TABFIT: STRUCTURE-AWARE POST-TRAINING FOR TABLE RECOGNITION

Anonymous authors
Paper under double-blind review

ABSTRACT

Table recognition can preserve cell text while assigning it to the wrong headers through erroneous row or column spans. This paper studies whether post-training can improve structural fidelity in PaddleOCR-VL 1.6 without changing its architecture. TABFIT (Table Fidelity Tuning) combines coverage-balanced supervised fine-tuning, preference pairs filtered for structural improvement without lower content-aware similarity, and online optimization with verifiable table-structure rewards. The selected checkpoint improves structure-only TEDS (TEDS-S) across all primary public cohorts. On a slice defined by rare ground-truth span geometry, both tree structure and exact merged-rectangle recovery improve, with the latter showing the larger absolute gain. These results describe the evaluated sweep: the public cohorts also selected the checkpoint, several transfer cohorts regress, and some training records have unresolved upstream provenance. The findings support span-focused adaptation in this setting while leaving independent selection and broader transfer open.

1 INTRODUCTION

Modern document parsers generate structured token sequences directly from images. Systems such as GOT-OCR, Dolphin, HunyuanOCR, and PaddleOCR-VL integrate text, formulas, layout, and tables within a common vision–language interface (Ye et al., 2024; Wang et al., 2024; HV Team, 2025; Sun et al., 2025). For table recognition, HTML and Optimized Table Structure Language (OTSL) make the output directly usable by downstream software (Nassar et al., 2023). However, accurate transcription does not guarantee accurate structure: a missing rowspan, an extended colspan, or a duplicated row can associate otherwise correct text with the wrong header.

This distinction matters to document retrieval-augmented generation (RAG): parsed tables may be retrieved as evidence, but a shifted span can attach a correctly transcribed value to the wrong row or column header. A downstream answer may then inherit the wrong relation. Similar structure errors can affect spreadsheet conversion, database ingestion, and extraction from financial or scientific documents. These are motivations rather than downstream outcomes measured here. Content-aware tree similarity alone does not isolate the topology failure. The evaluation therefore pairs structure-only TEDS (TEDS-S) with exact merged-cell geometry, while retaining content-aware TEDS for joint text and structure quality (Zhong et al., 2019).

The evaluation examines structural post-training at two resolutions: tree topology across complete tables and exact recovery of merged-cell anchors and spans. The evaluation holds sample identities, decoding, and failure handling fixed across checkpoints, then examines uncommon ground-truth merge geometries and external transfer cohorts. This design also exposes two limits on interpretation: the public cohorts are used for checkpoint selection, and some training records lack verified upstream partition provenance.

To target the topology errors in Figure 1, this study develops TABFIT (Table Fidelity Tuning), a three-stage post-training framework for PaddleOCR-VL 1.6. The name also denotes the selected final checkpoint. Supervised fine-tuning (SFT) establishes source and complexity coverage; direct preference optimization (DPO) favors naturally sampled outputs with better structure and non-decreasing content-aware similarity; group-relative policy optimization (GRPO) then optimizes a verifiable reward dominated by TEDS-S. Evaluation fixes sample identities and decoding settings

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

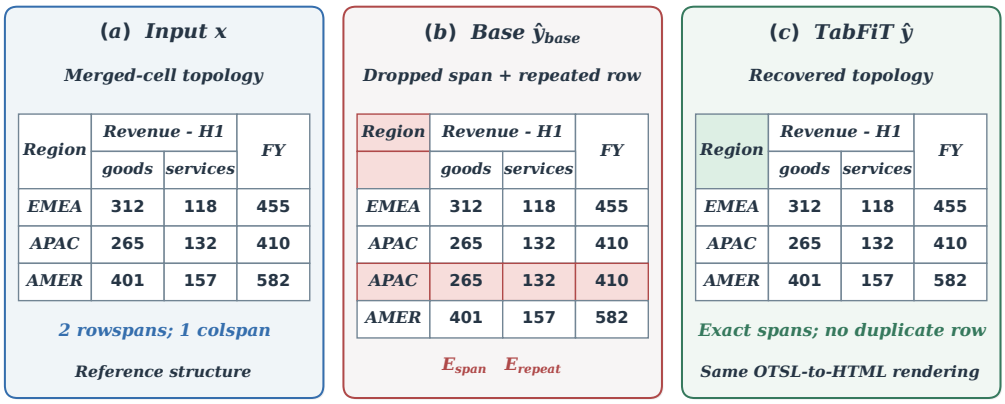


Figure 1: Structural errors and their intended correction by TABFIT. Given input x, the base output $\hat{y}_{\text{base}}$ drops a rowspan (E_{span}) and repeats a row (E_{repeat}), whereas the intended output $\hat{y}$ recovers the topology. Panels are schematic, rather than measured predictions from a held-out example.

across the base and post-trained configurations, and assigns zero evaluation credit to empty or malformed predictions.

