# Supplementary Material: Table Structure Recognition using Top-Down and Bottom-Up Cues

Sachin Raja, Ajoy Mondal, and C V Jawahar

Center for Visual Information Technology, International Institute of Information Technology, Hyderabad, India sachinraja130gmail.com, {ajoy.mondal,jawahar}@iiit.ac.in

# 1 Organization

We organize our supplementary paper in the following manner:

- Section 2 contains additional details about our methodology including the post-processing step to obtain the final XML output.
- Section 3.1 contains information about the datasets used for experiments and explains the ground-truth unification process.
- Section 3.2 briefly explains various previously published methods in the space of table structure recognition including the input modalities used for predictions.
- Section 3.3 talks about the very specific details of implementation for training our model.
- Section 3.4 explains the evaluation method in greater detail which will be useful for reproducing our results. It explains various assumptions and reasons for choosing the specific metrics.
- Section 3.5 lists out various challenges encountered for evaluating and comparing our method with the existing ones, and also talks about the ways in which we handled them in our work.
- Section 4 presents additional quantitative and qualitative results for a detailed analysis.

# 2 TabStruct-Net

Our Tabstruct-Net is a data-driven and an end-to-end trainable architecture for the prediction of table structure from a given table image, that combines topdown and bottom-up methods. As a first step, the input table image is broken down into individual cells using the cell detection network of the Tabstruct-Net. We call this as the top-down step of the process. After detecting individual cells, the next step is to obtain the entire table structure by building relevant row and column associations between the detected cells. This is done using the structure recognition network of the Tabstruct-Net and we call this as the bottom-up step of the process.

## 2.1 Post-processing to Get XML Output

After the cell bounding boxes along with the row and column adjacency matrices are obtained, an XML file is generated using an heuristic based algorithm. It works as follows:

- For row assignments, sort all bounding boxes by their  $start_y$  coordinates, and initialize a row belonging list for every cell.
- Assign a row belonging index (starting from 0) to the cell  $c_i$  and assign the same row index to all other cells that are connected to  $c_i$  in the row adjacency matrix.
- Increment the row index and repeat the above step until all the cells are assigned at least one row belonging index.
- For each cell, SR is the minimum of indexes in the row belonging list, and ER is the maximum of indexes in the row belonging list.
- Similarly, for column assignments, sort all bounding boxes by their  $start_x$  coordinates, and initialize a column belonging list for every cell.
- Assign a column belonging index (starting from 0) to the cell  $c_i$  and assign the same column index to all other cells that are connected to  $c_i$  in the column adjacency matrix.
- Increment the column index and repeat the above step until all the cells are assigned at least one column belonging index.
- For each cell, SC is the minimum of indexes in the row belonging list, and EC is the maximum of indexes in the column belonging list.

We use Tesseract [43] to extract the content of every predicted cell. Once SR, ER, SC and EC values (referred to as cell spanning values) and its content are obtained for every predicted cell, an XML file is created with these cell spanning values along with bounding box coordinates (top-left and bottom-right of the cell) and its content.

## 3 Experiments

#### 3.1 Dataset

We use various benchmark table structure recognition datasets — sciTSR [14], sciTSR-COMP [14], ICDAR-2013 table recognition [18], ICDAR-2019 cTDaR archival [19], UNLV [28], Marmot extended [12], TableBank [11] and PubTabNet [13] datasets for extracting structure information of tables. Statistics of these datasets are listed in Table 1 (main paper).

Our TabStruct-Net makes an assumption that all cells belonging to the same column are aligned with respect to x coordinates and cells belonging to the same row are aligned with respect to y coordinates. SciTSR[14], SciTSR-COMP [14] and ICDAR-2013 [18] datasets have ground truth bounding boxes at the level of cell's content (box is the smallest rectangular block that encapsulates the cell's content). To handle this, we expand the bounding boxes of every cell in a row and column to get maximum sized content-level box in a particular row and column.

	hyper		supern		ethers*		Freq	uency of sub	stance abuse	3			AGHed	usiad disab	ity rate			Uhadjusted	disability rate	
	1996	charge tince (990	1996	change since 1990	1996	Substance	In the last			Never			203	2	5		205		200	
Austria	12	+3	52	+11	36		month	year	a year				rale Marcin d		Namino	- E	Į.	lenin d	Mar	in di Dife
Belgium Luxenbourg	16	83	70	+5	14	Cannabis	9.6%	4.0%		80.1%			enor (s)	Fatirate	ence (a) D	therease in	Estimate la	nor (s)	simale enco	
Denmark	17	8 Ra	59	+8	24	Inhalants	2.7%	1.6%		91.6%	Alce	~	12.6			1-0.5	167	13	187	ũ
Finland	22	na Ka	51	+0	27	Ecstasy	2.1%	0.9%		94.0%	EL28	B		171		30	044	-	024	
Tance	51	+16	44		8	Anabolic steroids	2.1%	0.2%		97.0%	laie		173 14	170		100	21	57	22	8
iemany	24	+8	52	+7	24	Magic mushrooms	2.1%	0.5%		96.3%	Fenale		19.0	17.6 13.3		45	175 801	14	17.4 19.1	l
ireed	8	+5	51	8.5.	44	Heroin	0.5%	0.0%	0.5%	99.0%	CENSE		1276 87	32.0	-	-	843	-	124	-
reland	12	na	41	8.4.	47	Amphetamines	2.5%	0.7%	0.7%	96.1%	White alon		175 63	17.4	6	2.05	101	10	10.5	13
12h	13	+/3	39	8.4.	48	Ketamine	2.1%	0.0%	0.5%	97.4%	Not Hist					172	19.7	7	19.3	
ichethods	-	+3	82	+7	13	LSD	2.3%	0.9%	1.4%	95.4%	Backalone		181 64 232 6	222	64	1	18.5 19.7 20.4	i i	18.5 19.3 20.3	
tertugal	42	+42	28	+10	30	Cocaine	4.4%	1.6%	1.8%	92.2%	W/Hso	nic.	23.5	223		1	207		207	8
rein	34	+22	25	+1	31	GHB	1.1%	0.0%	0.0%	98.9%	Asian Alon		145 13	17.6 22.2 22.3 14.5	11		20.7 12.4 12.5		207 131	
ineder	13	84	64	+4	23	Tranquillisers	3.4%	1.4%	1.8%	93.4%	NetHiso	ńć	146 13		1	-0.2	123	3	101	11
								0.0%		98.9%			134 65	17.0		10	12.1	10	12.2	18
UK	45	+2§	42 Supe	+2	13 others*	Rtain	1.1%				Historica			uster desti	ib rate	NW	141	Unaduste		
IK	hyper	rmarkets	supe	markets	others		Fre	quency of su	bstance abu	ise	HISPICS		Ageat	usiaci disabi		New 1			é disability rat	
K					others <sup>e</sup> 1996	Substance	Free In the last	auency of su	bstance abu More than			997	Age ed 2005	20	0	-	205		d deach lity rai 2010	•
	hyper	rmarkets change	supe	markets	others <sup>e</sup> 1996	Substance	Free In the last month	auency of su In the last year	bstance abu More than a year	ISe Never		997	Age ed 2005 maia Marcin af	20	rð Margin af		1	i Nergin of	d disability rat 2010 Mar	¢ ģnaf 1
stria	hrper 1996	rmarkets change	supe 1996	markets change since 1991	others <sup>e</sup>	Substance Cannabis	Free In the last month 9.6%	In the last year 4.0%	bstance abu More than a year 6.3%	Never		997	Age ed 2005 maia Marcin af	20	0		1	i Nergin of	d deach lity rai 2010	¢ ģnaf 1
stria Igium/Lutembourg	hyper 1996 12	rmarkets change	supe 1996 52	markets change since 1991 +11	others <sup>4</sup> 1996	Substance Cannabis Inhalants	Free In the last month 9.6% 2.7%	auency of su In the last year 4.0% 1.6%	bstance abu More than a year 6.3% 4.1%	80.1%		997	Age ed 2005 maia Marcin af	20	rð Margin af		1	i Nergin of	d disability rat 2010 Mar	¢ ģnaf 1
stria Igium/Lutembourg arrark	https://www.international.com/ 1996 12 16 17 22	markets change since 1980 +3 -	52 59 59 59 51	markets change since 199 +11 +5	others <sup>4</sup> 1996 36 14 24 27	Substance Cannabis Inhalants Ecstasy	Free In the last month 9.6% 2.7% 2.1%	In the last year 4.0% 1.6% 0.9%	bstance abu More than a year 6.3% 4.1% 3.0%	80.1% 91.6%		igory Esti	Age ed 2005 maia Marcin af	20	rð Margin af	liference	Estinata e 187	i Nergin of	d disability rat 2010 Mar	e ginaf 1 r(±)
stria Igiural Lavembourg arrark Iland	hyper 1996 12 16 17 22 51	change since 1980 +3 	52 70 59 51 44	markets change since 1991 +11 +5 +8 -J -	others <sup>4</sup> 1996 36 14 24 27 5	Substance Cannabis Inhalants Ecstasy Anabolic steroids	Free In the last month 9.6% 2.7% 2.1% 2.1%	uency of su In the last year 4.0% 1.6% 0.9% 0.2%	bstance abu More than a year 6.3% 4.1% 3.0% 0.7%	80.1% 91.6% 94.0%	Ca Mpro	sgory Est	Age ed 2005 maia Marcin af	20 Estinate 18.1	rti Mangin of ernor (±) 0.3 0.4	ilference 140.5 40.3	1 Estinata 187 17.3	i Hangin of anor (±) E 11.3 11.4	d disability rat 2010 Mar	e ginaf [ 1] 0,3 0,4
stria Igium/Luxembourg arrark Iaad nice many	https://www.international.com/ 1996 12 16 17 22	markets change since 1980 +3 -	52 70 59 51 44 52	markets change since 199 +11 +5	others <sup>4</sup> 1996 36 14 24 27 5 24	Substance Cannabis Inhalants Enstasy Anabolic steroids Magic mustrooms	Free In the last month 9.6% 2.7% 2.1% 2.1% 2.1%	uency of su In the last year 4.0% 1.6% 0.9% 0.2% 0.5%	bstance abu More than a year 6.3% 4.1% 3.0% 0.7% 1.1%	80.1% 91.6% 94.0% 95.3%	Ca Mpro	igory Esti	Age ed 2005 maia Marcin af	20	rð Margin af	liference	Estinata e 187	i Nergin of	d disability rat 2010 Mar	e gindi E r(±) 0,3
stria Igium/Luxembourg arrark Itand Itand many ecce	https://www.international.com/ 1996 12 16 17 22 51 24 5	markets change since 1980 +3 -	52 70 59 51 44 52 51	markets change since 1991 +11 +5 +8 -J -	others* 1996 36 14 24 27 5 24 44	Substance Cannabis Inhalants Esstasy Anabolic steroids Macic mushrooms Heroin	Free In the last month 9.6% 2.7% 2.1% 2.1% 2.1% 0.5%	uency of su In the last year 4.0% 1.6% 0.9% 0.2% 0.5% 0.0%	bstance abu More than a year 6.3% 4.1% 3.0% 0.7% 1.1% 0.5%	ISE Never 91.6% 94.0% 97.0% 96.3% 99.0%	Ca Mpro	sgory Est	Age ed 2005 maia Marcin af	20 Estinate 18.1	rti Mangin of ernor (±) 0.3 0.4	ilference 140.5 40.3	1 Estinata 187 17.3	i Hangin of anor (±) E 11.3 11.4	d disability rat 2010 Mar	6 ginaf 1 r(±) 0.3 0.4
stria Igiun/Luxenbourg armark Iland ance armany recee Iand	hrper 1996 12 16 17 22 51 24 5 12	markets change since 1990 +3 -	supe 1996 52 70 59 51 44 52 51 41	markets change since 199 +11 +5 +8 -J - - +7	others <sup>6</sup> 1996           36           14           24           27           5           24           44           47	Substance Cannabis Inhalants Ecstasy Anabolic steroids Matic mushrooms Heroin Amphetamines	Free In the last month 9.6% 2.7% 2.1% 2.1% 2.1% 0.5% 2.5%	In the last year 4.0% 1.6% 0.9% 0.2% 0.2% 0.5% 0.0%	bstance abu More than a year 6.3% 4.1% 3.0% 0.7% 1.1% 0.5% 0.7%	80.1% 91.6% 94.0% 97.0% 96.3% 99.0% 99.0%	Ca Al pec Hale. Fenale	1907) Est	Agened 2005 mala Vangin d' encr (c) 185 0.0 173 0.4 190 0.0	20 Estinate 18.1	rti Mangin of ernor (±) 0.3 0.4	ilference 140.5 40.3	Estinata e 187 17.3 20.1 188	i Hangin of anor (±) E 11.3 11.4	d disability rat 2010 Mar	e ghof ( 1) 0,3 0,4 0,4 0,4
stria igiumLuxenbourg annak aneo aneo anany eeee land h	https://www.international.com/ 1996 12 16 17 22 51 24 5 12 13	markets change since 1980 + 3 -	52 70 59 51 44 52 51 41 39	markets change since 199 +11 +5 +8 -1 -1 - - 7 Ra. Ra. Ra. Ra.	others <sup>6</sup> 1996 36 14 24 27 5 24 44 47 48	Substance Cannabis Inhalants Esstasy Anabolic steroids Megic muskrooms Heroin Amphetamines Ketamine	Free In the last month 9.6% 2.7% 2.1% 2.1% 2.1% 0.5% 2.5% 2.5% 2.1%	Luency of su In the last year 4.0% 1.6% 0.9% 0.2% 0.5% 0.5% 0.0% 0.7%	bstance abu More than a year 6.3% 4.1% 3.0% 0.7% 1.1% 0.5% 0.5% 0.5%	80.1% 91.6% 94.0% 97.0% 96.3% 99.0% 99.1% 96.1%	Ca Mipe Penale Write admon	1907) Est	Age ed 2005 maia Marcin af	20 Estinate 18.1	rti Mangin of ernor (±) 0.3 0.4	ilference 140.5 40.3	1 Estinata 187 17.3	i Hangin of anor (±) E 11.3 11.4	6 disability rat 2010 Mari 18.7 17.4 18.8 18.5 18.8	e r(±) 0.3 0.4 0.4 0.4
stria igiumLuxenbourg anark anark ance anary cece land hr sherlands	httpp: 1996 12 16 17 22 51 24 5 12 13 5 5	markets change since (990) +3 - R.g. +16 +8 +5 R.g. +13 +3	52 70 59 51 44 52 51 44 52 51 44 39 82	markets change since 1999 +11 +5 +8 -1 +7 n.a. n.a. +7 +7	others*           1996           36           14           24           27           5           24           44           47           48           13	Substance Cannabis Inhalants Ecstasy Anaboio steroids Macio mushrooms Heron Amphetamines Kolamne LSD	Free           In the last month           9.6%           2.7%           2.1%           2.1%           2.1%           2.1%           2.5%           2.1%           2.1%	auency of su In the last year 4.0% 1.6% 0.9% 0.2% 0.9% 0.2% 0.5% 0.0% 0.7% 0.0%	bstance abu More than a year 6.3% 4.1% 0.7% 0.7% 0.7% 0.5% 1.4%	ISE Never 80.1% 91.6% 94.0% 97.0% 96.3% 99.0% 96.1% 96.1% 97.4% 95.4%	Ca All per Hele. Fanale White advess Back advess Back advess	1907) Eal 20	Agened 2005 mala Vangin d' encr (c) 185 0.0 173 0.4 190 0.0	20 Estinate 18.1	rti Mangin of ernor (±) 0.3 0.4	ilference 140.5 40.3	Estinata e 187 17.3 20.1 188	i Hangin of anor (±) E 11.3 11.4	0 disability rat 2010 18.7 18.7 18.8 18.8 18.8 20.3	0 r(H) 0.3 0.4 0.4 0.4 0.5
ssria Ingund.uxembourg annek ance many ecce Ind Y cherlands frigal	https://www.international.com/ 1996 12 16 17 22 51 24 5 12 13 5 42	markets change since 1990 +3 -	supe 1996 52 70 59 51 44 52 51 44 52 51 39 82 28	markets change since 199 +11 +5 +8 -1 - +7 R.a. R.a. R.a. +7 +7 +10	others*           1996           36           14           24           27           5           24           44           47           48           13           30	Substance Cannabis Inhafartis Ecstasy Anabolic stenids Mapic mushrooms Herroin Amphetamines Esp Cocaine	Free In the last month 9.6% 2.7% 2.1% 2.1% 2.1% 2.1% 2.5% 2.5% 2.3% 4.4%	auency of su In the last year 4.0% 1.6% 0.9% 0.2% 0.2% 0.2% 0.2% 0.0% 0.7% 0.0% 0.0% 0.0%	bslance abu More than a year 6.3% 4.1% 3.0% 0.7% 1.1% 0.5% 0.7% 0.5% 0.5% 1.4% 1.8%	ISE Never 91.6% 94.0% 96.3% 99.0% 96.1% 95.4% 95.4% 95.4%	Ca Miles Fenale Write abno Not Hass Badk along Not Hass	1997) Est 18 16	Ageed 2005 mais Vergin of encr.(2) 18.5 0.0 17.9 0.4 19.0 0.5 17.9 0.5 18.1 0.4 18.1 0.4 18.1 0.4 18.2 0.5 18.1 0.4 18.2 0.5 18.1 0.4 18.2 0.5 18.1 0.4 18.2 0.5 18.1 0.4 18.2 0.5 18.5 0.5 19.0	20 Estinate 18.1	rti Mangin of ernor (±) 0.3 0.4	ilference 140.5 40.3	Estinata e 187 17.3 20.1 185 197 20.4 20.7	i Hangin of anor (±) E 11.3 11.4	6 disability rat 2010 Mari 18.7 17.4 18.8 18.5 18.8	e rH 0.3 0.4 0.4 0.4 0.4 0.4 0.4 0.7 0.7
istria ligiumLuvernbourg smrark altad atec return return br sherlands writegal alian	hnpee 1996 12 16 17 22 51 24 5 12 13 5 42 34	markets change since (990) +3 - R.g. +16 +8 +5 R.g. +13 +3	supe           1996           52           70           59           51           44           52           51           41           39           82           28           25	markets change since 1999 +11 +5 +8 -1 - - +7 na. na. na. +7 +0 +7 +1 +8 - +7 +7 +8 +7 +7 +7 +7 +7 +7 +7 +7 +7 +7	others*           1996           36           14           24           27           5           24           44           47           48           13           30           31	Substance Carnabis Inhalants Ecstasy Anabolic steroids Macin mushrooms Heroin Amphetamines Kotamine SD Coccaine GHB	Free In the last month 9.6% 2.7% 2.1% 2.1% 2.1% 2.1% 2.5% 2.5% 2.5% 2.5% 2.3% 2.3% 2.3%	auency of su In the last year 4.0% 1.6% 0.3% 0.2% 0.2% 0.7% 0.0% 0.7% 0.0%	bstance abu More than a year 6.3% 4.1% 3.0% 0.7% 0.7% 0.5% 0.7% 0.5% 0.5% 1.4% 1.4% 0.0%	80.1% 91.6% 94.0% 96.1% 96.1% 95.4% 95.4% 95.4% 95.4% 95.4% 95.9%	Ca Miper Finite White sizes Data sizes Mittisse Bata sizes Mittisse Bata sizes Mittisse Bata sizes	1977 - Est 10	Ageed 2005 mais Vergin of encr (c) 188 00 173 04 190 05 173 05 191 05 193 05 193 05 193 05 194 193 193	20 Estinate 118.1 17.6 18.3 18.3 17.4 17.6 17.6 17.6 17.6 17.6 17.6 17.6 17.6	rti Mangin of ernor (±) 0.3 0.4	ilference 140.5 40.3	Estinata 187 17.3 20.1 188 19.7 20.4 20.4 20.7 12.4	i Hangin of anor (±) E 11.3 11.4	0 disability rat 2010 18.7 18.7 18.8 18.8 18.8 20.3	e r(H) 0.3 0.4 0.4 0.4 0.4 0.4 0.7 0.7 1.0
IK 2007a Information Informati	https://www.international.com/ 1996 12 16 17 22 51 24 5 12 13 5 42	markets change since 1990 +3 -	supe 1996 52 70 59 51 44 52 51 44 52 51 39 82 28	markets change since 199 +11 +5 +8 -1 - +7 R.a. R.a. R.a. +7 +7 +10	others*           1996           36           14           24           27           5           24           44           47           48           13           30	Substance Cannabis Inhafartis Ecstasy Anabolic stenids Mapic mushrooms Herroin Amphetamines Esp Cocaine	Free In the last month 9.6% 2.7% 2.1% 2.1% 2.1% 2.1% 2.5% 2.5% 2.3% 4.4%	auency of su In the last year 4.0% 1.6% 0.9% 0.2% 0.2% 0.2% 0.2% 0.0% 0.7% 0.0% 0.0% 0.0%	bslance abu More than a year 6.3% 4.1% 3.0% 0.7% 1.1% 0.5% 0.7% 0.5% 0.5% 1.4% 1.8%	ISE Never 91.6% 94.0% 96.3% 99.0% 96.1% 95.4% 95.4% 95.4%	Ca Miles Fenale Write abno Not Hass Badk along Not Hass	1907) Eali 10	Ageed 2005 mais Vergin of encr.(2) 18.5 0.0 17.9 0.4 19.0 0.5 17.9 0.5 18.1 0.4 18.1 0.4 18.1 0.4 18.2 0.5 18.1 0.4 18.2 0.5 18.1 0.4 18.2 0.5 18.1 0.4 18.2 0.5 18.1 0.4 18.2 0.5 18.5 0.5 19.0	20 Estinate 18.1	rti Mangin of ernor (±) 0.3 0.4	ilference 140.5 40.3	Estinata e 187 17.3 20.1 185 197 20.4 20.7	i Hangin of anor (±) E 11.3 11.4	0 disability rat 2010 18.7 18.7 18.8 18.8 18.8 20.3	e r(H) 0.3 0.4 0.4 0.4 0.4 0.4 0.5 0.4 0.7

