

Evaluating Table Structure Recognition: A New Perspective

Tarun Kumar and Himanshu Sharad Bhatt
 {tarun.kumar16, himanshu.s.bhatt}@aexp.com
 AI Labs, American Express

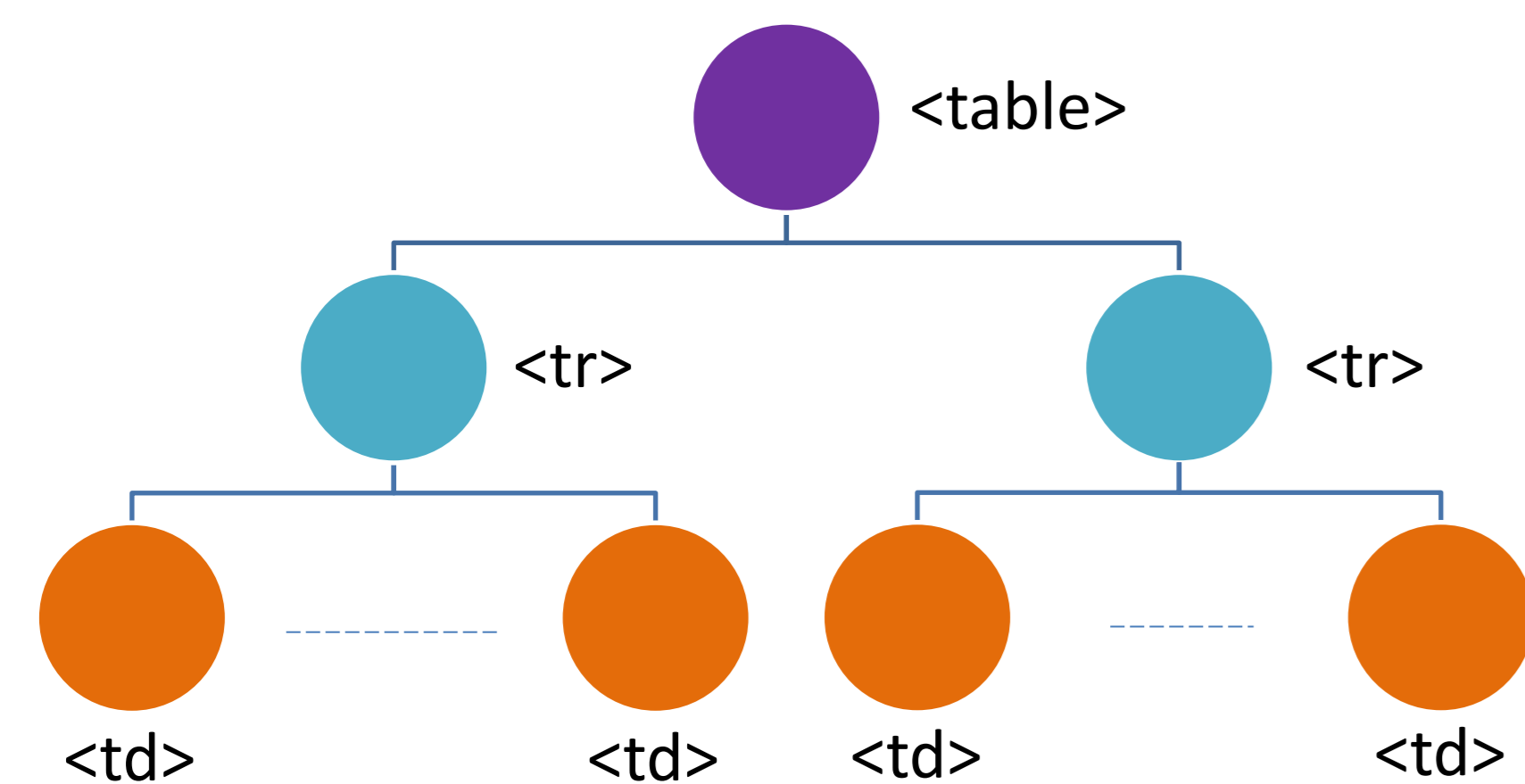
INTRODUCTION

- Table structure recognition algorithms aim at predicting bounding boxes of cells and their logical relation with one another, irrespective of the text in the cell.
- Thus, an evaluation metric should not penalize an algorithm's text inaccuracies.
- We emphasize that textual information depends on OCR, which is a separate problem all together.
- We present limitations of existing metrics and propose a new metric which is text agnostic and accounts only for layout of cells.

EXISTING METRICS

Existing Metric	Limitations
Adjacency Relation (Text)	Doesn't handle empty cells, misalignment of cells beyond immediate neighbours & text dependent
Adjacency Relation (IOU)	Doesn't handle empty cells, misalignment of cells beyond immediate neighbours
TEDS (Text)	Text dependent but less strict due to Levenshtein distance

- Adjacency Relation (Text) computes pair-wise relations between non-empty adjacent cells and their text.
- Adjacency Relation (IOU) is a text-independent metric where original non-empty cells are mapped to predicted cells by leveraging IOU thresholds.