The best checkpoint in this sweep, TABFIT, improves both TEDS and TEDS-S on all four primary public cohorts. Its gains also extend to a rare-span slice defined from ground truth alone, rather than from the base model's errors. However, it regresses on several transfer cohorts. The selected route also contains 167 HiTab records whose upstream partition could not be verified. These limitations and the use of public scores for checkpoint selection bound the interpretation of the measured improvement.

The study makes three contributions:

- An evaluation protocol separates global tree fidelity, exact merged-cell recovery, and serialization legality under fixed sample denominators.
- A three-stage post-training recipe uses coverage sampling, preference filtering, and verifiable rewards to target complementary aspects of structural recognition.
- An empirical evaluation reports public-cohort and rare-span improvements alongside route comparisons, transfer regressions, and explicit limits arising from checkpoint selection and unresolved training-data provenance.

2 RELATED WORK

2.1 DEVELOPMENT OF MULTIMODAL LARGE LANGUAGE MODELS

Multimodal large language models (MLLMs) have developed from visual instruction following toward richer visual representations, longer mixed modal contexts, and broader generation tasks. LLaVA established a widely used instruction-tuning interface (Liu et al., 2023); Qwen3-VL combines multi-level visual features with long interleaved image, video, and text contexts (Bai et al., 2025). InternVL3.5 couples post-training with adaptive visual-token allocation, while InternVL-U extends a unified model across understanding, reasoning, generation, and editing (Wang et al., 2025; Tian et al., 2026). These advances expand general multimodal capability, but their aggregate benchmarks do not measure whether a generated table assigns cells to the correct merged headers. The PaddleOCR-VL series follows the same vision-language modeling paradigm in a document-specialized, compact form: the original 0.9B model targets multilingual parsing, version 1.5 strengthens robustness and task coverage, and version 1.6 refines under-optimized regions through curated data and progressive post-training (Sun et al., 2025; Cui et al., 2026; Zhang et al., 2026b).

TABFIT retains the 1.6 architecture and targets exact table topology with established LoRA, DPO, and GRPO tools (Hu et al., 2021; Rafailov et al., 2023; Shao et al., 2024). Its contribution is a table-specific adaptation and evaluation, not a new general MLLM architecture or a claim that staged post-training itself is novel.

2.2 VISION-LANGUAGE MODELS FOR DOCUMENT PARSING

Document-parsing VLMs increasingly produce structured sequences directly from images. Donut and Pix2Struct are early image-to-sequence examples (Kim et al., 2022; Lee et al., 2023). Recent systems tackle different bottlenecks: MinerU2.5 separates page-level layout analysis from native-resolution crop recognition; Dolphin-v2 varies its parsing route by document type; and DeepSeek-OCR 2 changes visual-token ordering before decoding (Niu et al., 2025; Feng et al., 2026; Wei et al., 2026). PaddleOCR-VL-1.6 provides the compact document-parsing backbone used here (Zhang et al., 2026b).

Structural errors persist beyond any one parser. TableParseMap audits MinerU2.5-Pro, PaddleOCR-VL-1.6, GLM-OCR, and a general-purpose Qwen VLM, finding recurrent omissions, misalignment, duplication, and incorrect merges on selected difficult tables (Xiao et al., 2026). ParseBench likewise reports imperfect table fidelity for both general VLMs and specialized parsers (Zhang et al., 2026a). These diagnostics motivate targeted structure evaluation, but do not estimate the prevalence of each error on this paper's test sets. PaddleOCR-VL-1.6 offers a strong, compact starting point with an existing staged adaptation route; holding its backbone fixed lets TABFIT test topology-focused data, preferences, and rewards for exact span recovery.

For table output, OTSL provides a compact structural language and PubTabNet's TEDS evaluates canonical table trees (Nassar et al., 2023; Zhong et al., 2019). The evaluation pairs content-aware TEDS with structure-only TEDS-S and exact merged-rectangle F1: tree similarity and correct span anchors answer related but distinct questions. TRivia takes a different supervision route, using response consistency and generated visual questions to adapt on unlabeled table images with a proxy GRPO reward (Zhang et al., 2025). TABFIT instead requires labeled OTSL targets and scores sampled outputs against their reference structure. This contrast locates the method's dependency on annotations and the property its reward can verify.

3 METHODOLOGY

3.1 PRELIMINARY

Given a table image x, a model p_θ emits an OTSL sequence y, with ground-truth sequence y^*. The goal is to recover the table's cell topology while preserving its contents. The notation $T(y)$ and $S(y)$ denotes content-aware TEDS and structure-only TEDS-S against y^*, respectively; both lie in $[0, 1]$ during training and are reported as percentages in tables. Prediction and ground truth are canonicalized into table trees to compute

$$\mathrm{TEDS}(y, y^*) = \max\left(0, 1 - \frac{d_{\mathrm{tree}}(y, y^*)}{\max(|y|, |y^*|)}\right). \tag{1}$$

Here $|y|$ and $|y^*|$ count nodes in the canonical table trees. TEDS-S applies the same computation after discarding text nodes. It is the primary score because it isolates topology. Each merged cell is also represented as $(r_0, c_0, \mathrm{rowspan}, \mathrm{colspan})$ for exact set-F1 computation. A match requires both anchor and span dimensions to agree. Strict legality checks reject empty tables, invalid OTSL transitions, ragged grids, and non-renderable structures. Empty or malformed predictions receive zero and remain in the denominator. Fifteen metric fixtures, including official TEDS-equivalence cases and oversize predictions, pass before every run.

