# Supplemental material

Xu Zhong, Elaheh ShafieiBavani, and Antonio Jimeno Yepes

IBM Research Australia, 60 City Road, Southgate, VIC 3006, Australia
peter.zhong@au1.ibm.com

## A Example of perturbations

| gene | ALK | BRAF | CDKN1A | EGFR | KRAS | MET | MYC | TP53 | TP63 | PI3K |
|---|---|---|---|---|---|---|---|---|---|---|
| HCC827 | n | n | n | 4 | 3 | n | 4 | 3 | n | n |
| A549 | n | n | 1 | 3 | 3 | n | 3 | n | 1 | n |
| NCI-H460 | n | n | 1 | 1 | n | n | 4 | n | n | n |

(a) Before perturbation

| gene | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | ALK | | | | | | | | | |
| | | BRAF | | | | | | | | |
| | | | CDKN1A | | | | | | | |
| | | | | EGFR | | | | | | |
| | | | | | KRAS | | | | | |
| | | | | | | MET | | | | |
| | | | | | | | MYC | | | |
| | | | | | | | | TP53 | | |
| | | | | | | | | | TP63 | PI3K |
| HCC827 | n | n | n | 4 | 3 | n | 4 | 3 | n | n |
| A549 | n | n | 1 | 3 | 3 | n | 3 | n | 1 | n |
| NCI-H460 | n | n | 1 | 1 | n | n | 4 | n | n | n |

(b) After perturbation

Fig. S1: Example of cell shift perturbation, where 90% of the cells in the first row are shifted. TEDS = 34.9%. Adjacency relation F1 score = 80.3%.

| Method | Data Type | Mean (m) | RMSE (m) | P90% (m) | PGSD (%) |
|---|---|---|---|---|---|
| Improved FCM | Gaofen-3 | 5.77 | 5.89 | 10.07 | 94.37 |
| | Sentinel-1 | 6.30 | 5.83 | 14.03 | 80.00 |
| Original FCM | Gaofen-3 | 6.97 | 7.66 | 13.87 | 90.70 |
| | Sentinel-1 | 8.53 | 4.81 | 13.14 | 90.00 |

(a) Before perturbation

| Meahod | aatb Tapa | Mean (m) | RMSa (m) | P90% (m) | aGSD a%) |
|---|---|---|---|---|---|
| Improved FCM | Gbofen-3 | 5.77 | 5.89 | 1a.07 | 94.37 |
| | Sentiael-a | a.30 | 5.83 | 14.03 | 80.00 |
| Original FCM | Gaofena3 | 6.9a | 7.66 | 1a.87 | 90.7a |
| | Santinel-1 | 8.53 | 4.8a | 13.14 | 90.00 |

(b) After perturbation

Fig. S2: Example of cell content perturbation, where the chance that a character gets modified is 10%. TEDS = 93.2%. Adjacency relation F1 score = 19.1%.

## B Categories of table styles in the synthetic dataset

The synthetic dataset introduced in [4] contains four categories of table styles. Fig. S3 shows a sample table image in each style category.

| Book | on | assumed | structures Table | Primary | and | of | expenses Ford |
|---|---|---|---|---|---|---|---|
| t unsatu- | Alabama | Stuckert | PROTOCOL | Latin Net | above | Dividends | Static |
| | BI amount | 400 | TOWER Southwest | Total | 4 | 1992 | $ |
| 70,000 | million small | 35 | copper | land, | 145 | 42.1 | (264) |
| 11.4 | oil benefits | 3,000 | field | tax environment | 5 | 0 | $500 |
| 02 | | 801.9 | is | pair N.A. | 64,523 | | (14) |
| 706 | of exercised | 6.1.2.3.3 | Flight Tax | a N.A. | 0.44 | 188.6 | data(c) |
| 12.6 | plan b | 0.0 | moderately Depreciation, | borrowings | | | q2c |
| 19,545 | AIM | 2 | Dowden | Other Sleeve | .002 | 0.96 | * |

(a) Category 1

| traverse | None from | Ba | DATA criterion | BLOWFISH | property |
|---|---|---|---|---|---|
| J Modeling | Basin Tar | operations TO | Cut and | Atriplex limestone, | cash Martin, |
| 3,667 | 0.01 | GOODRICH | H | 0.001 | 497.7 |
| 65- | 113 | of or | -C unit. | 100 | 11 |
| 3.20 | 3.5 | while | Marketing for | 0.06 | 31 |
| 0.084 | 0980.873 | income and | gray, sorting | 220 | -5.1 |
| | 1,024 | expected of | Rainbow | 130 | 1989 |
| 79,467 | 102.7 | Products | Tornado Consolidated | 6 | 5.05 |
| 715.9 | | of | Nevares beginning | 4 | 847.00 |
| 56 | 1.6 | b | Saturated, | 12.0 | 3 |
| 9 | 17.420 | will | | 1.21 | 59.36 |
| 47. | 00 | study. income | ite Language | 2,400 | 0.89 |
| 966 | 504.0 | of | DIMENSION | 7.6 | 22.8 |
| 9.230 | 5 | .Perseid Same | WORLD monitor | 1,468 | 6.1.2.1.1 |

