2nd Place Solution for ICDAR 2021 Competition on Scientific Literature Parsing, Task B: Table Recogntion to HTML

Xianbiao Qi

Ping An Property and Casualty Insurance Company[1]

July 20, 2021

[1] Jiaquan Ye, Me, Yelin He, Yihao Chen, Dengyi Gu, Peng Gao, Rong Xiao

Outline

- Introduction to Table Recognition to HTML
- Our Solution[2]
  - Text Line Detection (TLD)
  - Table Structure Reconstruction (TSR)
  - Text Line Recognition (TLR)
  - Association Assignment
- Experiments

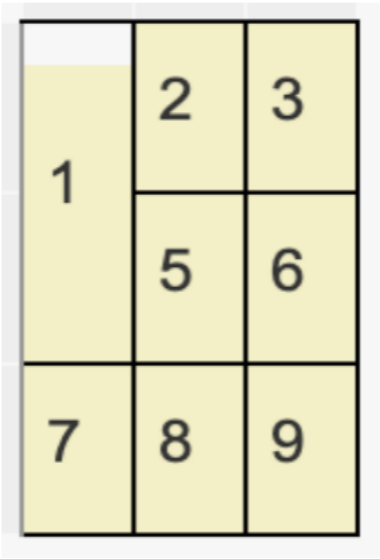
[2]Ye, Jiaquan, et al. "PingAn-VCGroup's Solution for ICDAR 2021 Competition on Scientific Literature Parsing Task B: Table Recognition to HTML." arXiv preprint arXiv:2105.01848 (2021).

Table2HTML

Company	2019-2020		2021
Microsoft	11	12	13
Apple	13	12	11
Google	11	13	12

```
<thead>
  <tr>
    <th class="tg-0lax">Company</th>
    <th class="tg-0lax" colspan="2">2019-2020</th>
    <th class="tg-0lax">2021</th>
  </tr>
</thead>
<tbody>
  <tr>
    <td class="tg-0lax">Microsoft</td>
    <td class="tg-0lax">11</td>
    <td class="tg-0lax">12</td>
    <td class="tg-0lax">13</td>
  </tr>
  <tr>
    <td class="tg-0lax">Apple</td>
    <td class="tg-0lax">13</td>
    <td class="tg-0lax">12</td>
    <td class="tg-0lax">11</td>
  </tr>
  <tr>
    <td class="tg-0lax">Google</td>
    <td class="tg-0lax">11</td>
    <td class="tg-0lax">13</td>
    <td class="tg-0lax">12</td>
  </tr>
</tbody>
```

