncla	y 12:30 ±
 Em	otional Intellig	ence Wednesd	lay			d><0>	12:30 ±
rmesnan="2"> He	ability	Thursday		8:30 - 10:30 am			
	lity	Friday		8:30 - 10	2-20	d> <td< td=""><td>>8:30</td></td<>	> 8:30

parameter	parameter		Squamous cell carcinoma (n = 229)			Ade	nocarcino	ma (n = 150)	p-value
- Surgical complications (%)						П			
- General cor	nplication	s (%)	37.9			33.7			0.447
-Total compli	ications (9	6)	40.7			28.0			0.012*
- Length of I	CU – stay	(d)	68.1			57.6			0.039*
- 30-day-mor	atility (%)				Ī			
- Mortality (%)		11 (2-17)	6)		8 (1-	-107)		0.0001*
			6.1			6.1			0.8569
			15.1			6.6			0.0001*
Speed (mph)	Driver		Cat ^a		Engine		Date		
407.447	C _{Telia} Bn	ee(dlov	e Spirit	Spirit of Annerica			8/5/63		
413.199	Tom Gre	en	Wingf	oot: Express	WE J46		10/2/64		
434.22	Art Arfo	15	Green	Monster	GE J79		10/5/64		
468.719	Craig Bn	edlove	Spirit	of Anerica	GE J79	1	10/13/64		
526.277	Craig Br	edlove	Spirit	pirit of Anerica			10/15/65		
536.712	Art Arfor	15	Green Monster		GE J79		10/27/65		
555.1.27	Craig Br	edlove	Spirit of Anerica, Sonic 1		GE J79		11/2/65		
576.553	Art Arfor	15	Greer,	reer, Monster GE			11/7/65		
600.601	C _{Traig} Br	eqdlov	e Spirit	of Anerica, Sonic 1	GE J79		1.1/15/65		
622.407	Gary Gal	oelich	Blue F	lame	Rocket		10/23/70		
633.468	Richard 1	Noble	Thrus	12	RR RG	146	10/4/83		
763.035	Antly Gr	een	Thrus	SSC	RR Spe	y	10/15/97		
Topics		Days 1	for Class Timing						
Design For Ux		Monc	8:30 - 10:30 am						
Design Thinking Monc		Monc	iay	12:30 ± 2:30 pm					
Empathy Map Tueso		Tuesd	ay	O:30 = 11:30 gm					
Emotional In	telligence	Wedn	esday	Q.30 - 11:30 gm					
Usability Thur		Thurs	sclay 8:30 – 10:30 am						

8:30 = 10:30 am

8:30 - 10:30 am

Utility

Accessibility

Friday

Fridzy

Changes made in MTL-TabNet and Layout Parser Code

- MTL-TabNet has been enhanced to process multiple images by directly reading them from a folder in a sequential manner. This modification improves efficiency when dealing with a large number of images.
- Removal of unncessary code like taking input images several times has been removed and edited the code of MTL-Tabnet to make it more generalised
- Removal of unnecessary code in Layout parser and used only for Layout analysis and extract tables from a PDF file. It detects the layout, crops the images around the table blocks, and saves them. It generates JSON data with table coordinates and writes it to a file
- JSON data generation has been integrated into the Layout parser, producing a file that contains the coordinates of the extracted tables. This information serves as a reference for further processing and integration into the final output.
- The updated code in both MTL-TabNet and the Layout parser removes redundancy and unnecessary steps, These changes contribute to improved overall performance and code maintainability.

July 10, 2023

Connecting Layout Parser to MTL Tabnet

- Table Detection with Layout Parser
 - Utilizes Layout Parser to detect tables within a document.
 - Provides cropped images of table parts for further processing.
 - Efficiently identifies table regions and extracts them.

Automated Whole PDF Processing:

- Enables processing of entire PDF documents.
- Detects all tables present within the document. Allows for bulk processing and saves manual effort.
- Extracted tables are saved as cropped images in a designated folder.
- Sending Cropped Images to MTL-Tabnet:
- Cropped images obtained from Layout Parser are sent as input to MTL-Tabnet and MTL-Tabnet extracts HTML table element with order of tables Benefits
- Enhaned table recognition accuracy with combined capabilities
- processing of multiple tables within a document or bulk pdf is easy

Overall Output image after integration:



Show Live Terminal output

Future work

- Integration of Borderless table work to RAVI output HTML
- To identify where to place the table in the output HTML .