Fig. 1. Ground truth unification. content-level bounding boxes are given in ground truth as shown in **First Row**. We make content-level bounding boxes into cell-level bounding boxes as shown in **Second Row**.

## 3.2 Baseline Methods

We compare the performance of our TabStruct-Net against seven following benchmark methods.

**DeepDeSRT** [7]: This method leverages semantic segmentation approach to localize each row and column from the table image. This method outputs table as a grid-like structure and fails to identify multiple row and multiple column spanning cells. Since no code is available, we implement our own version of this method. Since this method works by predicting every row and column, the SciTSR training dataset is pre-processed to obtain row and column level coordinates before training.

**TableNet** [12]: This method uses semantic segmentation approach to extract table and column masks, and segments rows by identifying words present in different columns (extracted using Tesseract OCR [43]) that occur at the same horizontal level. For comparison against other methods, we directly use the results reported by the authors.

**GraphTSR** [14]: This method consists of edge-to-vertex and vertex-to-edge graph attention blocks to compute vertex and edge representations in a latent space, and finally classify each edge as 'horizontal', 'vertical' or 'no-relation'. It uses absolute and relative positions of every cell extracted from the PDF to compute initial vertex and edge features.

**SPLERGE** [10]: This method comprises of two deep learning networks that perform splitting and merging operations in sequence to predict fine grid-like

table structure and to predict merged cells which span multiple rows/columns. Split method shows an improved performance when additional PDF extracted meta-features are provided along with the table image. For the split model (to obtain the basic grid of the table), authors pre-process the ground truth by maximizing the row and column separator regions without intersecting any non multiple row or column spanning cell. For the merge model (to identify cells that span multiple rows or columns), the authors prepare the ground truth by identifying grid elements that span multiple cells and set the merging probability in the respective directions. Further, for evaluating this method on ICDAR-2013 [18] dataset, the authors realized that merge method did not work with a good performance, and hence, introduced the following specific heuristics to merge cells instead:

- Merge cells where separators pass through text.
- Merge adjacent blank columns with cells that have been formed by merging many cells.
- Merge adjacent blank rows with content cells.
- Split columns that have a consistent white-space gap between vertically aligned text.

 $DGCNN^*$  [9]: Authors formulate the problem as a graph learning problem to predict whether every pair of words belongs to the same cell, row and/or column or not. Their architecture consists of a visual network, an interaction network and a classification network. For evaluation purposes, table image along with word-level bounding boxes is provided as inputs.

**Bi-directional GRU** [15]: Given the table image, this method uses two bidirectional GRUs to establish row and column boundaries in a context driven manner. This method however fails to localize multiple row and/or multiple column spanning cells.

**Image-to-Text** [11]: This method utilizes an Image-to-Markup model to predict a markup-like output of a given table image. It consists of a CNN based encoder to compute visual features and an LSTM based decoder to produce markup output.

## 3.3 Implementation Details

Our TabStruct-Net model<sup>1</sup> has been trained and evaluated with table images scaled to a fixed size of  $1536 \times 1536$  while maintaining the original aspect ratio as the input. While training, cell-level bounding boxes along with row and column adjacency matrices (prepared from start-row, start-column, end-row and end-column indices) are used as the ground truth. We use NVIDIA TITAN X GPU with 12 GB memory for our experiments and a batch-size of 1. Instead of using

<sup>&</sup>lt;sup>1</sup> Our code will be available publicly

 $3\times3$  convolution on the output feature maps from the FPN, we use a dilated convolution with filter size of  $2\times2$  and dilation parameter of 2. Also, we use the ResNet-101 backbone that is pre-trained on MS-COCO [48] dataset. To compute region proposals, we use 0.5, 1 and 2 as the anchor scale and anchor box sizes of 8, 16, 32, 64 and 128. Further, for generation of the row/column adjacency matrices, we use 2400 as the maximum number of vertices keeping in mind dense tables. Since every input table may contain hundreds of table cells, training can be a time consuming process.

To achieve faster training, we employ a 2-stage training process. In the first step, we use 2014 anchors and 512 RoIs. With this setting, the model is able to learn high and low level features but resulted in a large number of false negatives. To combat this, network is trained with 3072 anchors and 2048 RoIs. This significantly reduces the number of false negatives. For the first step, we train a total of 30 epochs, for the first 8, we train all FPN and subsequent layers, for the next 15, we train FPN + last 4 layers of ResNet-101 and for the last 7 epochs, we train all the layers of the model. For the second step, we train a total of 10 epochs, for the first 3, we train all FPN and subsequent layers, for the next 4, we train FPN + last 4 layers of ResNet-101 and for the next 4, we train FPN + last 4 layers of ResNet-101 and for the last 3 epochs, we train all the layers of ResNet-101 and for the last 3 epochs, we train all the layers of the model. During both the stages, we use 0.001 as the learning rate, 0.9 as the momentum and 0.0001 as the weight decay regularisation.

#### 3.4 Evaluation Measures

**Details of Evaluation Criteria:** For comparison against most of the existing methods, we use the precision, recall and F1 score [14,18,28]. Before evaluating performance of structure recognition, it is important to understand the specific cases in which detected cells are taken into consideration for structure recognition:

- We consider a detected cell to be a true positive if it overlaps with the ground truth cell bounding box is more than 0.5.
- During evaluation of structure recognition, cells that have no content (i.e., empty or blank cells) are discarded. It means that adjacency relations that involve a blank cell are not taken into consideration.

To evaluate the performance of structure recognition, adjacency relations between every cell (with content) are generated with their horizontal and vertical neighbors. This predicted relation list is then compared with the ground truth list to generate precision, recall and F1 measures.

As per [18], this method accounts for the standard evaluation measures for table structure recognition for the following reasons:

- It provides for a simple way to account for errors in the scenarios of complex table layouts containing blank cells, and cells that span multiple rows and/or columns.
- It accounts for evaluation of both physical as well as logical structure prediction methods as it is not dependent on the bounding box coordinates information.

Further, we present our results on both micro-averaged and macro-averaged precision, recall and F1 scores which are defined as following:

- Micro-averaged: In this case, the confusion matrix parameters are gathered across all the data points collectively in the test dataset to compute precision, recall and F1 scores.
- Macro-averaged: In this case, the confusion matrix parameters are gathered individually and then averaged across all the documents in the test dataset to compute precision, recall and F1 scores.

#### 3.5 Experimental Setup

One major challenge in the comparison study with the existing methods is the inconsistent use of additional information (e.g., meta-features extracted from the PDFs [10], content-level bounding boxes from ground truths [12,14] and cell's location features generated from synthetic dataset [9]).

For the unification of fair comparison for table structure recognition, we divide all inconsistencies into several levels - (i) inconsistency with respect to input modalities, (ii) inconsistency with respect to annotation levels, (iii) inconsistency with respect to representation of table structure, (iv) inconsistency with respect to evaluation methods, and (v) inconsistency with respect to way of computing evaluation scores.

Inconsistency with respect to input modalities: Section 3.2 describes that many methods for table structure recognition work with table images alone [7,10,15], several other assume additional information in the form of metafeatures or bounding boxes around every word or cell-content extracted from the PDFs [9,12,14]. This makes it difficult to compare these methods under a unified scenario. To take care of this problem, we define two different experimental setups - (a) Setup-A (S-A) where only table image is used as an input to the structure recognition model and (b) **Setup-B** (S-B) where table image along with additional meta-features such as low-level content bounding boxes are used as an input to the structure recognition model. We present our results under both the experimental setups for a thorough comparison of our work against most of the recent methods in this space. To achieve this, we train our model for cell detection as well as structure recognition collectively for S-A. For evaluation in S-B, instead of predicting cell bounding boxes from the image, we use the table image and the low-level bounding box information from OCR or ground truth to be able to directly and fairly compare our method.