Exact and soft merge scores. These diagnostics answer different questions and should not be confused with output-format validity. Let P and G be the predicted and reference sets of merged rectangles. The implementation name `format_exact` denotes

$$F_{\mathrm{exact}} = \frac{2|P \cap G|}{|P| + |G|}, \tag{2}$$

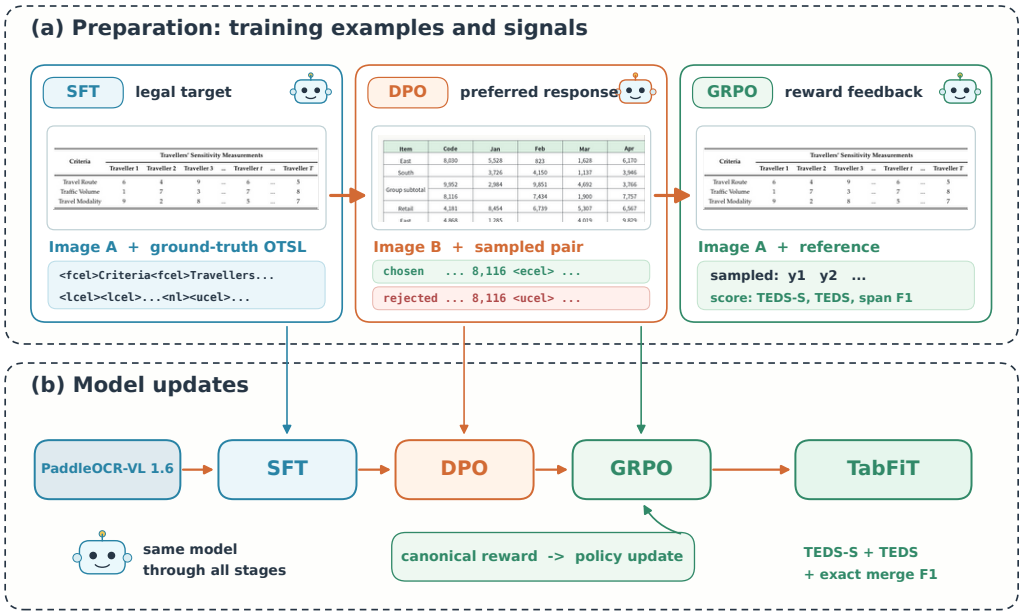


Figure 2: An example-led overview of TABFIT. (a) Two training images illustrate the different supervision consumed by SFT, DPO, and GRPO. The displayed OTSL fragments and chosen/rejected token difference are from the corresponding training records; the GRPO outputs are schematic. (b) The same model passes through supervised, preference, and online reward updates. GRPO scores sampled responses against labeled ground-truth structure using TEDS, TEDS-S, and exact merged-rectangle F1. The preparation/feedback presentation is informed by Zhang et al. (2025); unlike its unlabeled QA reward, this route uses labeled tables and canonical structural scores.

where two rectangles match only when anchor row, anchor column, rowspan, and colspan all agree. A correct 2×3 merge shifted by one column therefore receives no exact credit. The diagnostic `format_soft` gives partial credit to near misses. Pairwise rectangle IoU determines a maximum-weight one-to-one assignment M^* obtained with the Hungarian algorithm, giving

$$F_{\text{soft}} = \frac{2 \sum_{(p,g) \in M^*} \text{IoU}(p, g)}{|P| + |G|}. \tag{3}$$

Exact F1 is the reporting criterion because the downstream relation is either correct or incorrect; soft F1 is only a dense training signal in Q2. Neither score means approximately valid syntax: strict OTSL legality is checked separately.

The same structural scores also guide post-training. Figure 2 previews the supervision used in its three stages; the training subsections below specify the updates.

3.2 STRUCTURAL HALLUCINATION TAXONOMY

Figure 3 distinguishes four structural error types. An insertion creates a cell or column unsupported by the image; an omission drops a visible cell; a span mismatch assigns the wrong rowspan or colspan; and repetition copies a row, column, or tail segment. These categories expose different failure modes that aggregate TEDS-S can obscure. A missing cell and a shifted rowspan may incur similar tree-edit costs while altering different header relationships.

The metrics provide complementary views of these errors: TEDS measures content and tree structure, TEDS-S isolates the tree, and exact rectangle F1 tests span anchors and dimensions. Serialization validity is assessed separately before canonical scoring. The GRPO reward uses the fixed combination in Equation 5, together with a penalty for invalid serialization.