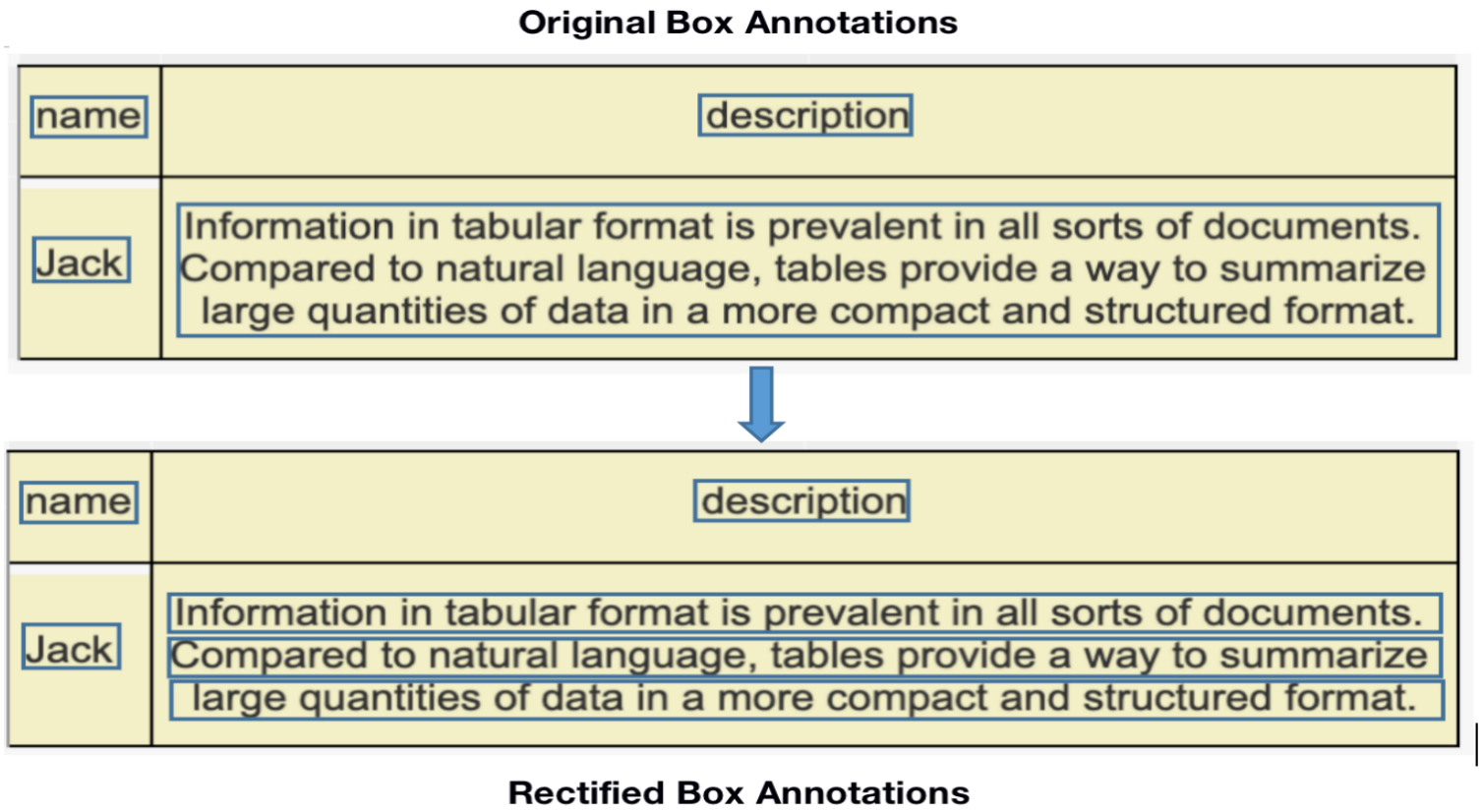
Table2HTML




```
<thead>
  <tr>
    <th class="tg-cly1" rowspan="2">1</th>
    <th class="tg-0lax">2</th>
    <th class="tg-0lax">3</th>
  </tr>
  <tr>
    <td class="tg-0lax">5</td>
    <td class="tg-0lax">6</td>
  </tr>
</thead>
<tbody>
  <tr>
    <td class="tg-0lax">7</td>
    <td class="tg-0lax">8</td>
    <td class="tg-0lax">9</td>
  </tr>
</tbody>
```

PSENet[3]

- Label Correction



[3]Wang, Wenhai, et al. "Shape robust text detection with progressive scale expansion network." CVPR. 2019.

Table Structure Reconstruction

Two key problems:

- Can we predict the item sequence?
- where is each item?