Inconsistency with respect to annotation levels: It is important to note that training of Tabstruct-Net assumes cell-level bounding boxes in a way that all cells that (a) have the same SR indices having same y1 coordinates, (b) have the same SC indices having same x1 coordinates, (c) have the same ER indices having same y2 coordinates, and (d) have the same EC indices having same x2 coordinates. This assumption is necessary for our alignment loss

function to work properly. However, different datasets for physical table structure recognition have ground truth annotations defined in different ways. UNLV and ICDAR-2019 archival datasets have ground truth annotated at the cell-level. SciTSR[14] and ICDAR-2013 [18] datasets have ground truth annotation defined at the content-level (cells' bounding box is the smallest rectangle that covers entire content of the cell). To be able to use those for training, we pre-process the ground-truth to obtain cell-level bounding boxes (as explained in Section 3.1). Please note that this pre-processing step is only done for the training process. Similarly, ground-truth bounding boxes of the synthetic dataset proposed in [9] are provided at the word-level. To obtain cell-level bounding boxes, we use the ground-truth cell adjacency matrix and word-level bounding boxes to obtain content-level bounding boxes. During the testing time in S-A, however, to compute if a detected cell is a true positive, we use the original ground-truth bounding boxes (either at cell-level or content-level), and not the pre-processed ones. Similarly while testing in S-B, we use the original content-level or cell-level bounding boxes as the additional input instead of the pre-processed ones. This ensures consistency while comparing our methods against previously published ones.

Inconsistency with respect to representation of table structure: We broadly classify table structure methods into two categories - (a) physical structure predicting methods that predict cell-level bounding boxes along with their associations [7,10,15] and (b) logical structure predicting methods [9,11,12,14] that predict only cell associations. In our work, we standardize our representation as described in Section 3.5, which allows us to directly compare methods in both the experimental setups. To compare the results of TabStruct-Net on logical structure prediction, we generate the mark-up string from the post-processed XML output of TabStruct-Net in the same format as TableBank [11] and PubTabNet [13] by extracting only the structure without cells' coordinates and content.

**Inconsistency with respect to evaluation methods:** While most of the previously published methods for table structure recognition use precision, recall and F1 scores as described in [18], there are some inconsistencies in this aspect as well. In [9], authors use true positive rate (TPR), false positive rate (FPR) and absolute accuracy on the predicted adjacency matrix to compute performance. In order to standardize evaluation with [9], we infer neighboring cell relations from their output to ensure consistency. Further, [11] use BLEU scores to compare their output with the ground truth. Since our method generates and XML output from the model's predictions, we bring our output to the same format as [11] to ensure a direct and fair comparison on the TableBank dataset [11].

Inconsistency with respect to way of computing evaluation scores: To fairly compare TabStruct-Net against previous methods, we list both micro as well as macro (document) - averaged results on all the test datasets. However, it is important to note that for TableBank [11] and PubTabNet [13] datasets, where

we evaluate our results on the markup output of the model, we only consider macro(document)-averaged results.

# 4 Results on Table Structure Recognition

## 4.1 Micro-averaged Results

Tables 1-4 show the micro-averaged results of various methods for structure recognition on multiple datasets. From the tables, it can be observed that our method outperforms previously published works under multiple kinds of experimental settings. Further, it is important to note that the tables use an IoU threshold of 0.5 to identify true positive cells for experiment setup S-A. We also show the precision, recall and F1 measures on various IoU thresholds to better interpret the performance of the cell detection module of TabStruct-Net.

Method	Training		Exp.	$\mathbf{P}\uparrow$	$\mathbf{R}\uparrow$	$\mathbf{F1}\uparrow$
	Dataset	#Images	Setup			
Deepdesrt [7]	ICDAR-2013-partial	0.124K	S-A	0.959	0.874	0.914
Splerge [10]	icdar-2013-partial	0.124 K	S-A	0.917	0.911	0.914
Bi-directional GRU [15]	icdar-2013-partial	0.124 K	S-A	0.969	0.901	0.934
Tabstruct-Net (our)	icdar-2013-partial	0.124 K	S-A	0.928	0.903	0.915
Tabstruct-Net (our)	SCITSR	12.124K	S-A	0.930	0.908	0.919
	+ icdar-2013-partial					
Tablenet [12]	Marmot extended	1.016K	S-B	0.931	0.900	0.915
Graphtsr [14]	Scitsr	12.124K	S-B	0.854	0.891	0.872
	+ icdar-2013-partial					
dgcnn [9]	Scitsr	12.124K	S-B	0.986	0.990	0.988
	+ ICDAR-2013-partial					
Tabstruct-Net (our)	ICDAR-2013-partial	0.124 K	S-B	0.991	0.989	0.990
Tabstruct-Net (our)	SCITSR	12.124K	S-B	0.991	0.993	0.992
	+ icdar-2013-partial					

Table 1. Comparison of results for physical structure recognition on ICDAR-2013partial dataset. P: indicates precision, R: indicates recall, F1: indicates F1 Score and #Images: indicates number of table images in the training set.

Method	Training		Exp.	$\mathbf{P}\uparrow$	$\mathbf{R}\uparrow$	$\mathbf{F1}\uparrow$
	Dataset	#Images	Setup			
NLPR-PAL [19]	CTDAR	0.6K	S-A	0.720	0.770	0.745
dgcnn [9]	CTDAR	0.6K	S-A	0.785	0.751	0.768
dgcnn [9]	SCITSR	12.0K	S-A	0.552	0.519	0.535
dgcnn [9]	CTDAR + SCITSR	12.6K	S-A	0.803	0.778	0.790
splerge [10]	CTDAR	0.6 K	S-A	0.774	0.783	0.778
splerge [10]	SCITSR	12.0 K	S-A	0.559	0.572	0.565
splerge [10]	CTDAR + SCITSR	12.6K	S-A	0.792	0.800	0.796
TabStruct-Net (our)	CTDAR	0.6K	S-A	0.803	0.768	0.785
Tabstruct-Net (our)	SCITSR	12.0K	S-A	0.595	0.572	0.583
Tabstruct-Net (our)	CTDAR + SCITSR	12.6K	S-A	0.822	0.787	0.804

Table 2. Comparison of results for physical structure recognition on ICDAR-2019 (CTDAR) archival dataset. For comparison against DGCNN[9], we use the cell bounding boxes detected from TabStruct-Net for a fair comparison. P: indicates precision, R: indicates recall, F1: indicates F1 Score and #Images: indicates number of table images in the training set.

Method	Exp. Setup	$\mathbf{P}\uparrow$	$\mathbf{R}\uparrow$	$F1\uparrow$
DeepDeSRT [7]	S-A	0.554	0.529	0.541
Splerge [10]	S-A	0.795	0.776	0.785
TabStruct-Net (our)	S-A	0.849	0.828	0.839
Graphtsr [14]	S-B	0.763	0.786	0.774
dgcnn[9]	S-B	0.921	0.898	0.909
TabStruct-Net (our)	S-B	0.992	0.994	0.993

Table 3. Comparison of results for physical structure recognition on UNLV-partial dataset. P: indicates precision, R: indicates recall, F1: indicates F1 Score. All models are trained on SciTSR and fine-tuned on UNLV-partial datasets.

Method	Exp.	Evaluation on						
	Setup	SciTS	$\mathbf{R}$		SciTSR-COMP			
		$\mathbf{P}\uparrow$	$\mathbf{R}\uparrow$	$\mathbf{F1}\uparrow$	$\mathbf{P}\uparrow$	$\mathbf{R}\uparrow$	$\mathbf{F1}\uparrow$	
DeepDeSrt [7]	S-A	0.906	0.887	0.890	0.863	0.831	0.846	
splerge [10]	S-A	0.922	0.915	0.918	0.911	0.880	0.895	
Tabstruct-Net (our)	S-A	0.927	0.913	0.920	0.909	0.882	0.895	
Graphtsr [14]	S-B	0.959	0.948	0.953	0.964	0.945	0.955	
dgcnn [9]	S-B	0.970	0.981	0.976	0.963	0.974	0.969	
Tabstruct-Net (our)	S-B	0.989	0.993	0.991	0.981	0.987	0.984	

Table 4. Comparison of results for physical structure recognition on SciTSR and SciTSR-COMP datasets. P: indicates precision, R: indicates recall, F1: indicates F1 Score. All the models are trained on SciTSR dataset.

## 4.2 Average Results on Markup Output

Tables 5-6 present compare our results for logical structure prediction from the table image on TableBank and PubTabNet dataset, respectively. The scores are obtained by averaging the score for every table across all the tables in the evaluation dataset. From the tables, it can be inferred that despite trained with a much smaller set of data, our model achieves better performance than [13]. Direct comparison, however would not be fair because of the use of different input modalities for training.

Method	Training			Experimental	BLEU	$\mathbf{BLEU}\uparrow$			
	Dataset	Type	#Images	Setup	Word	Latex	Both		
Image-to-Text [11]					0.751	0.673	0.7138		
Image-to-Text [11]	TableBank	Latex	87.597K	S-A	0.405	0.765	0.582		
Image-to-Text [11]	TableBank	Both	144.493K	S-A	0.712	0.765	0.738		
Tabstruct-Net (our)	SCITSR	Image	12K	S-A	0.914	0.937	0.916		

Table 5. Comparison of results for logical structure recognition on TableBank dataset.

Method	Experimental Setup	Training Dataset	#Images	$\mathbf{TEDS}\uparrow$
Acrobat Pro [13]	S-A	-	-	0.537
WYGIWYS [13]	S-A	PubTabNet	399K	0.786
EDD [13]	S-A	PubTabNet	399K	0.883
Tabstruct-Net (our)	S-A	Scitsr [14]	12K	0.901

**Table 6.** Comparison of results for logical structure recognition on PubTabNet dataset [13]. **TEDS:** indicates averaged tree edit distance based similarity [13].

We present our results on the output XML file that contains — (a) bounding box coordinates, (b) start and end row indices, (c) start and end column indices, and (d) content for every predicted cell given the table image. To evaluate our method, we compare this XML against the ground-truth prepared in the same format using BLEU, CIDRr and ROUGE scores as presented in Table 7. The table also compares our results against DGCNN [?] when cells detected from TabStruct-Net are provided as the input to their model.

Training Set	Evaluation Set	Model	$\mathbf{BLEU}\uparrow$	CIDEr↑	ROUGE↑
SCITSR	Scitsr	DGCNN	0.774	0.8	0.782
		Tabstruct-net	0.833	0.848	0.839
Scitsr	SciTSR-COMP	DGCNN	0.769	0.795	0.774
		Tabstruct-net	0.826	0.837	0.830
Scitsr + UNLV-partial	UNLV-partial	DGCNN	0.721	0.744	0.729
		TabStruct-net	0.804	0.826	0.813
Scitsr	ICDAR-2013	DGCNN	0.756	0.773	0.762
		Tabstruct-net	0.815	0.831	0.821
scitsr + ICDAR-2013-partial	ICDAR-2013-	DGCNN	0.772	0.801	0.78
	partial	Tabstruct-net	0.829	0.845	0.834

 Table 7. Results comparison of various methods for table structure recognition on various datasets.

## 4.3 Macro-Averaged Results

Tables 8-12 show the macro-averaged results of various methods for structure recognition on multiple datasets. From the tables, it can be observed that our method outperforms previously published works under multiple kinds of experimental settings. Further, it is important to note that the tables use an IoU threshold of 0.5 to identify true positive cells for experiment setup S-A. It can be inferred from the tables that the macro-averaged numbers follow the same trend as the micro-averaged results.

Method	Training		Experimental	$\mathbf{P}\uparrow$	$\mathbf{R}\uparrow$	$F1\uparrow$
	Dataset	#Images	Setup			
DeepDeSRT [7]	Scitsr	12K	S-A	0.603	0.591	0.597
splerge [10]	Private [10]	83K	S-A	0.914	0.897	0.905
Tabstruct-Net (our)	SCITSR	12K	S-A	0.883	0.871	0.877
TableNet [12]	Marmot Extended	1K	S-B	-	-	-
Graphtsr [14]	Scitsr	12K	S-B	0.819	0.855	0.837
Splerge [10]	Private [10]	83K	S-B	0.932	0.917	0.924
dgcnn [9]	Scitsr	12K	S-B	0.959	0.971	0.965
Tabstruct-Net (our)	SCITSR	12K	S-B	0.961	0.973	0.967

Table 8. Comparison of results for physical structure recognition on ICDAR-2013 dataset. P: indicates precision, R: indicates recall, F1: indicates F1 Score and #Images: indicates number of table images in the training set. SPLERGE [10] is the best performing model (with post-processing) on ICDAR-2013 dataset in S-A. TabStructNet is the best performing model on ICDAR-2013 dataset in S-B.

Method	Training		Exp.	$\mathbf{P}\uparrow$	$\mathbf{R}\uparrow$	$F1\uparrow$
	Dataset	#Images	Setup			
DeepDeSRT [7]	ICDAR-2013-partial	0.124K	S-A	0.853	0.795	0.823
Bi-directional GRU [15]	ICDAR-2013-partial	$0.124 \mathrm{K}$	S-A	-	-	-
Tabstruct-Net (our)	ICDAR-2013-partial	$0.124 \mathrm{K}$	S-A	0.907	0.889	0.898
Tabstruct-Net (our)	Scitsr	12.124K	S-A	0.922	0.896	0.909
	+ icdar-2013-partial					
Tablenet [12]	Marmot extended	1.016K	S-B	-	-	-
Graphtsr [14]	SCITSR	12.124K	S-B	0.846	0.879	0.862
	+ icdar-2013-partial					
DGCNN [9]	SCITSR	12.124K	S-B	0.978	0.984	0.981
	+ icdar-2013-partial					
Tabstruct-Net (our)	ICDAR-2013-partial	$0.124 \mathrm{K}$	S-B	0.985	0.986	0.985
Tabstruct-Net (our)	SCITSR	12.124K	S-B	0.985	0.989	0.987
	+ ICDAR-2013-partial					

Table 9. Comparison of results for physical structure recognition on ICDAR-2013partial dataset. P: indicates precision, R: indicates recall, F1: indicates F1 Score and #Images: indicates number of table images in the training set. TabStruct-Net is the best performing model on ICDAR-2013-partial dataset in both S-A and S-B.

Method	Training		Exp.	$\mathbf{P}\uparrow$	$\mathbf{R}\uparrow$	$F1\uparrow$
	Dataset	#Images	Setup			
NLPR-PAL [19]	CTDAR	0.6 K	S-A	-	-	-
dgcnn [9]	CTDAR	0.6K	S-A	0.704	0.649	0.675
dgcnn [9]	SCITSR	12.0K	S-A	0.427	0.395	0.410
dgcnn [9]	CTDAR + SCITSR	12.6K	S-A	0.728	0.672	0.699
Tabstruct-Net (our)	CTDAR	0.6K	S-A	0.729	0.667	0.697
Tabstruct-Net (our)	SCITSR	12.0 K	S-A	0.483	0.458	0.470
Tabstruct-Net (our)	CTDAR + SCITSR	12.6K	S-A	0.754	0.691	0.721

Table 10. Comparison of results for physical structure recognition on ICDAR-2019 (CTDAR) archival dataset. For comparison against DGCNN[9], we use the cell bounding boxes detected from TabStruct-Net for a fair comparison. P: indicates precision, R: indicates recall, F1: indicates F1 Score and #Images: indicates number of table images in the training set. TabStruct-Net is the best performing model on ICDAR-2019 CTDAR archival dataset in both S-A and S-B.