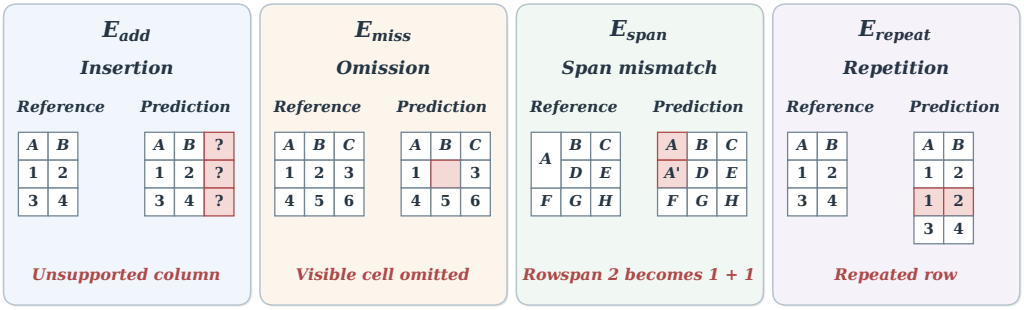


Figure 3: Four structural error types: insertion, omission, span mismatch, and repetition. Highlighted cells indicate the discrepancy from ground truth.

3.3 Expansion and split isolation

The SFT mixture covers multiple public sources; GRPO expands source coverage and stratifies prompts by merge geometry. Section 4 reports the realized source and complexity quotas. No exact image or ground-truth hash collision is reported between available training and evaluation manifests; this does not establish upstream partition provenance.

Lineage disclosure. The audit records source identifiers, available upstream split metadata, image and target digests, and local manifest digests. Image and target hashes are compared against locked evaluation cohorts before training. This two-level check matters because filenames can change when archives are repacked, while target identity can reveal copied annotations even if images were re-encoded. The later audit could not confirm the upstream partition of 167 HiTab records in the selected route. Accuracy remains the selection criterion, but this route is therefore not described as provenance-clean. The limitation is attached to every table in which the selected checkpoint appears.

Post-training route. Figure 2 summarizes TabFiT. The base→SFT→DPO→GRPO route uses one model architecture and three training distributions. The stages address different optimization needs: SFT supplies coverage of valid outputs, DPO ranks realistic model errors, and GRPO assigns online credit using ground-truth structure. This is a design rationale; the route comparisons below do not isolate the causal contribution of every stage.

3.4 Stage 1: coverage-balanced SFT

The SFT pool requires an upstream training split, an existing image, a parseable ground truth, and both original TEDS fields. Image bytes and canonical targets are checked against validation and test data. Exact image–target duplicates and repeated targets are collapsed, and repeated tables from one source document are capped. Complexity is computed only from ground-truth geometry: a table is complex when it is long, wide, merge-dense, has a deep header, a large span, or a two-dimensional merge. Stable seeded sampling balances simple, medium, and complex examples under source caps; it never selects by base-model error. Eligible targets are converted to strictly legal OTSL.

3.5 Stage 2: structure-preserving DPO

The SFT checkpoint samples natural candidates from a complexity-skewed training prompt pool. A candidate pair is retained only if

$$S(y^+) - S(y^-) \geq 0.03, \qquad T(y^+) \geq T(y^-) - 10^{-6}, \tag{4}$$

where S and T are defined in Section 3.1. Thus, a preferred candidate must improve structure without reducing content-aware TEDS beyond numerical tolerance. Because TEDS combines text and structure, this filter does not guarantee that every cell's transcription is preserved. The highest- and lowest-structure natural samples form at most two pairs per prompt; ground truth is never inserted as a synthetic chosen response, and illegal or truncated generations can only be rejected.

3.6 STAGE 3: SOURCE-BALANCED GRPO

The GRPO prompt pool contains unique images and is quota-balanced across sources and ground-truth complexity. This prevents the largest corpus or easy no-merge tables from dominating online optimization. For each prompt GRPO samples multiple completions. Let T denote TEDS, S TEDS-S, and F_{exact} exact merged-rectangle F1. The selected Q1 reward is

$$r(y, y^*) = \begin{cases} -1, & \text{if } y \text{ is empty, unparsable, or strictly illegal,} \\ 0.20T + 0.80S, & \text{if } y \text{ is legal and } |G| = 0, \\ 0.15T + 0.75S + 0.10F_{\text{exact}}, & \text{if } y \text{ is legal and } |G| > 0, \end{cases} \tag{5}$$

where G is the reference merged-rectangle set defined above. Missing or illegal ground truth instead raises an infrastructure error and aborts the run. For the K completions of one prompt, GRPO replaces a learned value baseline with the group's own rewards (Shao et al., 2024):

$$\bar{r} = \frac{1}{K}\sum_{j=1}^{K} r_j, \qquad \hat{A}_i = \frac{r_i - \bar{r}}{\sqrt{K^{-1}\sum_{j=1}^{K}(r_j - \bar{r})^2 + \epsilon}}. \tag{6}$$

Here $r_i = r(y_i, y^*)$, and ϵ stabilizes constant-reward groups. The group-relative advantage enters the clipped policy update, with a KL penalty to the frozen reference policy; the implementation uses $K = 4$.