(b) Category 2

| projects, | Proceeds Ramp | N.A. | | | |
|---|---|---|---|---|---|
| | | Bulk x | isolation component | OPERATING | share area |
| | pr144m | 0.011 | Maintenance WONDERFUL | 6 | Earthquake |
| • | 22:10:460 | 757 | Barney, WEAPONS | 320 | con- |
| 983e06 | outstanding: | 20.2 | Elev. | | temperatures Stephens |
| *3 | (25100) | 08-28-81 | Minor changes | 230 | |
| la19 | h1 | .73 | CONTROLLED | 1,959 | receivables, High-Level |
| | 015±025 | | and | 21 | of two-door |
| 43liter | $ | 60 | Underground | 25.4 | ENDED rate |
| r1945 | | 6 | OFF-SITE | 9.5 | Locking |
| (11) | c1 | 18,717 | operations THIRD | 21.6 | Rancho property, |

(c) Category 3

| equipment Material | RES operations | investments X | Craig before | others Comparison | | | Active | ASEMBL |
|---|---|---|---|---|---|---|---|---|
| | | | | borrowings Texas | increase | moderate-red-of | | |
| | (281) | of | (500 | 63 | 1.79 | Price | 242 | INTANGIBLE in |
| | psrv's | kg drilling | | -.030 | 729 | temperature | 20 | attributable Saturated |
| | 2043e08 | Vol- of | | 11.7 | 1.4 | Total | 3.075 | Percent |
| | stock: | Oil | sief77gliit'aility | 47,100 | 1992 | PVC of | -.003 | OCRWM |
| 0.20 | pa) | Number | spc2(11)2 | 33 | 567,527 | early Atlantic | 10.372 | |
| | solid/mthm | Procedure | 7000e+01 | 111111 | 5.2 | IIII | | confertifolia |
| | | Nonallocated Segment | (15) | 8 | 18.680 | CO | .82 | lock- |
| | | Assets | operations: | 43 | 1970 | YOU Day- | 23 | A. for |
| | 5812e02 | Car Net | pa) | 116,3 | 12,900 | in | 1.71 | all Numerous |
| 05 | 4/2) | L.A. | <1 | 133.0 | 219 | Orono | 175 | Job AC |
| 1 | 1±1 | zero | (14) | 00006 | 1,943 | Inventories Pumping | 21 | between |

(d) Category 4

Fig. S3: Sample table image of the four categories of table styles defined in [4].

## C Evaluation with adjacency relation metric

Table S1 compares the test performance between our EDD model and the 7 baseline approaches described in Section 6 under the conventional adjacency relation metric [2]. In this metric, for an adjacency relation to be true positive, the content of the two adjacent cells must be identical to the ground truth. In order to quantify the finer-grained performance on cell content recognition, we replace the exact match criterion with a softer criterion whether the normalized Levenshtein similarity of the two stringss $\geq$ a given threshold. To relax the dependence on the choice of the threshold, we follow the COCO competition [3] and the ICDAR 2019 Table Recognition Competition [1], where we compute the precision, recall, and F1-score over multiple thresholds $(0.5, 0.6, \cdots, 0.9)$. Table S1 depicts the mean precision, recall, and F1-score over the thresholds, which shows a similar pattern compare to the results in Table 2. EDD still considerably outperforms all the baselines.

| Input | Method | Mean precision (%) | | | Mean recall (%) | | | Mean F1-score (%) | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Simple | Complex | All | Simple | Complex | All | Simple | Complex | All |
| PDF | Tabula | 73.9 | 59.1 | 66.5 | 75.1 | 50.6 | 62.9 | 74.2 | 54.0 | 64.1 |
| | Traprange | 62.9 | 52.6 | 57.8 | 67.2 | 48.9 | 58.0 | 64.8 | 50.3 | 57.5 |
| | Camelot | 78.7 | 75.4 | 77.0 | 81.2 | 71.6 | 76.4 | 79.7 | 73.1 | 76.4 |
| | PDFPlumber | 38.4 | 35.4 | 36.9 | 39.9 | 32.8 | 36.3 | 38.7 | 33.6 | 36.1 |
| | Acrobat® Pro | 71.1 | 68.1 | 69.6 | 68.8 | 58.1 | 63.4 | 69.4 | 61.6 | 65.5 |
| Image | Acrobat® Pro | 59.3 | 64.3 | 61.8 | 54.0 | 51.2 | 52.6 | 55.5 | 55.3 | 55.4 |
| | WYGIWYS | 58.9 | 59.3 | 59.1 | 58.4 | 56.7 | 57.5 | 58.5 | 57.8 | 58.1 |
| | **EDD** | **85.1** | **80.2** | **82.6** | **84.6** | **78.4** | **81.5** | **84.7** | **79.0** | **81.9** |

Table S1: Test performance of EDD and 7 baseline approaches under the adjacency relation metric.