Method	Exp. Setup	$\mathbf{P}\uparrow$	$\mathbf{R}\uparrow$	<b>F1</b> ↑
DeepDeSrt [7]	S-A	0.469	0.426	0.446
splerge [10]	S-A	0.748		
Tabstruct-Net (our)	S-A	0.803	0.775	0.788
Graphtsr [14]	S-B	0.727	0.744	0.735
DGCNN <sup>*</sup> [9]	S-B	0.887	0.858	0.872
Tabstruct-Net (our)	S-B	0.975	0.981	0.978

Table 11. Comparison of results for physical structure recognition on UNLV-partial dataset. P: indicates precision, R: indicates recall, F1: indicates F1 Score. All models are trained on SciTSR and fine-tuned on UNLV-partial datasets. TabStruct-Net is the best performing model on UNLV-partial dataset in both S-A and S-B.

Method	Exp.									
	$\mathbf{Setup}$	SciTS	$\mathbf{SR}$		SciTS	R-CO	MP			
		$\mathbf{P}\uparrow$	$\mathbf{R}\uparrow$	$\mathbf{F1}\uparrow$	$\mathbf{P}\uparrow$	$\mathbf{R}\uparrow$	$F1\uparrow$			
DeepDeSrt [7]	S-A	0.872	0.854	0.863	0.806	0.779	0.792			
splerge [10]	S-A	0.907	0.892	0.899	0.884	0.857	0.870			
Tabstruct-Net (our)	S-A	0.910	0.897	0.903	0.889	0.864	0.876			
Graphtsr [14]	S-B	0.936	0.931	0.934	0.943	0.925	0.934			
dgcnn <sup>*</sup> [9]	S-B	0.956	0.965	0.960	0.947	0.955	0.951			
Tabstruct-Net (our)	S-B	0.963	0.977	0.970	0.955	0.969	0.962			

Table 12. Comparison of results for physical structure recognition on SciTSR and SciTSR-COMP datasets. P: indicates precision, R: indicates recall, F1: indicates F1 Score. All the models are trained on SciTSR dataset. TabStruct-Net is the best performing model on SciTSR and SciTSR-COMP datasets in both S-A and S-B.

## 4.4 Qualitative Results of Cell Detection

Figures 2-3 demonstrate some qualitative results of cell detection on all the evaluation datasets. From the figures, it can be seen that our model is able to work in the presence of archival table images, multiple row/column spanning cells, varied table layouts and multiple line spanning cells. This indicates the robustness of our method under multiple kind of table images.



Fig. 2. Sample intermediate cell detection results of TabStruct-Net on table images of ICDAR-2013 (in First Row), ICDAR-2019 (in Second Row), SCITSR (in Third Row), SCITSR-COMP (in Fourth Row) and TableBank (in Fifth Row) datasets.



Fig. 3. Sample intermediate cell detection results of TabStruct-Net on table images of PubTabNet (in First Row) and UNLV (in Second Row) datasets.

## 4.5 Qualitative Results of Structure Recognition

Figures 4-10 demonstrate some qualitative results of structure recognition on all the evaluation datasets. From the figures, it can be seen that our model is able to work in the presence of archival table images, multi row/column spanning cells, varied table layouts and multiple line spanning cells. This indicates the robustness of our method under multiple kind of table images.

	number of			franchisees		Eamala	atudanta		Total Costs	Lass: Fanlusions &	Indirect	Total Direct	Federal	NumFed
	1993	1994	1993	1994		Female	students	Cost Category	All Funds	Urailovables	Costs	Costs		Program
Austria	80	170	2500	2700				Salaries (#)	1,314,000		373,250	\$40,753	141,000	793
Belgium/Luxembourg	90	135	3200	2495	Faculty cluster	Sample	Population	Fringe Benefits (1) Consultert Services	357,500		99,960 14,000	252,012 12,000	37,772	214,
Denmark	42	42	500	500	i douity oldotoi	oumpic	ropulation	Stat Travel	94,000		20,000	74,000	11,100	
Finland		12	500	000	0.1	00 110 510		Dad Debts	10,000	10,000 (1)				
France	500	500	30000	30000	Sciences	63 (18.5%)	597 (16.4%)	Office Rent Consumable Supplies	170,000		170,000	150,000	22,500	12
	370	420	15500	18000	0001000	00 (10.070)	001 [10:130]	Subcontracta	175.000	107.000 (2)	1000	-	10,200	
Germany					Social Sciences	100 /EE C0/ \	007E (E7.0v/)	Purchase, Eccipment Losse	62,000	22,100 (2)				
Greece					Social Sciences	189 (55.6%)	2075 (57.0%)	Telephone	109,400	1,000 (1)	55,000	54,400	8,200	*
Ireland	20						<u> </u>	Entertainment Printing & Reproduction	1,800	1,000 (1)	11,000	37.000	5,500	31
Italy	318	361	16100	17500	Humanities	77 (22.6%)	755 (20.7%)	Insurance and Bonding	42,000		42,000			
Netherlands	331	340	12640	12120	i iui iiui iiuico	11 (22.0/0)	100 (20.170)	Fundmining	129,000			120,000		120
Portugal	55	70			01110.1	44.10.00(1)	010 /5 000	Postage and Delivery Decreciation	34,000		5,190	28,900	4,500	8 17
Spain	117	250	14500	20000	Civil Sciences	11 (3.2%)	213 (5.9%)	Allowances	148,000	148,000 (2)		20,000	2,000	
Sweden	200	200	900	900	on outproop	11 (0.2.10)	E10 (0.0.0)	Energency Assistance	54,000	54,000 (2)				
					<b>T</b> . 1	0.40	0010	Training Materials	603, 53	35,000 (2)		82,000	12,300	
116														
UK EU Total	373 2496 number of	396 2884 franchisors	18100 113940 number of	24900 129115	Total	340	3640	Participant Support Coals Total Coals	603,880,6	378,900		1.839.062	257,672	
	2496	2884	113940	129115		340 Female :				378,900	871.038	1.839.062 Total Direct Costs	257,672 Federal Program	Nce-F
EU Total	2496 number of	2884 franchisors	113940 number ol	129115 franchisees		•.•	students	Cost Cottagory	3.088.000 Total Costs All Funds 1.314.000	278,900 Loss: Exclusions &	Indirect Costs 373,250	Total Direct Costs 948,250	Federal Program 141,000	Nor-F Progra
EU Total	2496 number of 1993	2884 franchisors 1994	113940 number ol 1993 2500	129115 franchisees 1994 2700		Female :	students	Total Cools	3.088.000 Total Costs All Funds 1.314.000 352.000	278,900 Loss: Exclusions &	Indirect Costs 373,250 93,930	Total Direct Costs 948,750 252,012	Federal Program 141,000 27,772	Nor-F Progra
EU Total Austria Belgium/Luxembourg	2496 number of 1993 80 90	2884 franchisors 1994 170 135	113940 number of 1993 2500 3200	129115 franchisees 1994 2700 2495	Faculty cluster	•.•		Cost Cottagory	3.088.000 Total Costs All Funds 1.314.000	278,900 Loss: Exclusions &	Indirect Costs 373,250	Total Direct Costs 948,250	Federal Program 141,000	Non-F Prease 21
EU Total Austria Belgium/Luxembourg Denmark	2496 number of 1993 80 90 42	2884 franchisors 1994 170	113940 number ol 1993 2500 3200 500	129115 franchisees 1994 2700 2495 500	Faculty cluster	Female : Sample	students Population	Total Costs Cest Critegory Balaries (r) Friage Benefits (r) Cessillarit Services Bad Detes	3.088.000 Total Costs Al: Fands 1.314.000 302.000 25.000 94.000 10.000	278,900 Loss: Exclusions &	Indirect Ceess 373,250 93,888 14,000 23,000	Total Direct Costs 948,750 258,012 12,000 74,000	Federal Program 141,000 27,772 1,000 11,100	Non-F Progra 79 21 1 8
EU Total Austria Belgium/Luxembourg Denmark Pinland	2496 number of 1993 80 90 42	2884 franchisors 1994 170 135 42	113940 number ol 1993 2500 3200 500	129115 franchisees 1994 2700 2495 500 	Faculty cluster	Female : Sample	students Population	Total Coals	3.088.000 Total Costs All Fands 1.314.000 26.000 26.000 26.000 10.000 117.000	278,900 Lossi Exclusions & Unallowables	Indirect Costs 373,250 93,800 14,000 230,000	Total Direct Costs 948,750 259,812 12,000 74,000 159,900	Federal Program 141,000 27,772 1,800 11,000 22,500	Non-F Progra 21 1 6
EU Total Austria Belgium/Luxembourg Denmark Finland France	2496 number of 1993 80 90 42  500	2884 franchisors 1994 170 135 42 500	113940 number ol 1993 2500 3200 500  30000	129115 franchisees 1994 2700 2495 500  30000		Female :	students	Total Costs Cest Category Balance (I) Prices Rentlis (L) Consulter Services Stati Touvil Bad Debit Office Rent Consulter Services	3.088.000 Total Costs All Fands 1.314.000 26.000 26.000 04.000 10.000 16.000	273,900 Loss: Exclusions & Unallowables 10,000 (1)	Indirect Ceets 573,250 93,880 14,000 230,000 170,000 11,000	Total Direct Costs 948,750 258,012 12,000 74,000	Federal Program 141,000 27,772 1,000 11,100	Non-Fr Program 761 211
EU Total Austria Belgium/Luxembourg Denmark Finland Finnce Germany	2496 number of 1993 80 90 42	2884 franchisors 1994 170 135 42 500 420	113940 number ol 1993 2500 3200 500	129115 franchisees 1994 2700 2495 500 	Faculty cluster Sciences	Female : Sample 63 (18.5%)	students Population 597 (16.4%)	Total Costs	3.088.000 Total Costs Al: Funds 1.314.000 255.000 94.000 10.000 105.000 105.000 105.000 105.000 105.000 105.000	278,900 Lossi Exclusions & Unallowables	Indirect Costs 333,250 94,888 14,000 33,000 173,000 173,000 173,000 173,000	Total Direct Costs 940,750 250,012 12,000 74,000 150,000 68,000	Federal Program 141,000 27,772 1,800 11,100 22,500 11,200	Non-P Program 79 21 1 1 9 12 5
Austria Austria Belgium/Luxembourg Denmark Finand France Germany Germany	2496 number of 1993 80 90 42  500 370 	2884 franchisors 1994 170 135 42 500	113940 number ol 1993 2500 3200 500  30000	129115 franchisees 1994 2700 2495 500  30000	Faculty cluster Sciences	Female : Sample 63 (18.5%)	students Population 597 (16.4%)	Total Costs Total Costs Cost Cotegory Balaries (c) Prings Benefits (b) Costaluted Envices Balar Dotts Cotes Benefits Cotes Benefits Cotes Benefits Cotes Benefits Duccotes Bulcoctendes Bul	3,088,000 Total Coats All Farefs 1,314,000 305,000 94,000 10,000 105,000 105,000 105,000 105,000 105,000 105,400	278,900 Less: Exclusions & Unationables 10,000 (1) 107,000 (2) 22,100 (2)	Indirect Cests 5/3.250 99.880 14,000 230.000 170,000 11,000	Total Direct Costs 948,750 259,812 12,000 74,000 159,900	Federal Program 141,000 27,772 1,800 11,000 22,500	Non-P Program 79 21 1 1 9 12 5
EU Total Austria Belgium/Luxembourg Dennark Finland France Germany Greece Ficland	2496 number of 1993 80 90 42 500 370 20	2884 franchisors 1994 170 135 42 	113940 number ol 1993 2500 3200 500  30000 15500 	129115 franchisees 1994 2700 2495 500  30000 18000  	Faculty cluster	Female : Sample 63 (18.5%) 189 (55.6%)	students Population 597 (16.4%) 2075 (57.0%)	Total Doos Cold Category Marken (J) Frings Revefits (C) Consulter Services Starl Travia Bar Doost Official Revefits (S) Distances Distances Distances Total Revefits (S) Particular Sectors Total Revefits (S) Revefits (S) Revef	3.088.000 Total Costs All Funds 1.314.000 357.000 94.000 15.000 175.0000 175.000 175.0000 175.0000 175.0000 175.0000 175.0000 175.0000 175.0000 175.0000 175.0000 175.0000 175.0000 175.0000 175.0000 175.00000 175.00000 175.00000 175.00000 175.00000000 175.000000000000000000000000000000000000	273,900	Indirect Ceets 373,250 94,388 94,000 14,000 170,000 170,000 11,000 10,0000 10,0000 10,0000 10,0000 10,0000 10,00000 10,0000 10,0000 10,0000 10	Total Direct Costs 946.250 252,012 12,000 74,000 156,000 54,400	Federal Program 141,000 127,702 1800 11,000 22,500 13,200 8,200	Non-F Program 21 1 1 6 12 5 4
Austria Austria Belgium/Luxembourg Denmark Finland France Sermany Greece reland Talay	2496 number of 1993 80 90 42  500 370  20 318	2884 franchisors 1994 170 135 42  500 420  361	113940 number of 1993 2500 3200 500  30000 15500  16100	129115 franchisees 1994 2700 2495 500  30000 18000  17500	Faculty cluster Sciences Social Sciences	Female : Sample 63 (18.5%) 189 (55.6%)	students Population 597 (16.4%) 2075 (57.0%)	Total Costs Total Costs Cost Cotegory Balaries (c) Prings Benefits (b) Costaluted Envices Balar Dotts Cotes Benefits Cotes Benefits Cotes Benefits Cotes Benefits Duccotes Bulcoctendes Bul	3,088,000 Total Coats All Farefs 1,314,000 305,000 94,000 10,000 105,000 105,000 105,000 105,000 105,000 105,400	278,900 Less: Exclusions & Unationables 10,000 (1) 107,000 (2) 22,100 (2)	Indirect Costs 333,250 94,888 14,000 33,000 173,000 173,000 173,000 173,000	Total Direct Costs 940,750 250,012 12,000 74,000 150,000 68,000	Federal Program 141,000 27,772 1,800 11,100 22,500 11,200	Non-F Program 21 1 1 6 12 5 4
Austria Austria SelgianvLuxemboarg Jennark Finland France Jernary Greece reland taly	2496 number of 1993 80 90 42 500 370 20	2884 franchisors 1994 170 135 42 	113940 number ol 1993 2500 3200 500  30000 15500 	129115 franchisees 1994 2700 2495 500  30000 18000  	Faculty cluster Sciences	Female : Sample 63 (18.5%)	students Population 597 (16.4%)	Total Dool Copy Detailor Proga Benefic (1) Densulter Bervices Part Dool Part Dool	3.008.000 Total Costs All Funds 1.314.000 265.000 16.000 175.000 657.000 105.0000 105.0000 105.0000 105.0000	278,900 Less: Exclusions & Unationables 10,000 (1) 107,000 (2) 22,100 (2)	Indirect Costs 573,750 94,888 14,000 33,000 1770,000 11,000 45,000 41,000	Total Direct 2045 946,750 202,000 74,000 156,000 68,000 54,400 37,000 120,000	Pederal Pogram 141,000 27,772 18,000 11,000 22,500 11,000 8,000 5,500	Non-F Program 21 1 0 12 5 4 4 3 3 12
U Total Austria Belgium/Luxembourg Demmrk inland irance Bernary Bernary reland taly tethertands	2496 number of 1993 80 90 42  500 370  20 318	2884 franchisors 1994 170 135 42  500 420  361	113940 number of 1993 2500 3200 500  30000 15500  16100	129115 franchisees 1994 2700 2495 500  30000 18000  17500	Faculty cluster Sciences Social Sciences Humanities	Female : Sample 63 (18.5%) 189 (55.6%) 77 (22.6%)	students Population 597 (16.4%) 2075 (57.0%) 755 (20.7%)	Total Doos Cest Celepcy Basterin (1) Programmer (1) Programmer (1) Bas Dools Dots Rest Dots	3.008.000 Total Costs Al Fands 1.314.000 263.000 16.000 175.000 163.000 175.000 163.000 175.000 163.000 175.000 163.000 164.000 163.0000 163.000 163.000 163.000 163.0000 163.0000 163.0000 163.0000 163.0000 163.0000 163.0000 163.0000 163.0000 163.0000 163.0000 163.0000 163.0000 163.0000 163.00000 163.0000 163.0000 163.0000 163.0000 163.0000 163.00000 163.00000 163.00000 163.00000 163.000000 163.000000 163.000000 163.0000000000 163.000000000000000000000000000000000000	278,900 Less: Exclusions & Unationables 10,000 (1) 107,000 (2) 22,100 (2)	Indirect Costs 523,250 94,300 14,000 173,000 110,000 55,000 11,000 42,000 41,000	Total Direct Coats 946,750 252,870 125,800 155,800 66,400 54,400 237,600 120,000 28,500	Federal Program 141,000 22,772 1,800 11,300 22,500 13,200 8,200 5,500 4,300	Nov F Press 79 211 1 1 8 12 5 5 4 4 3 3 12 2 2 2
U Total Austria belgjum/Luxembourg bemark inland irance Germany Steece reland taly steherlands vetherlands	2496 number of 1993 80 90 42  500 370  20 318 331	2884 franchisors 1994 170 135 42  500 420  361 340	113940 number of 1993 2500 3200 500  30000 15500  16100 12640	129115 Tranchisees 1994 2700 2495 500  30000 18000   17500 12120	Faculty cluster Sciences Social Sciences Humanities	Female : Sample 63 (18.5%) 189 (55.6%) 77 (22.6%)	students Population 597 (16.4%) 2075 (57.0%) 755 (20.7%)	Total Dool Copy Detailor Proga Benefic (1) Densulter Bervices Part Dool Part Dool	3.008.000 Total Costs All Funds 1.314.000 265.000 16.000 175.000 657.000 657.000 100.400 14.000 142.000 142.000 120.000	278,900 Less: Exclusions & Unationables 10,000 (1) 107,000 (2) 22,100 (2)	Indirect Cests 273,250 93,380 14,000 11,000 11,000 55,000 11,000 42,000	Total Direct 2045 946,750 202,000 74,000 156,000 68,000 54,400 37,000 120,000	Pederal Pogram 141,000 27,772 18,000 11,000 22,500 11,000 8,000 5,500	Nov F Press 79 211 1 1 8 12 5 5 4 4 3 3 12 2 2 2
U Total Austria Austri	2496 number of 1993 80 90 42  500 370  20 318 331 55	2884 franchisors 1994 170 135 42  500 420  361 340 70	113940 number of 1993 2500 3200 30000 15500  16100 12640 	129115           franchisees           1994           2700           2495           500	Faculty cluster Sciences Social Sciences	Female : Sample 63 (18.5%) 189 (55.6%)	students Population 597 (16.4%) 2075 (57.0%)	Testi Guesa Cett Cetapory Riskete (1) Proga Randha (2) Riskete (2)	3.088.000 4// Funds 1.314.000 1.314.000 10.000 10.000 107.000 05.000 107.0000 107.0000 107.0000 107.0000 107.0000 10	278,000 Lossi Enclusions & Unationables 10,000 (1) 10,000 (2) 1,000 (2) 1,000 (2)	Indirect Ceets 533,259 94,380 14,000 130,000 11,000 55,000 11,000 42,000 41,000 43,000	Total Direct Ox85 94:0750 2502,602 150,000 150,000 54:400 37,000 120,000 28,000 28,000	Pederal Program 141,000 11,000 11,000 11,000 122,500 11,000 13,000 4,000 1,000	Nce-F Praga 20 12 12 5 5 4 4 3 3 12 2 2 1
EU Total Austria Belgium/Luxembourg Dennark Finland France Germany Greece Ficland	2496 number of 1993 80 90 42 500 370 20 318 331 55 117	2884 franchisors 1994 170 135 42  500 420  361 340 70 250	113940 number of 1993 2500 3200 500 16100 12640 14500	129115 Tranchisees 1994 2700 2495 500  30000 18000   17500 12120  20000	Faculty cluster Sciences Social Sciences Humanities	Female : Sample 63 (18.5%) 189 (55.6%) 77 (22.6%)	students Population 597 (16.4%) 2075 (57.0%) 755 (20.7%)	Total Costs Total Cost Sector 201 Statistics Sector 201 Consult of Diversity Consult of Diversity Derivation D	3.008.000 Total Costs 1.314.000 023.000 04.000 177,000 170,0	278,000 Lossi Exclusions & Unationables 18,000 (1) 10,7,000 (2) 1,800 (1) 1,800 (1)	Indirect Costs 323,256 14,000 1750,000 11,000 55,000 41,0000 41,0000 41,000 41,0000 41	Total Direct Coats 946,750 252,870 125,800 155,800 66,400 54,400 237,600 120,000 28,500	Federal Program 141,000 22,772 1,800 11,300 22,500 13,200 8,200 5,500 4,300	Non-Program 769 214 11 65