The reward remains deliberately limited. TEDS-S may score a nearly correct tree whose important merge is misplaced, exact F1 is discontinuous, and a group of identical errors provides GRPO no positive direction. Combining TEDS, TEDS-S, and exact F1 reduces but does not eliminate objective misspecification. A single checkpoint is selected globally rather than separately for each dataset: TABFIT maximizes public macro TEDS-S and also improves macro TEDS and both pre-specified rare-span criteria.

4 EXPERIMENTS

4.1 FIXED PUBLIC EVALUATION

Table 1 defines the evaluation denominator. Four public clean tests form the primary macro score. "Clean" means that a small, pre-frozen set of exceptionally long examples is excluded uniformly for every checkpoint; no additional length filtering is performed. The same sample IDs are scored for every model. Four public transfer tests bring the total to 170,909 examples.

Rare-span slice. To test uncommon topology without selecting by model error, the evaluation uses a frozen GT-only slice. Each merge is encoded by $(\text{rowspan}, \text{colspan}, q_r, q_c)$, where q_r and q_c are row and column anchor quintiles. A signature occurring at most 100 times within a source is rare; a table enters the slice if it contains at least one rare signature. The resulting slice contains 17,959 examples. Metrics are first averaged within each source and then equally across the four sources, so PubTables-1M cannot dominate the result. Its TEDS-S and exact F1 separate tree-level improvement from correct span placement; the pre-specified check requires both to improve by at least one percentage point.

4.2 IMPLEMENTATION

All models start from PaddleOCR-VL 1.6 (0.9B parameters) (Zhang et al., 2026b). Training uses four GPUs, bfloat16, gradient checkpointing, and frozen vision parameters unless stated otherwise.

Training sets and sampling. SFT retains at most ten tables per source document, targets 5,000 simple, 6,000 medium, and 9,000 complex examples, and caps a source at 25% before shortage relaxation. After filtering and quota filling, the realized set has 20,651 training and 512 development examples. DPO samples six candidates per prompt at temperature 0.7 and top-p 0.95 from a pool targeting 5,000 complex, 2,000 medium, and 1,000 simple prompts. The filter in Equation 4 retains 8,383 pairs: 5,994 complex, 2,038 medium, and 351 simple, including 652 with an illegal rejected

Table 1: Fixed evaluation cohorts. The primary public macro is the equal-source mean over the first four rows.

Cohort	Samples	Role
FinTabNet clean	10,197	public test
PubTables-1M clean	93,656	public test
PubTables-v2 clean	7,663	public test
SynthTabNet clean	43,433	public test
PubTabNet/ICDAR	9,064	transfer test
SPRINT-MUSTARD	1,429	transfer test
TableVerse-5K	4,955	transfer test
OmniDocBench tables	512	transfer test
Total	170,909	fixed per checkpoint

response. The GRPO pool has 8,000 unique prompts: 1,400 each from FinTabNet and PubTabNet, 1,300 from PubTables-1M, 1,100 from PubTables-v2, 800 each from SynthTabNet and ControlledRender, and 400 each from HiTab, SciTSR, and TABLET. Its ground-truth strata contain 4,635 complex, 1,930 medium, and 1,435 simple prompts.

Optimization and inference. SFT trains rank-32 LoRA for one epoch at learning rate 2×10^{-5}. DPO uses rank-32 LoRA, $\beta = 0.1$, learning rate 2×10^{-6}, and 0.5 epoch. GRPO samples four completions per prompt and updates rank-32 LoRA for 100 steps at learning rate 2×10^{-7}, KL coefficient 0.02, and maximum input/output lengths 12,288/6,144. Decoding and scoring settings are fixed across checkpoint comparisons.

Checkpoint selection and scope. The headline checkpoint maximizes public macro TEDS-S over 24 evaluated configurations, including the base. The same checkpoint is reported across all cohorts. This avoids per-dataset oracle selection, but the primary public cohorts are still used for both selection and reporting. Their scores are therefore descriptive results of the sweep, not an unbiased estimate from an untouched test set. A stronger evaluation would select on separate development data and assess the chosen checkpoint once on a held-out test set. All results marked † additionally share the unresolved 167-record HiTab provenance limitation.

4.3 Global structural fidelity

Table 2 gives per-source results and published reference scores. TabFiT improves both TEDS and TEDS-S on the four primary public tests, with the largest structural gains on SynthTabNet and PubTables-v2. PubTabNet/ICDAR is shown separately because it is a transfer test and regresses. The published systems use their authors' data and scoring protocols, so their numbers contextualize table-recognition performance without forming a paired ranking against this study's fixed cohorts.

The equal-source macro TEDS-S gain is 3.2468 points. All 154,949 examples, including empty and malformed generations, remain in the denominator. Figure 4 shows the public and rare-span results.