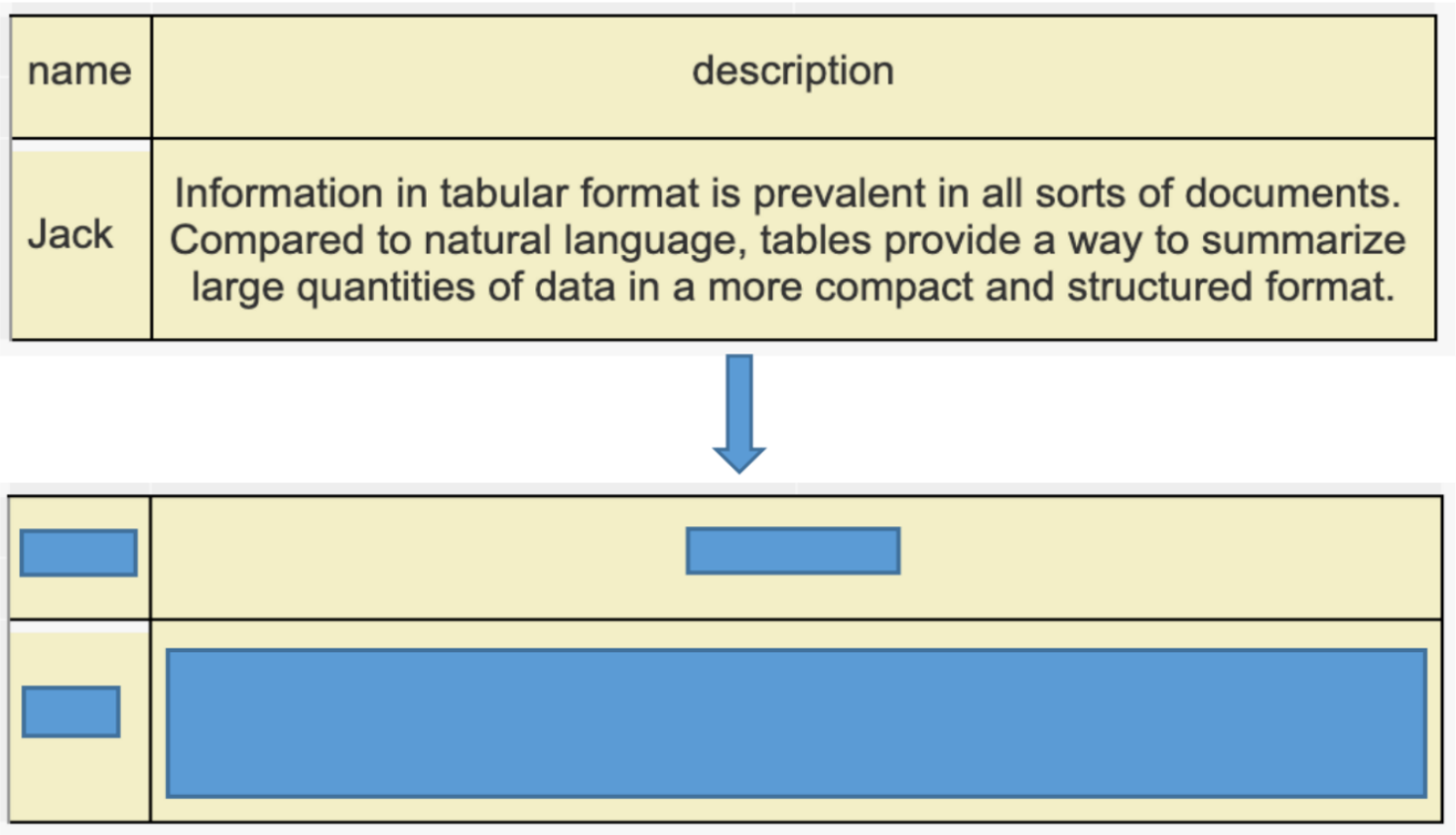


Table Structure Reconstruction

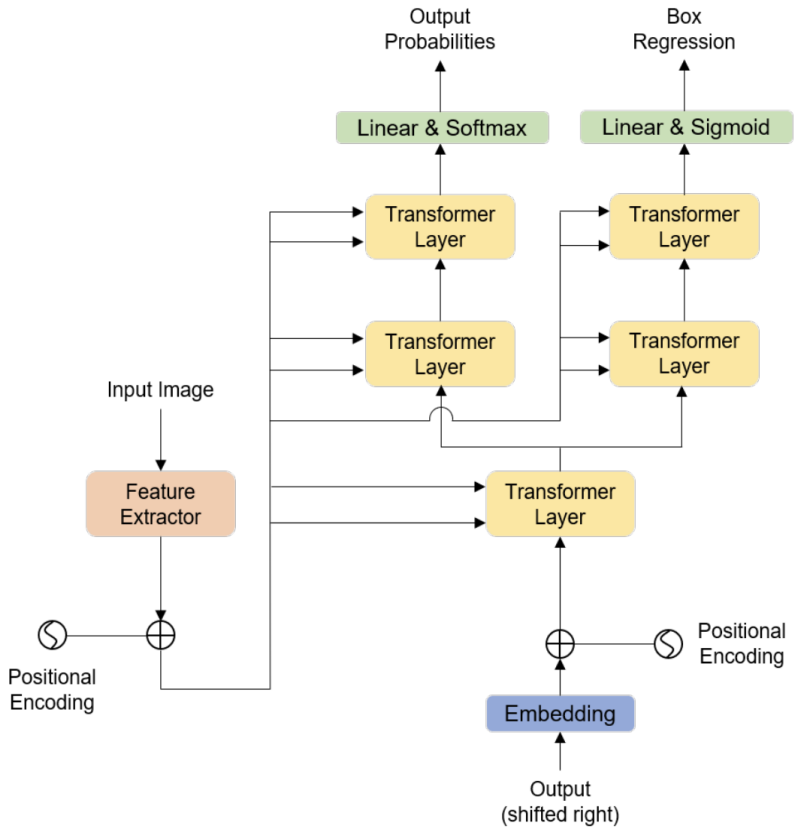
Table Structure Reconstruction

39 classes used for TSR

1. <thead>	11. rowspan=”3”	21. colspan=”4”	31. <td> </td>
2. </thead>	12. rowspan=”4”	22. colspan=”5”	32. <td><sep></sep><td>
3. <tbody>	13. rowspan=”5”	23. colspan=”6”	33. <td></td>
4. </tbody>	14. rowspan=”6”	24. colspan=”7”	34. <td></td>
5. <tr>	15. rowspan=”7”	25. colspan=”8”	35. <td> </td>
6. </tr>	16. rowspan=”8”	26. colspan=”9”	36. <td></td>
7. <td></td>	17. rowspan=”9”	27. colspan=”10”	37. <td> </td>
8. <td	18. rowspan=”10”	28. <td></td>	38. <td><sep></sep></hd>
9. >	19. colspan=”2”	29.<td> </td>	39. </td>
10. rowspan=”2”	20. colspan=”3”	30. <td></td>	

Table Structure Reconstruction

Our model structure, a customed Master[4] model



[4] Lu, N., Yu, W., Qi, X., Chen, Y., Gong, P., Xiao, R., and Bai, X. (2021). Master: Multi-aspect non-local network for scene text recognition. PR 2021.

Table Structure Reconstruction

How to encode the table structue into a sequence?

Company	2019-2020		2021
Microsoft	11	12	13
Apple	13	12	11
Google	11	13	12

```
<thead>
  <tr>