Fig. 4. Sample structure recognition output of TabStruct-Net on table images of ICDAR-2013 dataset. First Row: prediction of cells which belong to the same row. Second Row: prediction of cells which belong to the same column. Cells marked with orange colour represent the examine cells and cells marked with green colour represent those which belong to the same row/column of the examined cell.



Fig. 5. Sample structure recognition output of TabStruct-Net on table images of ICDAR-2019 dataset. First Row: prediction of cells which belong to the same row. Second Row: prediction of cells which belong to the same column. Cells marked with orange colour represent the examine cells and cells marked with green colour represent those which belong to the same row/column of the examined cell.



Fig. 6. Sample structure recognition output of TabStruct-Net on table images of SciTSR dataset. First Row: prediction of cells which belong to the same row. Second Row: prediction of cells which belong to the same column. Cells marked with orange colour represent the examine cells and cells marked with green colour represent those which belong to the same row/column of the examined cell.

	Nu	mber of itera	tions			Moh	vdick	Tos	W)	En	likh	Wik	ipedia		Dataset	Support	I	C	DTLB L	1M/L2H		1M/1.2M
Instance	Original	UPLA	UPLA+ER		1	men	DIOC	men	DICC	mem	proc	mem	Droc		accidents	0.25	Cik 0.59589	Cluster 0.603259	Cilk 0.00048	Cluster 0.000046	Cik 0.00061	Cluster
aim - 50 - no - 1.6	39,223	4,201	4.958		0	_		a. (c)	· .	_					chass	0.6	0.560538	0.68955	0.000797	0.0(0242	0.001016	0.000032
aim - 50 - no - 2.0	67,105	6,240	7,026		U	0.25	0.061	0.40	0.156	2.36	0.886	14.41	7.131		comect	0.8	0.513099	0.803038	0.000249	0.0(0112	0.001204	0.000141
aim-100-no-1.6	433,398	134,572	120,122		1	133	0.320	1.79	0.576	8.55	3.450	55.84	32.287		kosarak	0.0013	0.692103		0.00040	0.)(0185	0.00659	
aim - 100 - no - 2.0	3,266,805	160,971	174,131		2	4.57	1.272	6.91	2.483	30.49	12.596	170,79	107.289		punsb	0.75	0.294539	0.719072		0.0(0114	0.001276	0.000125
aim-200-no-1.6	5,902,875	1,390,580	1,788,612		2	9.78	4.044	15.18	7.458	61.37	36.309	342.18	270.506		punsb.star mushroom	0.3	0.527659	0.636358 0.705003		0.0(0145	0.001082	0.000113
aim - 200 - no - 2.0	9,802,581	3,066,713	3,762,236		9										TICIICDICCK	0.05	0.617272	0.727288		0.00005		0.000021
uu f 50 - 218	174,752,209	43,620,471	5,139,0176		4	16.09	14.647	27.20	28.144	105.75	117.970	003.35	922.521		11014D100K	0.0006	0.555330			0.0(0144		
				느										느	·							
	Nur	nber of iterat	ions			Moł	wike	To	<b>7</b> 1	En	likh	Wik	neis	iF		0	I	20	DTLB 1	1M/L2H	DTLB I	.1M/L2M
Instance		nber of iterat UPLA					ydick	To		1	lish		ipedia		Dataset	Support	Cik	Cluster	Gik	IM/L2H Chster	Cik	Cluster
	Original	UPLA	UPLA+ER		d	mem	p100	mem	proc	menn	proc	m(m	proc		accidents	0.25	Cik 0.525689	Cluster 0.603059	Cilk 0.000048	Chister 0.000346	Cík (((0)161	Cluster 0.000110
aim-50-no-1.6	Original 39,223	UPLA 4,201	UPLA+ER 4.958		d 0					1			-		accidents chess	0.25 0.6	Cik 0.525689 0.560538	Cluster 0.603559 0.663955	Cilk 0.000048 0.000797	Chister 0.000046 0.000242	Cik (000161 (0010)6	Chister 0.000110 0.000002
aim - 50 - no - 1.6 aim - 50 - no - 2.0	Original 39,223 67,105	UPLA 4,201 6,240	UPLA+ER 4.958 7.026		d 0	mem	p100	mem 0.46	ргсс 0.156	mem 2.36	proc	mem 14.41	proc 7.131		accidents dress connect	0.25 0.6 0.8	Cik 0.55589 0.560538 0.513069	Cluster 0.603259 0.663955 0.80308	Cilk 0.000048 0.000797 0.000249	Chster 0.00046 0.00242 0.000112	Cik (0.00161 (0.0010)6 (0.0012)4	Chister 0.000110 0.000032 0.000141
$\begin{array}{l} aim-50-no-1.6\\ aim-50-no-2.0\\ aim-100-no-1.6 \end{array}$	Original 39,223 67,105 433,398	UPLA 4,201 6,240 134,572	UPLA+ER 4.958 7.026 120,122		d 0 1	mem 0.25 1.33	proc 0.061 0.320	mem 0.46 1.79	ргос 0.156 0.576	mem 2.36 8.55	proc 0.886 3.450	mem 14.41 55.84	ргос 7.131 32.287		accidents chess	0.25 0.6	Cik 0.525689 0.560538 0.543069 0.692103	Cluster 0.603559 0.663955	Cilk 0.000048 0.000797 0.000249 0.000440	Chister 0.000046 0.000242	Cik (000161 (0010)6	Cluster 0.000110 0.000002 0.000141 0.000123
aim - 50 - no - 1.6 aim - 50 - no - 2.0	Original 39,223 67,105 433,398 3,266,805	UPLA 4,201 6,240 134,572 160,971	UPLA+ER 4.958 7.026 120,122 174,131		d 0 1 2	mem 0.25 1.33 4.57	proc 0.061 0.320 1.272	mem 0.46 1.79 6.91	proc 0.156 0.576 2.483	mem 2.36 8.55 30.49	proc 0.886 3.450 12.596	mem 14.41 55.84 170.79	риос 7.131 32.287 107.289		accidents dress connect kosarak	0.25 0.6 0.8 0.0013 0.75 0.3	Cik 0.525689 0.560538 0.543069 0.622103 0.694539 0.527659	Cluster 0.603359 0.668965 0.800008 0.717599 0.719072 0.608358	Cilk 0.000048 0.000249 0.000249 0.000400 0.000230 0.000230	Cheter 0.000046 0.000242 0.000112 0.000185 0.000114	Cik (0.00161 (0.0010)6 (0.0012)1 (0.00659	Cluster 0.000110 0.000022 0.000141 0.000123 0.000123 0.000125
$\begin{array}{l} aim-50-no-1.6\\ aim-50-no-2.0\\ aim-100-no-1.6\\ aim-100-no-2.0 \end{array}$	Original 39,223 67,105 433,398	UPLA 4,201 6,240 134,572	UPLA+ER 4.958 7.026 120,122		d 0 1 2 3	mem 0.25 1.33	proc 0.061 0.320	mem 0.46 1.79	ргос 0.156 0.576	nem 2.36 8.55 30.49 61.37	proc 0.886 3.450 12.596 36.309	mem 14.41 55.84 170.79 342.18	ргос 7.131 32.287 107.289 270.505		accidents dress connect losserak punsb punsb_star unskroom	0.25 0.6 0.013 0.75 0.3 0.10	Cik 0.525689 0.560538 0.543049 0.82203 0.494539	Cluster 0.603059 0.668965 0.80008 0.717579 0.719072 0.608358 0.705003	CIL 0.00048 0.00029 0.00029 0.00020 0.00020 0.000315 0.000477	Chater 0.00046 0.00012 0.00112 0.00185 0.00145 0.00145	Cik (0.00161 (0.010)6 (0.012)4 (0.0059 (0.0059 (0.001276 (0.00182 (0.00185)	Chaster 0.000110 0.000022 0.000141 0.000123 0.000125 0.000113 0.000122
$\begin{array}{l} aim-50-no-1.8\\ aim-50-no-2.0\\ aim-100-no-1.6\\ aim-100-no-2.0\\ aim-200-no-1.6\end{array}$	Original 39,223 67,105 433,398 3,266,805 5,902,875	UPLA 4,201 6,240 134,572 160,971 1,390,580	UPLA+ER 4.958 7.026 120,122 174,131 1,788,612		d 0 1 2 3 4	mem 0.25 1.33 4.57	proc 0.061 0.320 1.272 4.044	mem 0.46 1.79 6.91 15.18	proc 0.156 0.576 2.483	mem 2.36 8.55 30.49	proc 0.886 3.450 12.596 36.309	mem 14.41 55.84 170.79 342.18	риос 7.131 32.287 107.289		accidents dross connect kosarak punsb punsb_star	0.25 0.6 0.8 0.0013 0.75 0.3	Cik 0.525689 0.560538 0.543069 0.622103 0.694539 0.527659	Cluster 0.603359 0.668965 0.800008 0.717599 0.719072 0.608358	Cilk 0.00048 0.00797 0.00249 0.00240 0.00230 0.00230 0.000477 0.00048	Chater 0.00046 0.00212 0.00112 0.0015 0.00114 0.00145 0.00267 0.00265	Cik (0.00161 (0.01204 (0.002204 (0.002276 (0.002276 (0.002276 (0.00250 (0.00250	Cluster 0.000110 0.000032 0.000141 0.000123 0.000125 0.000125 0.000125 0.00022

Fig. 7. Sample structure recognition output of Tabstruct-Net on table images of SciTSR-COMP dataset. First Row: prediction of cells which belong to the same row. Second Row: prediction of cells which belong to the same column. Cells marked with orange colour represent the examine cells and cells marked with green colour represent those which belong to the same row/column of the examined cell.