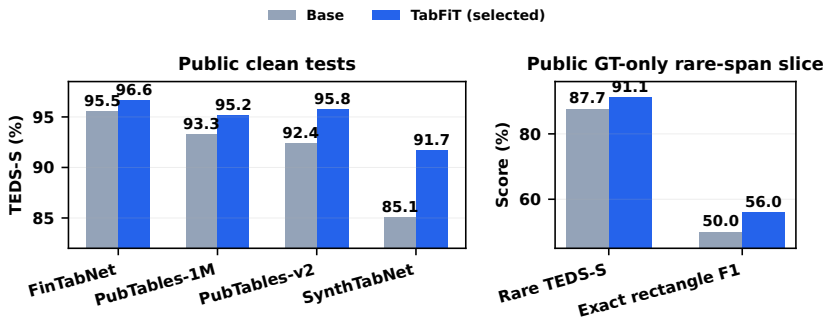


Figure 4: TEDS-S on four public clean tests (left) and the public GT-only rare-span slice (right). The same accuracy-selected TABFIT checkpoint is compared with the base in every cohort.

The public plot summarizes the same fixed-denominator comparison reported below. Its source-level spread matters: strong performance on FinTabNet alone would not explain the larger gains on SynthTabNet and PubTables-v2.

Table 2: Public TEDS/TEDS-S (percent). The fixed-cohort macro covers the first four sources; published scores use different protocols. †Unresolved provenance for 167 HiTab training records.

Cohort	Base	TABFIT†
FinTabNet	94.6719/95.5369	**95.6984/96.6438**
PubTables-1M	91.0400/93.2523	**94.0222/95.1577**
PubTables-v2	90.1274/92.3650	**94.7846/95.7662**
SynthTabNet	81.1677/85.0966	**89.6940/91.6701**
Macro	89.2518/91.5627	**93.5498/94.8095**
PubTabNet/ICDAR test	92.7225/94.4282	92.4283/93.8664
Published references: model (cohort)		*TEDS / TEDS-S*
TFLOP (Khang & Hong, 2024) (FinTabNet)		99.45/99.56
TFLOP (PubTabNet test)		96.66/98.38
TFLOP (SynthTabNet)		99.40/99.42
Dolphin (Qin et al., 2026) (PubTables-1M)		95.48/96.82
TDATR (Qin et al., 2026) (PubTables-1M)		97.97/98.39

4.4 RARE SPANS AND TRANSFER

The frozen rare slice tests the target failure mode at finer granularity. Table 3 shows that TABFIT clears both one-point criteria: TEDS-S rises by 3.4280 points and exact merged-rectangle F1 by 6.0088 points. The larger gain in rectangle F1 indicates that the update improves exact anchor/span recovery beyond what is visible in average tree similarity.

Table 3: Public rare-span macro (17,959 examples; percent). †Unresolved provenance for 167 HiTab training records.

Checkpoint	TEDS	TEDS-S	Rect. F1	ΔS/ΔF1
Base	84.9233	87.6986	50.0157	–
TABFIT†	**89.3797**	**91.1266**	**56.0245**	**+3.4280/+6.0088**

To resolve which uncommon geometries improve, Table 5 partitions the frozen slice using ground truth alone. Categories may overlap: for example, one table can contain both a rare rowspan and a two-dimensional merge. Scores are averaged within source and then across represented sources, as in the headline rare result.

All five overlapping subsets improve on both metrics (Appendix Table 5). The size of the gain varies by geometry: rare rowspans rise by 4.98 TEDS-S points and 12.34 exact-F1 points, while rare colspans rise by 3.54 and 2.17 points. Thus, a better tree score does not imply an equally large gain in exact span placement. These overlapping subsets are descriptive breakdowns, not independent tests (Figure 5); they do not identify which training stage corrected each case.

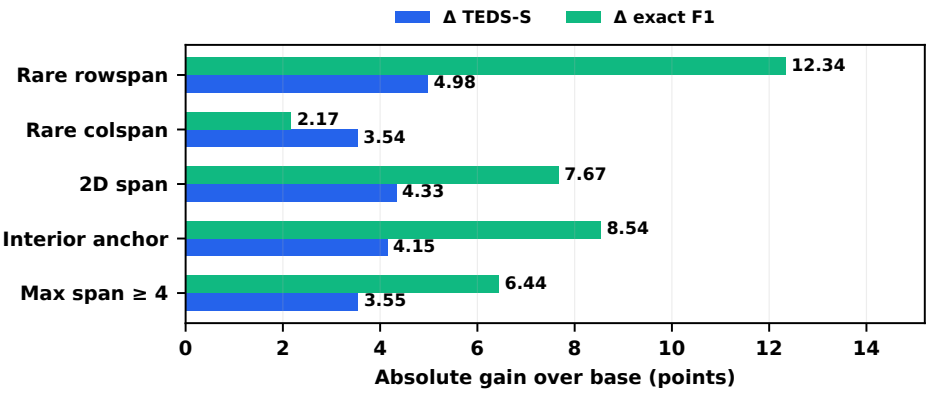


Figure 5: Absolute gains of the selected TABFIT checkpoint over the base on GT-defined rare-geometry subsets. Categories overlap.