- TEDS (text) considers the structure as a HTML tree and computes the edit distance between cells' texts as compared to the exact match in Adjacency Relation (Text).

PROPOSED METRIC

	A	B	
C			
D	E	F	G
H	I	J	K

Ground truth

C	A	B	
D	E	F	G
H	I	J	K

Predicted

- In the above figure, the predicted table missed one entire row. The only extra predicted relation is $\{C, A, Horizontal\}$.
- Due to this, only precision was affected, but recall is 100%, which is clearly inaccurate.
- In IOU based metric, "C" in ground truth is mapped to "C" cell in prediction and we still get the relation $\{C, A, Horizontal\}$.
- This demonstrates its inability to capture empty cells and mis-alignments.
- TEDS (Text) metric solves the problems of empty cells and multi-hop mis-alignments by considering all cells.

$$TEDS (IOU) = 1 - \frac{EditDistIOU(T_a, T_b)}{\max(|T_a|, |T_b|)}$$

- We propose TEDS (IOU) which replaces the edit distance between cells' text with the IOU distance between their bounding boxes.
- This removes dependency on text, while also preserving the benefits of TEDS (text) metric.
- Cost of edit is $1 - IOU(n_s, bbox, n_t, bbox)$ if both n_s & n_t are cells with same span.
- IOU distance ($IOU_d = 1 - IOU$) being a Jaccard index, is a metric as it satisfies:

- $IOU_d(A, B) \Leftrightarrow A = B$ (Identity)
- $IOU_d(A, B) = IOU_d(B, A)$ (Symmetry)
- $IOU_d(A, C) \leq IOU_d(A, B) + IOU_d(B, C)$ (Triangle Inequality)

RESULTS & DISCUSSION

Parameter	Controls	Active RA	OA	RA in remission
n	34	28	12	36
Age (mean ± standard deviation [range]; years)	48 ± 16 (24–62)	51 ± 17 (20–83)	60 ± 9 (49–73)	48 ± 11 (25–67)
Sex (male/female)	6/17	9/28	3/9	7/29
Disease duration (mean ± standard deviation [range]; years)	NA	5.1 ± 7.5 (0.1–37)	NA	9.3 ± 6.8 (2–28)
Remission duration (mean ± standard error [range]; months)	NA	NA	NA	29 ± 29 (6–144)
CRP (mean ± standard deviation [range]; mg/l), below /above detection ^a	NA	55 ± 52 (5–164), 0/28	NA	3.5 ± 5.2 (0–12), 23/13

Ground truth table

Parameter	Controls	Active RA	OA	RA in remission
n	34	28	12	36
Age (mean ± standard deviation [range]; years)	48 ± 16 (24–62)	51 ± 17 (20–83)	60 ± 9 (49–73)	48 ± 11 (25–67)
Sex (male/female)	6/17	9/28	3/9	7/29
Disease duration (mean ± standard deviation [range]; years)	NA	5.1 ± 7.5 (0.1–37)	NA	9.3 ± 6.8 (2–28)
Remission duration (mean ± standard error [range]; months)	NA	NA	NA	29 ± 29 (6–144)
CRP (mean ± standard deviation [range]; mg/l), below /above detection ^a	NA	55 ± 52 (5–164), 0/28	NA	3.5 ± 5.2 (0–12), 23/13

Red-lines: Predicted structure | Blue Lines: True structure

- In this example, OCR was unable to recognize the \pm symbol (it got recognized as +) and unable to detect "NA".
- Adjacency Relation (Text) got a poor score of 13.7 F1 due to the exact text match constraint.
- Adjacency Relation (IOU), being text independent, is more robust and achieves a F1 of 59.8.
- TEDS (text) matches text through edit distances and scores 71.6.
- TEDS (IOU) computing IOU_d between cells, assigns a higher score of 80.6 which seems the most representative.
- A possible extension is to introduce different thresholds for the IOU as in Adjacency Relation (IOU), instead of using absolute no.

REFERENCES

[1] Gao, L., Huang, Y., D'ejean, H., Meunier, J.L., Yan, Q., Fang, Y., Kleber, F., Lang, E.: Icdar 2019 competition on table detection and recognition (ctdar). In: 2019 ICDAR. pp. 1510–1515. IEEE (2019)

[2] Gobel, M., Hassan, T., Oro, E., Orsi, G.: Icdar 2013 table competition. In: 2013 ICDAR. pp. 1449–1453. IEEE (2013)

[3] Jimeno Yepes, A., Zhong, P., Burdick, D.: Icdar 2021 competition on scientific literature parsing. In: ICDAR. pp. 605–617. Springer (2021)

[4] Kosub, S.: A note on the triangle inequality for the jaccard distance. Pattern Recognition Letters 120, 36–38 (2019)

[5] Zhong, X., ShafieiBavani, E., Jimeno Yepes, A.: Image-based table recognition: data, model, and evaluation. In: ECCV. pp. 564–580. Springer (2020)