	Olare my		In case of a local division of the	- ( */+ don 1885 com							14	12	Damage	Count		(	1 (hun)	-
Revenue and conserves	an Lanaga Miningat	na Linnapa 2006 nega	cycone .	- 5	Vanasaans		0	huma oora	ALL DAY		NCCIEROB	RCGLO	Purpose	Count	a	40 (KIII)	( (km)	C
Государствонная кронная буласть - волос	14557	5231	\$2,4	u	Категория		0	ценка сост	(UNHEN)		1110	MOCTOR	education	12704	3.05	31.07	8.15	57
a tan wear					800008		_			_		_	and straining	40322	2.85	95.15	5.10	10
блиции факульта збот с технологі склай цистего нація (MD40)	3152,8	27982	1502	42	de las e	X00.	YILOB.	Heygos.	Abapwe.	KOETH'L			shopping	40,055	4.00	30.15	5.19	14
Company (and the company of the company of the sector (party (par	1058,6	791.2	-345,4	-21,1									4004	25508	2.71	38.95	17.40	16
абликари фурральни займ с прочикам алих ок (993-130)	1 098,0	13452	353	347	Междуна	296	67	43	23	8	437	451	territor da	00716	0.51	45.1.9	0.00	50
блицро (крузляте забы с нанклузием линнали ОФГПБ)	U	141.0	Ini.8		Decement	204	122	39	n	10	424	172	CITALOS	23716	2.51	40.13	8.10	9;
1043	161,6	- 10	-18%	-118,0	l bevuap.	404	135	23	24	10	424	4/3	accompanying driver	6447	250	32.30	7.74	47
аударственые сперетехные областия с постоянся продоктовые сперетехные областия с постоянся	5485	34.6	28.9	w	Beem	500	200	Ø	55	24	861	Q04	1 7 0	61152	0.15	30.38	10.77	22
пудренных обсатичных абласция с фасаралений родиний станай относой сондол СО 2003	132.0	122.0	0,0	93	DACID.	Jun:	100	10	au	1.4	001		me time	0.110	3.10	30.38	9.99	40,
						550.				382	93%	100%		2706	1.00	19.15	20.59	80
Славана механа «Славарноних зайка. Российний >-урнун об РОЦ	96,2 Othere may	9.Li	0,0 Euros	9 (SI	7	0070	23%	10%	070	3:0	2071	107	0.SHESS		1.64	12.35	30.38	00
бланцин зансано: «Аланцин-мес зайол. Российский 9-уулуч об РСС); Належитсяние годарууся	Officers may response	an lanaye				100		10%	070	0:0	алл Исследов	Bcero	Purpose	Count	1.02 a	12.55	( (km)	00
densen sense dienseren sins Provinsi Propried (R): Remainserenseren	Office my support	14: 10:00 1	Fares	e (194 des Millione	70 Kareropee	0070		иченка сост	070 ICRHHM	3:0	l.	Bcero			a 3.05	(	14.5	57
Rowsing of DOS Rowsing suggests	Olive my representation a la sempe 205 mp2	art anage at 1 anage 2015 rep	Races United - Ju Quan	e (15) des Millerse N	70 Kareroper accors	2070		и или сост	070 ICRHIBS	3:0	Исследов		Purpose education	Count 12704	a 3.05	ℓ <sub>0</sub> (km) 31.07	ℓ (km) 8.15	57
hypper of ROS Bensmann maarten Rospersona const basis- oon 199 konst	Oliare may my per 2015 mg/ 2475.7	art anage at 1 anage 2015 rep	Races United - Ju Quan	e (15) des Millerse N		3070 Хор.		иченка сост Неудов.		570 Кретич	Исследов	Bcero	Ршроне	Count 12704 40322	a	€ <sub>6</sub> (km)	( (km)	57
heppeyer of FUS; Rammana enancem	Oliver my respective 2015 reprint 2015 reprint 14557	artinari internet internet istance	Barrow Hanga - Ja Quan S2.4	n (15) Ge Millere N	, коротн	зол Хор.	0			oro Kortan	Исследов	Bcero Mocros	Purpose education	Count 12704	a 3.05	ℓ <sub>0</sub> (km) 31.07	ℓ (km) 8.15	57
нуруни (1975) Валектание начатите Подарствие слова булит-лого стоя наст стоя наст облаща (програм) и вогот и настание спорт стор (20-да) облаща (програм) и стор сладет (20-да)	Oliver may request to 1 mappe 2015 mpt <sup>4</sup> 1 4757 2 556,0	ang managan ang mangan ang mang mang mangan ang mangan ang mangan ang mangan ang mangan ang mangan	100000 10000 1000 1000 1000	N COL Con Millione N LL D		30,4	0			5:0 Кретич. 8	Исследов	Bcero	Purpose education shopping work	Count 12704 40322 25508	a 3.05 2.85 2.71	ℓ <sub>6</sub> (km) 31.07 35.15 38.95	ℓ (km) 8.15	57 19 16
нородна (ПР) Колектован на валоти Подарителна конан булот-зака такота Собща барала и било такота пода и полого Собща барала и пода такота пода и полого Собща барала собща с претака и на СОСА, Собща барала собща с претака и на СОСА,	History mapping and 1 manager 2015 mapping 2 47557 2 475577 2 475577 2 475577 2 4755777 2 47557777 2 47557777777777777777777777777777777777	200 100000 1 000 200 200 5 2731 2 776/2 5 523	2000 2000 200 200 200 200 200 200 200 2	e (Pa) Sa Milan Si U U U U U	, акропн Междуна	Хор. . 296	О Удов. 67	Heyace. 43		5:0 Кретич. 8	Исследов ано 437	Всего Мостов 451 472	Purpose education shapping	Count 12704 40322 25508 23716	a 3.05 2.85 9.71	l <sub>0</sub> (km) 31.07 35.15 38.95 45.13	ℓ (km) 8.15	57 19 16
Reported INTS Reconstruction conserves Inspacements conserves conserves conserves	08644 1997 1997 1997 1995 1995 1995 1995 1995	278/3 5231 278/3 5231 278/3 5231 278/3	9000 924 1942 2473	e (Pa) Sa Milan Si U U U U U	, коротн	зол Хор.	0	Heyace.		5 то Критачи. 8 16	Исследов	Bcero Mocros	Purpose education shopping aork errunds	Count 12704 40322 25508	a 3.05 2.85 2.71	ℓ <sub>6</sub> (km) 31.07 35.15 38.95	ℓ (km) 8.15	57
нороди (ПР) Вланные надота Подотника сама булить нак подотника бездат прадотна за подотна за как и конст бездат прадотна за подотна са как и конст бездат прадотна са как и констрана, са констрана, бездат прадотна са констрана, са констрана, бездат прадотна са констрана, са констрана, оконстрана, са констрана, са констрана, оконстрана, са констрана, са констрана, оконстрана, са констрана, са констрана, оконстрана, са констрана, са констрана, са констрана, са констрана, оконстрана, са констрана, са констрана, са констрана, оконстрана, са констрана, са констрана, са констрана, оконстрана, са констрана, са констра	00000000000000000000000000000000000000	197 100000 1 1000000 1 1000000 52731 52731 52735 52735 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 100000 100000 100000 100000 100000 100000 100000 100000 100000 100000 100000 1000000	9000 900 90 90 90 90 90 90 90 90 90 90 9	4 (19) 5	, акропн Междуна	Хор. . 296	О Удов. 67	Heyace. 43		5:0 Кретичь 8 16 24	Исследов ано 437	Всего Мостов 451 472	Purpose education shopping out erruids accomparying driver	Count 12704 40322 25508 23716 6447	a 3.05 2.85 2.71 2.51	\$0 (km) \$1.07 \$5.15 \$8.95 45.13 \$2.30	( (km) 8.15 5.19 17.40 8.06 7.74	57 19 16 59 47
нарадна (1921) Влажная калагта тотост Самана бразона или (тотоска сама сама) Самана бразона сама (тотоска сама) Самана бразона сама (тотоска сама) Самана бразона сама) Самана бразона сама (тотоска сама) (тотоска Самана бразона сама)	010 response response 2015 response 2015 res	ар конско из анадия 388 года 3281 года 3291 года 3201 г	Reserve           Quase           Quase           SEA           DD2           D2           D2           D2	400 40 Million 9 10 10 10 10 10 10	дороги Междуна Государ.	Хор. . 296	О Удов. 67	Heyace. 43	Аварий. 23 32	8 16 24	Исследов ано 437 424	Всего Мостов 451 472	Purpose education shopping aork errunds	Count 12704 40322 25508 23716	a 3.05 2.85 2.71 2.51	l <sub>0</sub> (km) 31.07 35.15 38.95 45.13	( (km) 8.15 5.19 17.40 8.06 7.74 13.55	57 19 16 59 47

Fig. 8. Sample structure recognition output of Tabstruct-Net on table images of TableBank dataset. First Row: prediction of cells which belong to the same row. Second Row: prediction of cells which belong to the same column. Cells marked with orange colour represent the examine cells and cells marked with green colour represent those which belong to the same row/column of the examined cell.



Fig. 9. Sample structure recognition output of TabStruct-Net on table images of PubTabNet dataset. First Row: prediction of cells which belong to the same row. Second Row: prediction of cells which belong to the same column. Cells marked with orange colour represent the examine cells and cells marked with green colour represent those which belong to the same row/column of the examined cell.