Transfer is mixed. On SPRINT-MUSTARD, TABFIT obtains 21.9718/87.3767; the large TEDS–TEDS-S gap is consistent with a content or normalization mismatch, but the aggregate scores alone do not identify its cause. On PubTabNet/ICDAR it scores 92.4283/93.8664, below the base 92.7225/94.4282. On TableVerse-5K it obtains 84.9792/90.4380, below the base 85.9554/90.9021 on both metrics. OmniDocBench tables score 94.3323/97.1659 versus 95.7142/97.8760 for the base. These results delimit the claim: the selected route improves the primary public macro and rare topology, but does not uniformly improve every external domain.

Cross-system reference comparisons are reported in Supplementary Section A.1; their differing protocols preclude a controlled ranking.

4.5 WHAT ROUTE COMPARISONS ESTABLISH

The selected route's intermediate SFT and DPO checkpoints were not evaluated on the same cohorts, preventing stage attribution. The strongest audited-data alternative, Strict-LoRA, reaches 92.3335 public and 89.0826 rare-span TEDS-S, below the selected route's 94.8095 and 91.1266 (Appendix A.2). These configurations differ in several choices and cannot isolate provenance. The separate full-SFT branch changes modestly and non-monotonically (Figure 6; Appendix A.3); its deltas cannot substitute for selected-route stage results. Two eval10k examples illustrate error types but lack audited checkpoint lineage (Appendix A.5).

The public macro serves as both selection signal and headline comparison, so its gain is not an untouched-test estimate. The GT-only rare slice shares its sources, avoiding selection by base-model errors but not forming an independent holdout. PubTabNet/ICDAR, TableVerse-5K, and OmniDocBench regress, limiting transfer beyond the primary public cohorts.

5 CONCLUSION

TABFIT combines SFT, DPO, and GRPO for structural table recognition. The selected route raises public macro TEDS/TEDS-S from 89.2518/91.5627 to 93.5498/94.8095, with rare-span TEDS-S and exact rectangle F1 gains of 3.4280 and 6.0088 points. Their different behavior on rare spans motivates reporting both global tree fidelity and exact merge recovery. Public-score selection and mixed transfer bound this finding to the evaluated protocol.

REFERENCES

Shuai Bai, Yuxuan Cai, Ruizhe Chen, et al. Qwen3-VL Technical Report. *arXiv preprint arXiv:2511.21631*, 2025. URL https://arxiv.org/abs/2511.21631.

Cheng Cui, Ting Sun, Suyin Liang, et al. PaddleOCR-VL-1.5: Towards a Multi-Task 0.9B VLM for Robust In-the-Wild Document Parsing. *arXiv preprint arXiv:2601.21957*, 2026. URL https://arxiv.org/abs/2601.21957.

Hao Feng, Wei Shi, Ke Zhang, et al. Dolphin-v2: Universal Document Parsing via Scalable Anchor Prompting. *arXiv preprint arXiv:2602.05384*, 2026. URL https://arxiv.org/abs/2602.05384.

Edward J Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. Lora: Low-rank adaptation of large language models. *arXiv preprint arXiv:2106.09685*, 2021.

HV Team. Hunyuanocr technical report. *arXiv preprint arXiv:2511.19575*, 2025.

Minsoo Khang and Teakgyu Hong. TFLOP: Table structure recognition framework with layout pointer mechanism. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence*, 2024. URL https://www.ijcai.org/proceedings/2024/0105.pdf.

Geewook Kim, Teakgyu Hong, Moonbin Yim, JeongYeon Nam, Jinyoung Park, Jinyeong Yim, Wonseok Hwang, Sangdoo Yun, Dongyoon Han, and Seunghyun Park. Ocr-free document understanding transformer. In *European Conference on Computer Vision (ECCV)*, 2022.

Kenton Lee, Mandar Joshi, Iulia Turc, Hexiang Hu, Fangyu Liu, Julian Eisenschlos, Urvashi Khandelwal, Peter Shaw, Ming-Wei Chang, and Kristina Toutanova. Pix2struct: Screenshot parsing as pretraining for visual language understanding. In *Proceedings of the International Conference on Machine Learning (ICML)*, 2023.

G. Li et al. Hunyuanocr-1.5: Making lightweight ocr vlms faster and better. *arXiv preprint arXiv:2607.04884*, 2026.

Haotian Liu, Chunyuan Li, Qingyang Wu, and Yong Jae Lee. Visual instruction tuning. In *Advances in Neural Information Processing Systems*, 2023.

Islam Nassar, Austin P Darrow, Rohit Smith, and Andreas Savakis. Optimized table tokenization for table structure recognition. In *International Conference on Document Analysis and Recognition (ICDAR)*, 2023.

Jun Niu, Zhuangcheng Gu, et al. Mineru2.5: A decoupled vision-language model for efficient high-resolution document parsing. *arXiv preprint arXiv:2509.22186*, 2025.