    <th class="tg-0lax">Company</th>
    <th class="tg-0lax" colspan="2">2019-2020</th>
    <th class="tg-0lax">2021</th>
  </tr>
</thead>
<tbody>
  <tr>
    <td class="tg-0lax">Microsoft</td>
    <td class="tg-0lax">11</td>
    <td class="tg-0lax">12</td>
    <td class="tg-0lax">13</td>
  </tr>
  <tr>
    <td class="tg-0lax">Apple</td>
    <td class="tg-0lax">13</td>
    <td class="tg-0lax">12</td>
    <td class="tg-0lax">11</td>
  </tr>
  <tr>
    <td class="tg-0lax">Google</td>
    <td class="tg-0lax">11</td>
    <td class="tg-0lax">13</td>
    <td class="tg-0lax">12</td>
  </tr>
</tbody>
```

1. <thead>	11. rowspan=”3”	21. colspan=”4”	31. <td> </td>
2. </thead>	12. rowspan=”4”	22. colspan=”5”	32. <td><sep></sep><td>
3. <tbody>	13. rowspan=”5”	23. colspan=”6”	33. <td></td>
4. </tbody>	14. rowspan=”6”	24. colspan=”7”	34. <td></td>
5. <tr>	15. rowspan=”7”	25. colspan=”8”	35. <td> </td>
6. </tr>	16. rowspan=”8”	26. colspan=”9”	36. <td></td>
7. <td></td>	17. rowspan=”9”	27. colspan=”10”	37. <td> </td>
8. <td	18. rowspan=”10”	28. <td></td>	38. <td><sep></sep></hd>
9. >	19. colspan=”2”	29.<td> </td>	39. </td>
10. rowspan=”2”	20. colspan=”3”	30. <td></td>	

Structure Sequence:

SOS->1->5->7->19->7->6->2->3->5->7->7->7->7->6->5->7->7->7->7->6->5->7->7->7->7->6->4->EOS

Text Line Recognition

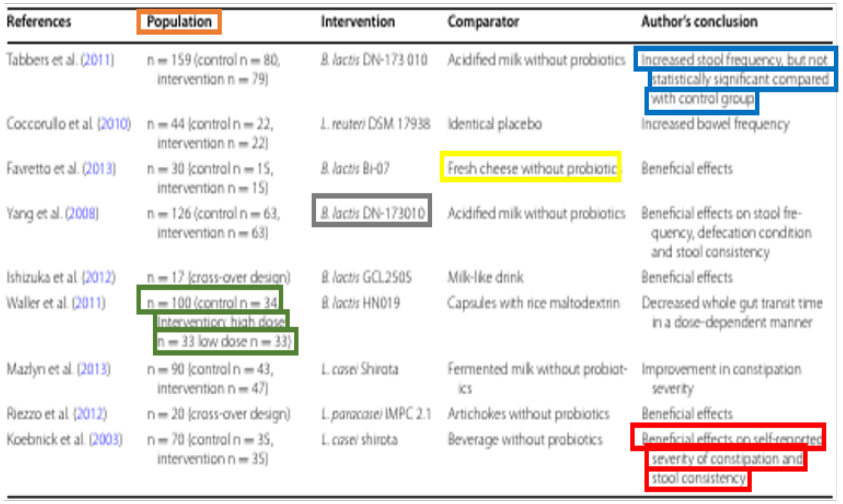
Master OCR

- Training stage: text images with multiple lines
- Testing stage: text image with single line (output by PSENet)

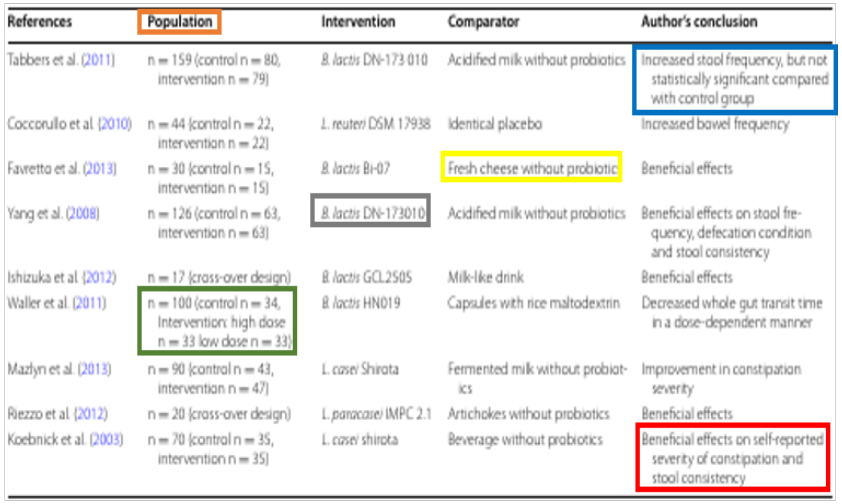
Association Assignment

Three steps consists of:

- Center Point Rule
- IOU
- Distance Rule



References	Population	Intervention	Comparator	Author's conclusion
Tabbers et al. (2011)	n = 159 (control n = 80, intervention n = 79)	*B. lactis* DN-173 010	Acidified milk without probiotics	Increased stool frequency, but not statistically significant compared with control group
Coccorullo et al. (2010)	n = 44 (control n = 22, intervention n = 22)	*L. reuteri* DSM 17938	Identical placebo	Increased bowel frequency
Favretto et al. (2013)	n = 30 (control n = 15, intervention n = 15)	*B. lactis* Bi-07	Fresh cheese without probiotic	Beneficial effects
Yang et al. (2008)	n = 126 (control n = 63, intervention n = 63)	*B. lactis* DN-173010	Acidified milk without probiotics	Beneficial effects on stool frequency, defecation condition and stool consistency
Ishizuka et al. (2012)	n = 17 (cross-over design)	*B. lactis* GCL2505	Milk-like drink	Beneficial effects
Waller et al. (2011)	n = 100 (control n = 34, intervention: high dose n = 33 low dose n = 33)	*B. lactis* HN019	Capsules with rice maltodextrin	Decreased whole gut transit time in a dose-dependent manner
Mazlyn et al. (2013)	n = 90 (control n = 43, intervention n = 47)	*L. casei* Shirota	Fermented milk without probiotics	Improvement in constipation severity
Riezzo et al. (2012)	n = 20 (cross-over design)	*L. paracasei* IMPC 2.1	Artichokes without probiotics	Beneficial effects
Koebnick et al. (2003)	n = 70 (control n = 35, intervention n = 35)	*L. casei* shirota	Beverage without probiotics	Beneficial effects on self-reported severity of constipation and stool consistency