(Dollars in millions			Set	Per Share Car	mos Stark			10-6	1.61.67	5 N.m.	30mo \$0	61 (18) 85 224	Cramps was among the sweeted tested Lonco lime had a big huit favor. Both ert a very signt being of threat burn.	- 11 -	2/3/7	PUIDICE	SISTERATION .	
except per share		Gross	(124)	Ecel)		Stock Pr				• 00mp			favor, Both et a very signt being of theat burn. Both fevora had a bio half fevor, Lamon-Ime was among the	- 11	Dectecters	Norschen, S-staped vallers		0206002
and stock prices)	Renerate	Profit	Earnings	Earnings	Bividents	High	Levi	All Sport	L 78 8	4 7 40		N N	sweethert and slightly satiringant. Both ware way slightly firsty. Lemma interference was sevent with a slight artificial supernover tests	- 11	late Table stands at present lavel	subsuch graded to present	cold and motet	Tiops
1993 First quarter	\$ 13,058	\$ 5,162	\$ (399)*	\$ 1,707	\$ .54	\$ 57.13	\$ 45.88	All Sport Like Bolly's 1st Ade				25 N 510 28	Tary slightly fizzy (Change not available).	- 11	Repid retreat of	No recentional retraines		
Second quarter	15,519	5,974	(8.036)	(14.10)	54	54,38	\$7.13	Inned	1.78.7	8		34 N	Band both favors were bein sweetness and salts ingression. Create had a sunderscript size favor, skattashev rom.	- 11	glacters	from Pollien Leaf Loke to Utick Page	Tabon Basis precipitation exceeds evaporation	9900+800 ar (A=5F3) (Bevice
Third guarter	14,743	5.602	(48)	6125	.25	49.75	40.63	Inned	P 70 7	30	45 9	40 B	Lamon-Ane-was anly slightly sweet, asthingent. (Shange not texted.)					to 99904000 22
Fourth quarter	19,396	7,410	382	.62	.25	59.88	42.13	Gotorada	L 50 8	110		e e	Lamon-line had a bit loss that favor than other products. Orange was only slightly sweet, with a favor the baby sopier, and a slight pray role.	1	Luke false drops	Eveningal cuts on stifesters; much dures stiffsters; much dures	Take hain experation extends precipitation	
Total	\$ 62,716	\$26,168	\$ (8,101)	\$ (14.22**	\$1.58			Getorada	P 40 8		25 3	50C 11	Lemon lime had bigger that flever than most (Onergenot basied)		Report habitation of			
								Getorada Light	1 1 25 3	3 10	8 1	6 a	Both favors were sweet, with a slight artificial-sweetener laste. Orange had tangetine-like favor.		Table Bade	Artifacts at Spioner Lake site (Elater 15(1))		T100+140 MP
1992 First quarter	\$ 14,037	\$ 7,155	\$ 2,542	\$ 4.407	\$1.21		\$ 83.13	Bysics Feel	L 56 T	8	50 Z	or 52	Lamon-line had thicker lealing in the mosth than most, with leas that favor than powder, slightly sevent. Orange had mondescript citrut-size favor, seek tote.	×	Lake Takes rises	Sreatoral cuts are Rooded: trees are	Table Basin precipitation	(1-30n) 9869-241 m
Second quarter	16,224	7,863	734	1.29	1.21	98.63	81.63	Bydre Fool	P 58 T	25	50 tž	n Da	Letton-line was slightly sweet, with a slight chally last. [Change not actual.]	α		killed at Kiwa Beach	encode evaporation	9909450 SP (12-504) 9556450 SP
Third quarter	14,702	6,779	(2,7/8)	(4.87)	1.21	100.38	80.00	<b>Bearing Plus</b>	L 40 T	12	87 0	80 22	Reformulated since Cit's test.)		Netter, Kingt Beach			(12-805)
	19,560	7,679	(5.663)	(9.57)	1.21	81.13	48.75	Pewankda	1 67 8	- 78	88 O	96 N	amon-lime had a bit less huit favor than most. Grange had so intercelue characteristics.		pocupation; and	Archeological sites oriented towards present topography	Take hade precipitation	3500 BP to
Fourth quarter													Lamon-line had little salts into resident, rendes priot oinus favor.			Spical projectile points	exceeds evaporation	present
Feerth gearner Teeal	\$ 64,523	\$29,454	\$ 49.965)	\$ (8.70)	\$4.84	_		Supple Sup-1	<b>h</b> 1.10 8	8	4 1	3 3	bolied of peol and canned juice, with a birt of oppierment. Ouropi had orangeade-ike flavor, basied of peol and canned juice.		These at present level			
Total (Dollars in millions	\$ 64,523		\$ (4,965) Net	\$ (8.78) PerShare Co	\$4.84	_		0000			0 7	0 (Q	benef op of and carried pilo, with a hird of pole-mint. During had ourspeed- the favor, tanked of peel and correct pilon.	╡┝	noe a proez lest	WIRK .	Servers (Liverso .	
Total (Dolars in millions except per share		6nsi	doel	1.081		Stock		194		8 7 8 5 5 %	0 ( 31 mg \$2	0 ( 0 8 29	ballet of point and carrent pipe, with a hird of optimized Coungs is ad comparable for flavor, used of point of carrent pipes. U Theore discription Orange were among the sensent tablet. Lorono instruct a top but theore, before it way sight ballet of theore favor.	╡┝		NUBIC	SINDICARE	apostuar
Total (Dollars in millions	\$ 61.523 Iterene				\$4.84 mon Stak Dividents	Stati High	Mos Lav	19-E Al Spot	700/100/ L 60 P	8 7 8 5 5 15 5	8 (	N N N 29	beard of prior and come gains, who and not any parent. O long it ad- compands that not instantial of the advancement of the prior of the second second second second second second second second second second secon		TANK OF DURING THE		Denomin (Liberro) - Sibilitioex mild and emist	CEOKLON Tinga
Total (Dolars in millions except per share		6nsi	doel	1.081				13 Sport	1/00/100/ L 80 P L 30 S	6 7 6 5 55 mg 55 8 4	10 21 mg 52 55 c 51 c	0 (0 8 29 8 8	beit of per of complex for the special Courge's at expended for the counterplane. The special council of the special council of the special council of the special of the special council of the special council of the special feed the special council of the special council of the special for the special council of the special council of the special feed the special council of the special council of the special restinct of the special council of the special council of the special council of the special council of the special restinct of the special council of the special		FileCastion Statistics Statistics Strengthere and Statistics Stati	NUDE: Number, L-stapet villeys streach protect to present iced)	mild and moint	Tinga
Total (Dollars in millions except per share and stock prices)	levenue	Gross Profit	doel	Loei Eamings	Dividents		Low	19-E Al Spot	1 <b>0</b> / <b>0</b> / <b>0</b> / 1 00 P 1 00 P	6 7 6 5 55 mg 55 8 4	8 0 31mg 52 55 0 51 0 25 0	N N N 29	bank J per la sama Lipi, with Lint Fragmenic Drogin Sel angebie history, taket per la classification ( <u>Uters includes</u> ) ( <u>Uters includes</u> ) ( <u>Uters includes</u> ) ( <u>Distributed</u> ) (Distributed) (Distribut			INIDEZ Nerskes, U-shpet villeys miteski gradei to present level	SINDICARE	75 aga 1900-608 29
Teed (Dolars in millions except per share and stock prices) 1993 First quarter Second quarter	Bovenue 8 13,058 15,519	6ros Profit \$ 5,162	4.cel Earnings \$ (399)* (8,036)	Looi Earnings \$ (.30)* (14.10)	Dividents \$54	High \$ 57.13	tow \$ 45.88 47.13	19860 19-6 all Spert All Spert Lite Delive Lite Lite	(0/10) L 60 P L 70 I L 7 I L 8 7	8 4 5 5 5 5 5	8 0 8 0 8 0 8 0 8 0 8 0	0 (0 8 29 8 8 8 8 8 8 8 8	beint of the disc with the disc week of the disc of th		Toto: Statistics Statistics Notification Rectification Statistics Statis	NTERE Brains, Schopel willings schesh proble to present ierd) W recentioni, arrites from PAINs (and lake to Sido Rus	mid and mint	75aga 2900-608 27 18-7351 (Ser
Teed (Dollars in millions except per share and steck prices) 1993 First quarter	berrue \$ 13,058	Gross Pedit \$ 5,162 5,974	Loei Earnings \$ (395%)*	Loon Earrings \$ (,30)*	Dividents \$54	High \$ 57.13 54.38	low \$ 45.88	Parties 19-6 All Speet All Speet the bally List Ma Esseet Caseed Calcula	10707 L 60 F L 70 I L 8 F L 8 7 L 70 7 P 70 7 L 92 5	6 7 6 5 35 mg 5 4 5 4 55 55 55 55 55 55 55 55 55 55 55 55 55	8 0 8 0 8 0 8 0 8 1 8 1 8 1	0 (0 8 28 8 8 8 8 8 8 8 8 8 8 8 8 8 8	beind priper of omenuity with of or greater Deep late support and the support of the support (result has the support of the support of the support of the support of the support of the support of the support of the support of the support of the support of the su	a a 60 v	FileCastion Statistics Statistics Strengthere and Statistics Stati	FUENCE Notices, Lisboot valleys minute grade to present less Nr notestaal northese from Follow Later Jake to Citle Fues Provinged Long of FUENDER and Alexe	mid and mint	Tinga associate on
Teed (Delars in millions except par share and steck prices) 1993 First quarter Second quarter Third quarter	Revenue \$ 13,058 15,519 14,743	6ross Profit \$ 5,162 5,974 5,602	4.001 Earnings \$ (3999)* (8,036) (48) 382	LON Emirgs \$ (.30)* (14.10) (.12) .62	Deidents 5.54 .54 .25 .25	#ah \$ 57.13 54.38 49.75	tow \$ 45.88 47.13 40.63	Parties 19-1 21 Spert 21 Spert Unit Dealer Land Econol Contende Contende	70707 L 60 F L 70 B L 70 B L 80 7 L 80 7 L 70 7 L 70 7 L 90 8 F 70 8	6 7 6 5 35 mg 5 5 5 6 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	8 0 3 c 3 c 3 c 3 c 3 c 3 c 3 c 3 c 3 c 3 c	0 0 8 28 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8	beind oppen open und ihr wich of a system Chapter to any of any open and the system of an analysis. Chapter and any open and analysis of a system o	a a 60 v	Even of protect jets!	FIDDE Protes, Labor villoys means grade to present less In means and include from Robust cases to this fue to this fue to this fue the tage and integrates	SURPLOAD mid and mint These facts precipitation encode expendion	75aga 2900-608 27 18-7351 (Ser
Teed Delass in millions except per shase and steck prices 1993 First quarter Second quarter Third quarter Fourth quarter Teed	Romue \$ 13,058 15,519 14,743 19,396 \$ 62,716	6ross Profit \$ 5,162 5,974 5,602 7,410 \$24,148	4.001 Earnings \$ (3999)* (8,036) (48) 382	L094 Earning: \$ (.30)* (14.10) (.12) .62 \$ (14.22)*	Dividents \$54	114p \$ 57.13 54.38 49.75 59.88	Low \$ 45.88 47.13 40.63 42.13	Parties 19-6 All Speet All Speet the bally List Ma Esseet Caseed Calcula	70707 L 60 F L 70 B L 70 B L 80 7 L 80 7 L 70 7 L 70 7 L 90 8 F 70 8	6 7 6 5 35 mg 5 5 5 6 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	8 0 3 c 3 c 3 c 3 c 3 c 3 c 3 c 3 c 3 c 3 c	0 (0 8 28 8 8 8 8 8 8 8 8 8 8 8 8 8 8	Interfairper comparison of the for descent Comparison of the formation of the sector comparison of the formation of the formation of the sector comparison of the formation of the formation of the sector comparison of the formation of the formation of the sector comparison of the formation of the formation of the sector comparison of the formation of the formation of the sector comparison of the formation of the formation of the formation of the formation of the formation of the formation of the formation of the formation of the formation of the formation of the formation of the formation of the formation of the formation of the formation of the formation of the formation of the formation of t	2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	Pose Dactarios Dactarios Date finos stands to protect Javel Roto stands Roto stands Roto stands Roto stands Roto stands Roto stands Roto stands Roto stands	FUENCE Notices, Lisboot valleys minute grade to present less Nr notestaal northese from Follow Later Jake to Citle Fues Provinged Long of FUENDER and Alexe	SURPLOAD mid and mint These facts precipitation encode expendion	71aga 900-608 SF 18-551 (Ren 19 90-601 b by Laboratory 7008-148 SF
Teel Delas: in millions except per shase and steck prices 1993 First quarter Second quarter Third quarter Total 1992 First quarter 1992 First quarter	Romone \$ 13,058 15,519 14,743 19,396 \$ 62,716 \$ 14,037	6088 Politi \$ 5,162 5,974 5,602 7,410 \$24,148 \$ 7,133	4.001 Earrings \$ (399)/* (8,036) (48) 382 \$ (8,101) \$ 2,512†	Lose Lering: \$ (30)* (14.10) (.12) 62 \$ (14.22)* \$ 4.45†	Deidents 5.54 .54 .25 .25	8 57.13 54.38 49.75 59.88 5 98.13	Low \$ 45.88 47.13 40.63 42.13 \$ 83.13	Parties 19-1 21 Spert 21 Spert Unit Dealer Land Econol Contende Contende	70707 L 60 F L 70 B L 70 B L 80 7 L 80 7 L 70 7 L 70 7 L 90 8 F 70 8	6 7 6 5 35 mg 5 5 5 6 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	B ( Sing Si Si () Si ()S	0 0 8 28 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8	Unit of program july or the of a years of the comparison of the program is a program of the prog	2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	Even of protect jets!	THEME Throws, i-stope sullays intel proof to prove the proof to prove the proof of the proof to club hus to club hus brister, and dees intel and brister, and dees brister, and dees brister, and dees brister, and dees brister, and dees brister, and and brister, and brist	322020ax mid and mint These Sacts precipitation encode expension block hats expension science precipitation	Toga 900-500 pr 14-951 Oter to 990-500 p ty Lateratory 7004-34 pp 12-852
Teed (Dolars in millions except per shase and steek poices) 1993 First quarter Second quarter Third quarter Teed 1992 First quarter Second quarter Second quarter Second quarter	Romue \$ 13,058 15,519 14,743 19,396 \$ 62,716 \$ 14,037 16,224	6ross Profit \$ 5,162 5,602 7,410 \$24,148 \$ 7,133 7,863	4 cel toring \$ (3999/ (8,036) (48) 382 \$ (8,101) \$ 2,542† 734	Loss Lening: \$ (30)* (14.10) (.12) 62 \$ (14.22)* \$ 4.45† 1.29	Dividents \$.54 .54 .55 .35 \$1.58 \$1.58 \$1.21 .121	8 57.13 54.38 49.75 59.88 \$ 98.13 98.63	Low \$ 45.88 47.13 40.63 42.13 \$ 83.13 81.63	namer 19-4 28 Speet 28 Speet 2		5 7 6 5 35 mg 5 6 4 5 5 5 8 9 10 10 10 10 2 10 5	0 3 mg 22 55 0 58 0 58 0 58 0 58 1 56 5 56 1 56 5 56 1 56 5 56 1 56 5 57 1 56 5 57 1 56 5 57 1 57 5 57 1 57 5 57 5 57 5 57 5 57	0 0 8 28 8 3 8 3 8 3 8 3 8 3 8 3 8 3 8 3 8 3 8	The second secon		Total a product joint Exercision Don Total and the Don Total and Don Total and Don Total and Ref of network at glacines Linke Total Crops Remo Nolitation at Total Rest	DIDOS Anties (stage valleys action grade to present the action of the stage of the to the factor of the to the factor to the fac	SERVICES mid wid mint These facts precipitation exceeds responsition moveds precipitation	730ga 9004505 pp 14-953 (New 14-953) (New 14-953) (New 13-950-160 17-004-16 pp 13-061) 1460-250 pp (13-961) 1460-250 pp (13-961)
Teel Delas: in millions except per shase and steck prices 1993 First quarter Second quarter Third quarter Total 1992 First quarter 1992 First quarter	Romone \$ 13,058 15,519 14,743 19,396 \$ 62,716 \$ 14,037	6ross Profit \$ 5,162 5,602 7,410 \$24,148 \$ 7,133 7,863 6,779	4 cel toring \$ (3999/ (8,036) (48) 382 \$ (8,101) \$ 2,542† 734	Lose Lering: \$ (30)* (14.10) (.12) 62 \$ (14.22)* \$ 4.45†	Deidents 5.54 .54 .25 .25	8 57.13 54.38 49.75 59.88 5 98.13	Low \$ 45.88 47.13 40.63 42.13 \$ 83.13	1980 1984 All Spert Ha Bally Lin Ma Eneral Eneral Eneral Caturals Caturals Caturals Caturals	1 10 / 10 / 1 66 P 1 86 P 1 86 P 1 8 1 1 8 1 1 8 7 1 8 7 1 8 1 1 8 1	5 7 6 5 35 mg 5 6 4 5 5 5 8 9 10 10 10 10 10 2 10 5	8 ( 3 mp 20 55 ( 55 (	0 (0 8 25 8 3 8 3 8 3 8 3 8 3 8 3 8 3 8 3 8 3 8 3	The second secon		International and a second and a second	NUMP: NUMPA, Langer village mean pade to prove the pade to prove the pade to prove the pade to pade the pade NUMPA (the pade NUMPA (th	32200000 nil wi mint Theo Rain presidentia toodi reporting here Rain responden norma presidente. Theo Rain presidente.	734pa 9900-608 pr 14-951 (Net 14-951) (Net 14-950) (Net 14-950) 14-6500 14-650 pr (L-160)
Teed (Dolars in millions except per shase and steek poices) 1993 First quarter Second quarter Third quarter Teed 1992 First quarter Second quarter Second quarter Second quarter	Romue \$ 13,058 15,519 14,743 19,396 \$ 62,716 \$ 14,037 16,224	6ross Profit \$ 5,162 5,602 7,410 \$24,148 \$ 7,133 7,863	4 cel toring \$ (3999/ (8,036) (48) 382 \$ (8,101) \$ 2,542† 734	Loss Lening: \$ (30)* (14.10) (.12) 62 \$ (14.22)* \$ 4.45† 1.29	Dividents \$.54 .54 .55 .35 \$1.58 \$1.58 \$1.21 .121	8 57.13 54.38 49.75 59.88 \$ 98.13 98.63	Low \$ 45.88 47.13 40.63 42.13 \$ 83.13 81.63	namer 19-4 28 Speet 28 Speet 2	<ul> <li>I II / III /</li> <li>L 60 P</li> <li>L 80 P</li> <lil 80="" <="" p<="" td=""><td>5 7 6 5 35 mp 55 55 53 53 53 53 53 53 53 53 53 53 53</td><td>0 ( 3 mp 20 55 0 58 0 58 0 58 0 58 1 56 5 55 1 56 5 55 1 55 2 51 2 51 2 51 2 51 2 51 2 51</td><td>0 0 0 8 29 84 28 84 28 8</td><td>The second secon</td><td></td><td>EVER &amp; proset jest DESC (Constraint) Statistics No forestand No forest lend Red names at glaters Deb faces Cryp Immo baltation of Dave Mait Like Table Time</td><td>THEME Throws, i-stope sullays intel proof to prove the proof to prove the proof of the proof to club hus to club hus brister, and dees intel and brister, and dees brister, and dees brister, and dees brister, and dees brister, and dees brister, and and brister, and brist</td><td>322020ax mid and mint These Sacts precipitation encode expension block hats expension science precipitation</td><td>73aga 9500-808 pr 14-953 (Net 13-950-800 by Landinov 13-851) 1460-953 pr (L3-861) 1460-953 pr (L3-861)</td></lil></ul>	5 7 6 5 35 mp 55 55 53 53 53 53 53 53 53 53 53 53 53	0 ( 3 mp 20 55 0 58 0 58 0 58 0 58 1 56 5 55 1 56 5 55 1 55 2 51 2 51 2 51 2 51 2 51 2 51	0 0 0 8 29 84 28 84 28 8	The second secon		EVER & proset jest DESC (Constraint) Statistics No forestand No forest lend Red names at glaters Deb faces Cryp Immo baltation of Dave Mait Like Table Time	THEME Throws, i-stope sullays intel proof to prove the proof to prove the proof of the proof to club hus to club hus brister, and dees intel and brister, and dees brister, and dees brister, and dees brister, and dees brister, and dees brister, and and brister, and brist	322020ax mid and mint These Sacts precipitation encode expension block hats expension science precipitation	73aga 9500-808 pr 14-953 (Net 13-950-800 by Landinov 13-851) 1460-953 pr (L3-861) 1460-953 pr (L3-861)

Fig. 10. Sample structure recognition output of TabStruct-Net on table images of UNLV dataset. First Row: prediction of cells which belong to the same row. Second Row: prediction of cells which belong to the same column. Cells marked with orange colour represent the examine cells and cells marked with green colour represent those which belong to the same row/column of the examined cell.

## 4.6 Failure Examples

Figure 11 shows some failure cases of our model in presence of empty spaces along both horizontal and vertical axes.

20 Raja et al.



Fig. 11. Sample intermediate cell detection results of Tabstruct-Net on table images of ICDAR-2013, ICDAR-2019 CTDAR, SCITSR, SCITSR-COMP, TableBank, PubTabNet and UNLV datasets illustrate failure of Tabstruct-Net.