## D Runtime of TEDS

Table S2 shows the runtime of TEDS on the test results predicted by EDD and the 7 baseline approaches described in Section 6. The runtime is measured by running TEDS with 28 parallel threads on a 28 core[1] Linux machine.

| Input | Method | Average TEDS (%) | Runtime (minutes) |
|---|---|---|---|
| PDF | Tabula | 67.9 | 15.6 |
| | Traprange | 55.4 | 15.6 |
| | Camelot | 73.0 | 15.8 |
| | PDFPlumber | 40.4 | 17.3 |
| | Acrobat® Pro | 65.3 | 11.0 |
| Image | Acrobat® Pro | 53.7 | 9.7 |
| | WYGIWYS | 78.6 | 11.1 |
| | EDD | 88.3 | 13.7 |

Table S2: Runtime of TEDS on the test results predicted by EDD and 7 baseline approaches

## E Settings of the encoder

Table S3 depicts the validation performance of the 5 different settings of the encoder described in Section 6.1.

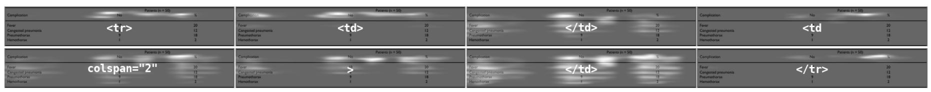
| Setting | Average TEDS (%) | | |
|---|---|---|---|
| | Simple | Complex | All |
| EDD-S1 | 91.0 | 84.7 | 87.9 |
| EDD-S2 | 87.0 | 79.5 | 83.3 |
| EDD-S1S1 | 91.3 | 85.4 | **88.4** |
| EDD-S2S1 | 91.3 | 85.0 | 88.1 |
| EDD-S2S2 | 88.8 | 82.0 | 85.4 |

Table S3: Validation performance of different settings of the encoder.

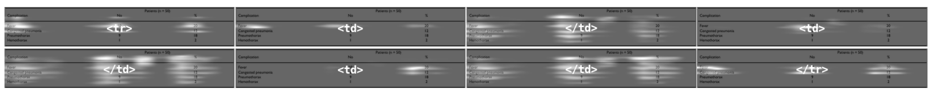
[1]Intel(R) Xeon(R) Platinum 8280M CPU @ 2.70GHz

## F Visualization of attention

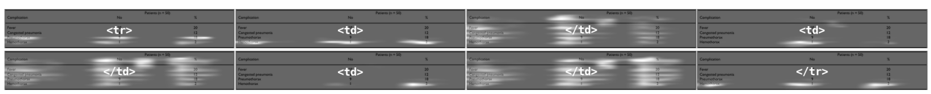
Figures S4 (a) - (c) illustrate the attention of the structure decoder when processing an example input table.



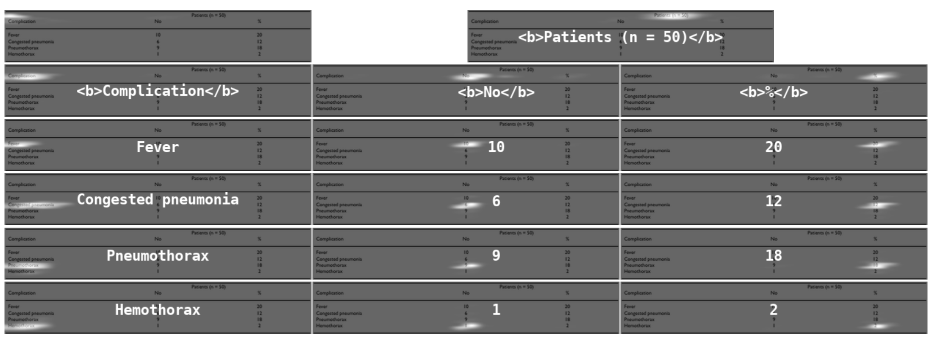

(a) Attention of structure decoder on the first header row




(b) Attention of structure decoder on the first body row




(c) Attention of structure decoder on the last body row




(d) Aggregated attention of cell decoder on each cell

Fig. S4: Attention distribution of the structure decoder (a - c) and the cell decoder (d) on an example input table. The texts at the center of the images are predictions of the EDD model. The structure decoder focuses its attention around table cells when recognizing new rows and cells, whereas the cell decoder places more focused attention on cell content.

## References

1. Gao, L., Huang, Y., Li, Y., Yan, Q., Fang, Y., Dejean, H., Kleber, F., Lang, E.M.: ICDAR 2019 competition on table detection and recognition. In: 2019 International Conference on Document Analysis and Recognition (ICDAR). pp. 1510–1515. IEEE (Sep 2019). https://doi.org/10.1109/ICDAR.2019.00166

2. Hurst, M.: A constraint-based approach to table structure derivation (2003)
3. Lin, T.Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., Zitnick, C.L.: Microsoft coco: Common objects in context. In: European conference on computer vision. pp. 740–755. Springer (2014)
4. Qasim, S.R., Mahmood, H., Shafait, F.: Rethinking table recognition using graph neural networks pp. 142–147 (Sep 2019). https://doi.org/10.1109/ICDAR.2019.00166