References	Population	Intervention	Comparator	Author's conclusion
Tabbers et al. (2011)	n = 159 (control n = 80, intervention n = 79)	*B. lactis* DN-173 010	Acidified milk without probiotics	Increased stool frequency, but not statistically significant compared with control group
Coccorullo et al. (2010)	n = 44 (control n = 22, intervention n = 22)	*L. reuteri* DSM 17938	Identical placebo	Increased bowel frequency
Favretto et al. (2013)	n = 30 (control n = 15, intervention n = 15)	*B. lactis* Bi-07	Fresh cheese without probiotic	Beneficial effects
Yang et al. (2008)	n = 126 (control n = 63, intervention n = 63)	*B. lactis* DN-173010	Acidified milk without probiotics	Beneficial effects on stool frequency, defecation condition and stool consistency
Ishizuka et al. (2012)	n = 17 (cross-over design)	*B. lactis* GCL2505	Milk-like drink	Beneficial effects
Waller et al. (2011)	n = 100 (control n = 34, intervention: high dose n = 33 low dose n = 33)	*B. lactis* HN019	Capsules with rice maltodextrin	Decreased whole gut transit time in a dose-dependent manner
Mazlyn et al. (2013)	n = 90 (control n = 43, intervention n = 47)	*L. casei* Shirota	Fermented milk without probiotics	Improvement in constipation severity
Riezzo et al. (2012)	n = 20 (cross-over design)	*L. paracasei* IMPC 2.1	Artichokes without probiotics	Beneficial effects
Koebnick et al. (2003)	n = 70 (control n = 35, intervention n = 35)	*L. casei* shirota	Beverage without probiotics	Beneficial effects on self-reported severity of constipation and stool consistency

Ablation Study

TSR

- TSR: Adam vs Ranger
- Evaluation of SyncBN and Feature Concatenation for TSR and TLR

Optimizer	Structure prediction Acc.
Adam	0.7784
Ranger	**0.7826**

(a) Comparison of optimizer.

Feature Concatenation	Text line recognition Acc.
No	0.9313
Yes	**0.9347**

(b) Comparison of with or without feature concatenation.

SyncBN	FC	Structure prediction Acc.
		0.7734
✓		0.7750
✓	✓	**0.7785**

(c) Evaluation of label encoding, SyncBN and feature concatenation.

Ablation Study

TEDS

- Text Line Detection (TLD)
- Table Structure Recognition (TSR)
- Text Line Recognition (TLR)
- Model Ensemble (ME)
- Empty Space Box (ESB)
- Synchronized BN (SyncBN)
- Feature Concatenate (FeaC)
- Format Correction (ForC)

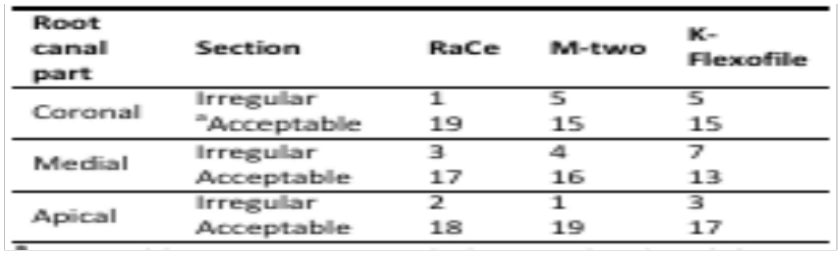
TLD	TSR			TLR	BA	ME	ForC	TEDS
PSE	ESB	SyncBN	FeaC	FeaC	Extra Insert			
✓								0.9385
✓	✓	✓	✓		✓			0.9621
✓	✓	✓	✓	✓	✓			0.9626
✓	✓	✓	✓	✓	✓	✓		0.9635
✓	✓	✓	✓	✓	✓	✓	✓	**0.9684**