# 4.7 Robustness of TabStruct-Net

CD Network	SR Network	IoU	CD S	Scores	5	$\mathbf{SR} \ \mathbf{S}$	$\mathbf{cores}$	
		TH	$\mathbf{P}\uparrow$	$\mathbf{R}\uparrow$	$\mathbf{F1}\uparrow$	$\mathbf{P}\uparrow$	$\mathbf{R}\uparrow$	$\mathbf{F1}\uparrow$
		0.5	0.942	20.948	30.945	0.933	80.915	<b>0.924</b>
		0.6	0.937	0.941	0.939	0.930	0.908	0.919
Mask R-CNN+TD+BU+AL	DGCNN+P2+LSTM	0.7	0.828	0.831	0.829	0.800	0.791	0.795
		0.8	0.651	0.670	0.660	0.638	0.624	0.631
		0.9	0.314	0.336	0.325	0.291	0.284	0.287

Table 13. Physical structure recognition results on ICDAR-2013-partial dataset for varying IoU thresholds to demonstrate TabStruct-Net's robustness. ES: indicates Experimental Setup, CD: indicates Cell Detection, TH: indicates IoU threshold value, SR: indicates Structure Recognition, P2: indicates using visual features from P2 layer of the FPN instead of using separate convolution blocks, LSTM: indicates use of LSTMs to model visual features along center-horizontal and center-vertical lines for every cell, TD+BU: indicates use of Top-Down and Bottom-Up pathways in the FPN, AL: indicates addition of alignment loss as a regularizer to TabStruct-Net, P: indicates precision, R: indicates recall, F1: indicates F1 Score.

CD Network	SR Network	IoU	CD S	Scores	;	$\mathbf{SR} \ \mathbf{S}$	cores	
		$\mathbf{TH}$	$\mathbf{P}\uparrow$	$\mathbf{R}\uparrow$	$\mathbf{F1}\uparrow$	$\mathbf{P}\uparrow$	$\mathbf{R}\uparrow$	$F1\uparrow$
		0.5	0.865	50.857	0.861	0.864	10.842	0.853
		0.6	0.84	0.836	0.838	0.822	0.787	0.804
Mask R-CNN+TD+BU+AL	DGCNN+P2+LSTM	0.7	0.694	0.681	0.687	0.641	0.625	0.633
		0.8	0.454	0.428	0.441	0.404	0.376	0.389
		0.9	0.201	0.153	0.174	0.175	0.138	0.154

Table 14. Physical structure recognition results on ICDAR-2019 dataset for varying IoU thresholds to demonstrate Tabstruct-Net's robustness. **ES**: indicates Experimental Setup, **CD**: indicates Cell Detection, **TH**: indicates IoU threshold value, **SR**: indicates Structure Recognition, **P2**: indicates using visual features from P2 layer of the FPN instead of using separate convolution blocks, LSTM: indicates use of LSTMs to model visual features along center-horizontal and center-vertical lines for every cell, TD+BU: indicates use of Top-Down and Bottom-Up pathways in the FPN, **AL**: indicates addition of alignment loss as a regularizer to Tabstruct-Net, **P**: indicates precision, **R**: indicates recall, **F1**: indicates F1 Score.

CD Network	SR Network	IoU	CD S	Scores		$\mathbf{SR} \mathbf{S}$	cores	
		$\mathbf{TH}$	$\mathbf{P}\uparrow$	$\mathbf{R}\uparrow$	$\mathbf{F1}\uparrow$	$\mathbf{P}\uparrow$	$\mathbf{R}\uparrow$	$\mathbf{F1}\uparrow$
		0.5	0.871	10.879	0.875	0.864	10.842	0.853
		0.6	0.858	0.864	0.861	0.849	0.828	0.839
Mask R-CNN+TD+BU+AL	DGCNN+P2+LSTM	0.7	0.751	0.773	0.762	0.735	0.711	0.723
		0.8	0.595	0.622	0.608	0.558	0.532	0.545
		0.9	0.214	0.237	0.225	0.173	0.148	0.160

Table 15. Physical structure recognition results on UNLV-partial dataset for varying IoU thresholds to demonstrate Tabstruct-Net's robustness. **ES**: indicates Experimental Setup, **CD**: indicates Cell Detection, **TH**: indicates IoU threshold value, **SR**: indicates Structure Recognition, **P2**: indicates using visual features from P2 layer of the FPN instead of using separate convolution blocks, LSTM: indicates use of LSTMs to model visual features along center-horizontal and center-vertical lines for every cell, TD+BU: indicates use of Top-Down and Bottom-Up pathways in the FPN, **AL**: indicates addition of alignment loss as a regularizer to Tabstruct-Net, **P**: indicates precision, **R**: indicates recall, **F1**: indicates F1 Score.

CD Network	SR Network	IoU	CD S	cores		$\mathbf{SR} \mathbf{S}$	cores	
		$\mathbf{TH}$	$\mathbf{P}\uparrow$	$\mathbf{R}\uparrow$	$F1\uparrow$	$\mathbf{P}\uparrow$	$\mathbf{R}\uparrow$	$F1\uparrow$
		0.5	0.939	0.944	0.941	0.930	0.922	0.926
		0.6	0.932	0.938	0.935	0.927	0.913	0.920
Mask R-CNN+TD+BU+AL	DGCNN+P2+LSTM	0.7	0.808	0.820	0.814	0.793	0.775	0.784
		0.8	0.639	0.652	0.645	0.618	0.594	0.606
		0.9	0.297	0.324	0.310	0.271	0.258	0.264

Table 16. Physical structure recognition results on SciTSR dataset for varying IoU thresholds to demonstrate TabStruct-Net's robustness. **ES**: indicates Experimental Setup, **CD**: indicates Cell Detection, **TH**: indicates IoU threshold value, **SR**: indicates Structure Recognition, **P2**: indicates using visual features from P2 layer of the FPN instead of using separate convolution blocks, LSTM: indicates use of LSTMs to model visual features along center-horizontal and center-vertical lines for every cell, TD+BU: indicates use of Top-Down and Bottom-Up pathways in the FPN, **AL**: indicates addition of alignment loss as a regularizer to TabStruct-Net, **P**: indicates precision, **R**: indicates recall, **F1**: indicates F1 Score.

4.8 Ablation Study

$\mathbf{ES}$	CD Network	SR Network	CD S	Scores		SR S	cores	
			$\mathbf{P}\uparrow$	$\mathbf{R}\uparrow$	$F1\uparrow$	$\mathbf{P}\uparrow$	$\mathbf{R}\uparrow$	$F1\uparrow$
	Mask R-CNN	DGCNN	0.835	0.843	0.839	0.885	0.864	0.874
	Mask r-cnn	dgcnn+P2	0.837	0.846	0.841	0.887	0.865	0.876
	Mask r-cnn	DGCNN+P2+LSTM	0.840	0.848	0.844	0.903	0.889	0.896
	Mask R-CNN+TD+BU	DGCNN	0.898	0.900	0.899	0.889	0.882	0.885
S-A	Mask r-cnn+td+bu	DGCNN+P2	0.902	0.905	0.903	0.893	0.885	0.889
	Mask r-cnn+td+bu	DGCNN+P2+LSTM	0.924	0.928	0.926	0.911	0.896	0.903
	Mask R-CNN+TD+BU+AL	DGCNN	0.920	0.924	0.922	0.915	0.892	0.903
	Mask R-CNN+TD+BU+AL	DGCNN+P2	0.924	0.927	0.925	0.918	0.894	0.906
	Mask R-CNN+TD+BU+AL	DGCNN+P2+LSTM	0.937	0.941	0.939	0.930	0.908	0.919
	-NA-	DGCNN	-NA-	-NA-	-NA-	0.986	0.990	0.988
S-B	-NA-	DGCNN+P2	-NA-	-NA-	-NA-	0.987	0.990	0.989
	-NA-	DGCNN+P2+LSTM	-NA-	-NA-	-NA-	0.991	0.993	0.992

Table 17. Ablation study for physical structure recognition on ICDAR-2013-partial dataset. ES: indicates Experimental Setup, CD: indicates Cell Detection, SR: indicates Structure Recognition, P2: indicates using visual features from P2 layer of the FPN instead of using separate convolution blocks, LSTM: indicates use of LSTMs to model visual features along center-horizontal and center-vertical lines for every cell, TD+BU: indicates use of Top-Down and Bottom-Up pathways in the FPN, AL: indicates addition of alignment loss as a regularizer to TabStruct-Net, P: indicates precision, R: indicates recall, F1: indicates F1 Score.

$\mathbf{ES}$	CD Network	SR Network	CD S	ores		SR S	cores	
			$\mathbf{P}\uparrow$	$\mathbf{R}\uparrow$	$F1\uparrow$	$\mathbf{P}\uparrow$	$\mathbf{R}\uparrow$	$F1\uparrow$
	Mask R-CNN	DGCNN	0.770	0.752	0.761	0.744	0.706	0.725
	Mask r-cnn	dgcnn+P2	0.774	0.761	0.767	0.751	0.718	0.734
	Mask r-cnn	DGCNN+P2+lstm	0.797	0.785	0.791	0.775	0.750	0.762
	Mask r-cnn+td+bu	DGCNN	0.775	0.761	0.768	0.751	0.713	0.732
S-A	Mask r-cnn+td+bu	DGCNN+P2	0.781	0.768	0.774	0.756	0.721	0.738
	Mask r-cnn+td+bu	DGCNN+P2+LSTM	0.803	0.790	0.796	0.782	0.754	0.768
	Mask R-CNN+TD+BU+AL	DGCNN	0.821	0.814	0.817	0.797	0.748	0.772
	Mask R-CNN+TD+BU+AL	DGCNN+P2	0.823	0.818	0.820	0.800	0.753	0.776
	Mask R-CNN+TD+BU+AL	DGCNN+P2+LSTM	0.840	0.836	0.838	0.822	0.787	0.804
	-NA-	DGCNN	-NA-	-NA-	-NA-	0.904	0.889	0.896
S-B	-NA-	DGCNN+P2	-NA-	-NA-	-NA-	0.932	0.921	0.927
	-NA-	DGCNN+P2+LSTM	-NA-	-NA-	-NA-	0.975	0.958	0.966

Table 18. Ablation study for physical structure recognition on ICDAR-2019 dataset. ES: indicates Experimental Setup, CD: indicates Cell Detection, SR: indicates Structure Recognition, P2: indicates using visual features from P2 layer of the FPN instead of using separate convolution blocks, LSTM: indicates use of LSTMs to model visual features along center-horizontal and center-vertical lines for every cell, TD+BU: indicates use of Top-Down and Bottom-Up pathways in the FPN, AL: indicates addition of alignment loss as a regularizer to TabStruct-Net, P: indicates precision, R: indicates recall, F1: indicates F1 Score.

$\mathbf{ES}$	CD Network	SR Network	CD S	Scores		SR S	cores	
			$\mathbf{P}\uparrow$	$\mathbf{R}\uparrow$	$F1\uparrow$	$\mathbf{P}\uparrow$	$\mathbf{R}\uparrow$	<b>F1</b> ↑
	Mask r-cnn	DGCNN	0.835	0.843	0.839	0.795	0.764	0.779
	Mask r-cnn	dgcnn+P2	0.837	0.846	0.841	0.812	0.788	0.800
	Mask r-cnn	DGCNN+P2+lstm	0.840	0.848	0.844	0.838	0.821	0.829
	Mask R-CNN+TD+BU	DGCNN	0.837	0.845	0.841	0.797	0.766	0.781
S-A	Mask r-cnn+td+bu	DGCNN+P2	0.840	0.849	0.844	0.815	0.790	0.802
	Mask r-cnn+td+bu	DGCNN+P2+LSTM	0.844	0.851	0.847	0.841	0.823	0.832
	Mask R-CNN+TD+BU+AL	DGCNN	0.847	0.855	0.851	0.802	0.775	0.788
	Mask R-CNN+TD+BU+AL	DGCNN+P2	0.853	0.860	0.856	0.823	0.797	0.810
	Mask R-CNN+TD+BU+AL	DGCNN+P2+LSTM	0.858	80.864	0.861	0.849	0.828	0.839
	-NA-	DGCNN	-NA-	-NA-	-NA-	0.921	0.898	0.909
S-B	-NA-	DGCNN+P2	-NA-	-NA-	-NA-	0.950	0.935	0.942
	-NA-	DGCNN+P2+LSTM	-NA-	-NA-	-NA-	0.992	0.994	0.993

Table 19. Ablation study for physical structure recognition on UNLV-partial dataset. ES: indicates Experimental Setup, CD: indicates Cell Detection, SR: indicates Structure Recognition, P2: indicates using visual features from P2 layer of the FPN instead of using separate convolution blocks, LSTM: indicates use of LSTMs to model visual features along center-horizontal and center-vertical lines for every cell, TD+BU: indicates use of Top-Down and Bottom-Up pathways in the FPN, AL: indicates addition of alignment loss as a regularizer to TabStruct-Net, P: indicates precision, R: indicates recall, F1: indicates F1 Score.

$\mathbf{ES}$	CD Network	SR Network	CD Scores			SR Scores		
			$\mathbf{P}\uparrow$	$\mathbf{R}\uparrow$	$\mathbf{F1}\uparrow$	$\mathbf{P}\uparrow$	$\mathbf{R}\uparrow$	$F1\uparrow$
S-A	Mask r-cnn	DGCNN	0.896	0.900	0.898	0.888	0.874	0.881
	Mask r-cnn	dgcnn+P2	0.904	0.907	0.905	0.892	0.879	0.885
	Mask r-cnn	DGCNN+P2+lstm	0.911	0.915	0.913	0.903	0.894	0.898
	Mask R-CNN+TD+BU	DGCNN	0.901	0.909	0.905	0.893	0.880	0.886
	Mask r-cnn+td+bu	DGCNN+P2	0.905	0.917	0.911	0.896	0.882	0.889
	Mask r-cnn+td+bu	DGCNN+P2+LSTM	0.918	0.924	0.921	0.905	0.898	0.902
	Mask R-CNN+TD+BU+AL	DGCNN	0.908	0.919	0.913	0.908	0.894	0.901
	Mask R-CNN+TD+BU+AL	DGCNN+P2	0.921	0.926	0.923	0.913	0.901	0.907
	Mask R-CNN+TD+BU+AL	DGCNN+P2+LSTM	0.932	20.938	0.935	0.927	0.913	0.920
S-B	-NA-	DGCNN	-NA-	-NA-	-NA-	0.970	0.981	0.976
	-NA-	DGCNN+P2	-NA-	-NA-	-NA-	0.973	0.986	0.979
	-NA-	DGCNN+P2+LSTM	-NA-	-NA-	-NA-	0.989	0.993	0.991

Table 20. Ablation study for physical structure recognition on SciTSR dataset. ES: indicates Experimental Setup, CD: indicates Cell Detection, SR: indicates Structure Recognition, P2: indicates using visual features from P2 layer of the FPN instead of using separate convolution blocks, LSTM: indicates use of LSTMs to model visual features along center-horizontal and center-vertical lines for every cell, TD+BU: indicates use of Top-Down and Bottom-Up pathways in the FPN, AL: indicates addition of alignment loss as a regularizer to TabStruct-Net, P: indicates precision, R: indicates recall, F1: indicates F1 Score.