Final Result

Team Name	TEDS Simple	TEDS Complex	TEDS all
Davar-Lab-OCR	97.88	94.78	**96.36**
VCGroup	**97.90**	94.68	96.32
XM	97.60	**94.89**	96.27
YG	97.38	94.79	96.11
DBJ	97.39	93.87	95.66
TAL	97.30	93.93	95.65
PaodingAI	97.35	93.79	95.61
anyone	96.95	93.43	95.23
LTIAYN	97.18	92.40	94.84

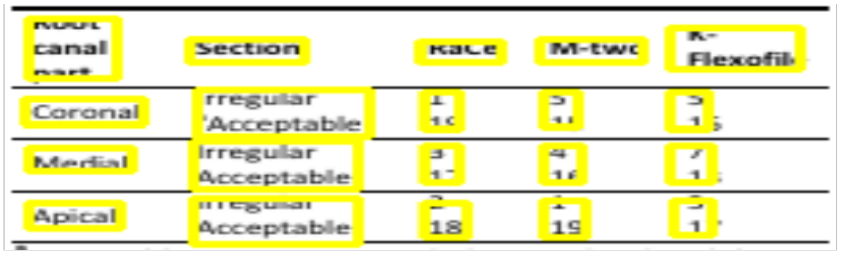
Table 4. Task B top TEDS results. The overall result (TEDS all) is decompose into simple and complex tables [16]

Error Analysis



Root canal part	Section	RaCe	M-two	K-Flexofile
Coronal	Irregular	1	5	5
	°Acceptable	19	15	15
Medial	Irregular	3	4	7
	Acceptable	17	16	13
Apical	Irregular	2	1	3
	Acceptable	18	19	17

(a) input image



(b) visulization of structure prediction

```
<td></td>  <td></td>  <td></td> <td></td>  <td></td>
<rsp=2></td>  <td></td>  <td></td> <td></td>  <td></td>
                <td></td>  <td></td> <td></td>  <td></td>
<rsp=2></td>  <td></td>  <td></td> <td></td>  <td></td>
                <td></td>  <td></td> <td></td>  <td></td>
<rsp=2></td>  <td></td>  <td></td> <td></td>  <td></td>
                <td></td>  <td></td> <td></td>  <td></td>
```

(c) structure GT

```
<td></td>  <td></td>  <td></td> <td></td>  <td></td>
<td></td>  <td></td>  <td></td> <td></td>  <td></td>
<td></td>  <td></td>  <td></td> <td></td>  <td></td>
<td></td>  <td></td>  <td></td> <td></td>  <td></td>
```

(d) structure prediction

```
<thead><tr><td><b>Root canal part</b
></td><td><b>Section</b></td><td><b>
RaCe</b></td><td><b>M-two</b></td><t
d><b>K-Flexofile</b></td></tr></thea
d><tbody><tr><td rowspan="2">Coronal
</td><td>Irregular</td><td>1</td><td
>5</td><td>5</td></tr><tr><td>aAccep
table</td><td>19</td><td>15</td><td>
15</td></tr><tr><td rowspan="2">Medi
al</td><td>Irregular</td><td>3</td><
td>4</td><td>7</td></tr><tr><td>Acce
ptable</td><td>17</td><td>16</td><td
>13</td></tr><tr><td rowspan="2">Api
cal</td><td>Irregular</td><td>2</td>
<td>1</td><td>3</td></tr><tr><td>Acc
eptable</td><td>18</td><td>19</td><t
d>17</td></tr></tbody>
```

(e) HTML code GT

```
<thead><tr><td><b>Root canal part</b
></td><td><b>Section</b></td><td><b>
RaCe</b></td><td><b>M-two</b></td><t
d><b>K- Flexofile</b></td></tr></the
ad><tbody><tr><td>Coronal</td><td>Ir
regular <sup>a</sup>Acceptable</td><
td>1 19</td><td>5 15</td><td>5 15</t
d></tr><tr><td>Medial</td><td>Irregu
lar Acceptable</td><td>3 17</td><td>
4 16</td><td>7 13</td></tr><tr><td>A
pical</td><td>Irregular Acceptable</
td><td>2 18</td><td>1 19</td><td>3 1
7</td></tr></tbody>
```

(f) HTML code prediction

Our Code is Released

https://github.com/wenwenyu/MASTER-pytorch


中国平安
PING AN
金融·科技