Chunxia Qin, Chenyu Liu, Pengcheng Xia, Jun Du, Baocai Yin, Bing Yin, and Cong Liu. Tdatr: Improving end-to-end table recognition via table detail-aware learning and cell-level visual alignment. *arXiv preprint arXiv:2603.22819*, 2026.

Rafael Rafailov, Archit Sharma, Eric Mitchell, Christopher D Manning, Stefano Ermon, and Chelsea Finn. Direct preference optimization: Your language model is secretly a reward model. In *Advances in Neural Information Processing Systems*, 2023.

Zhihong Shao, Peiyi Wang, Qihao Zhu, Runxin Xu, Junxiao Song, Xiao Bi, Haowei Zhang, Mingchuan Zhang, YK Li, Y Wu, et al. Deepseekmath: Pushing the limits of mathematical reasoning in open language models. *arXiv preprint arXiv:2402.03300*, 2024.

Ting Sun, Suyin Liang, Tingquan Gao, Zelun Zhang, et al. Paddleocr-vl: Boosting multilingual document parsing via a 0.9b ultra-compact vision-language model. *arXiv preprint arXiv:2510.14528*, 2025.

Changyao Tian, Danni Yang, Guanzhou Chen, et al. InternVL-U: Democratizing Unified Multimodal Models for Understanding, Reasoning, Generation and Editing. *arXiv preprint arXiv:2603.09877*, 2026. URL https://arxiv.org/abs/2603.09877.

Bin Wang, Zhibo Yu, Haidong Tang, Yuxin Lai, Zeming Lu, Chengjin Yang, Xianyu Wang, Jian Li, Jun Xie, et al. Dolphin: Document image parsing via heterogeneous anchor network. *arXiv preprint arXiv:2412.03671*, 2024.

Weiyun Wang, Zhangwei Gao, Lixin Gu, et al. InternVL3.5: Advancing Open-Source Multimodal Models in Versatility, Reasoning, and Efficiency. *arXiv preprint arXiv:2508.18265*, 2025. URL https://arxiv.org/abs/2508.18265.

Haoran Wei, Yaofeng Sun, and Yukun Li. DeepSeek-OCR 2: Visual Causal Flow. *arXiv preprint arXiv:2601.20552*, 2026. URL https://arxiv.org/abs/2601.20552.

Jutao Xiao, Yuan Qu, Dongsheng Ma, Fan Wu, Tianyao He, et al. From diagnosis to correction: Benchmarking and improving real-world table parsing. *arXiv preprint arXiv:2608.09842*, 2026. URL https://arxiv.org/abs/2608.09842.

Haibin Ye, Jiaxin Liu, Guanheng Liu, Ming Qiao, Jin Zhang, Xing Wu, Shuai Zhang, Lix Ning, Jie Zhou, Wei Zhang, et al. Got-ocr 2.0: A general ocr theory towards ocr-2.0 via generative vision-geometry pre-training. *arXiv preprint arXiv:2409.01704*, 2024.

Boyang Zhang, Sebastián G. Acosta, Preston Carlson, Sacha Bron, Pierre-Loïc Doulcet, et al. Parsebench: A document parsing benchmark for ai agents. *arXiv preprint arXiv:2604.08538*, 2026a. URL https://arxiv.org/abs/2604.08538.

Junyuan Zhang, Bin Wang, Qintong Zhang, et al. Trivia: Self-supervised fine-tuning of vision-language models for table recognition. *arXiv preprint arXiv:2512.01248*, 2025.

Zelun Zhang, Hongen Liu, Suyin Liang, et al. PaddleOCR-VL-1.6: Expanding the Frontier of Document Parsing with Under-Optimized Region Refinement and Progressive Post-Training. *arXiv preprint arXiv:2606.03264*, 2026b. URL https://arxiv.org/abs/2606.03264.

Xu Zhong, Jianbin Tang, Jose Enrique Alonso-Pablo, and Jie Shang. Image-based table recognition: data, model, and evaluation. *arXiv preprint arXiv:1911.10683*, 2019.

A Supplementary Material

A.1 Comparison with reported systems

Cross-paper values differ in crops, decoding, and scorer implementation, so Table 4 is a reference-gap analysis rather than a controlled leaderboard. Under the canonical protocol TabFiT exceeds the PaddleOCR-VL base but remains below the strongest specialized references listed here.

Table 4: Reference gaps on two difficult public datasets. Values are reported TEDS/TEDS-S percentages; protocols are not identical. †Unresolved provenance for 167 HiTab training records.

System	PubTables-1M	TableVerse-5K
Dolphin (Wang et al., 2024)	95.48/96.82	–
TDATR (Qin et al., 2026)	**97.97/98.39**	–
HunyuanOCR-1.5 (Li et al., 2026)	–	79.37/86.05
StrucTab (reported)	–	**87.81/92.74**
Base, study protocol	91.0400/93.2523	85.9554/90.9021
TabFiT†, study protocol	94.0222/95.1577	84.9792/90.4380

The PubTables-1M result closes part of the gap but remains 3.23 TEDS-S points below TDATR. TableVerse is more resistant: the selected post-training route does not improve TEDS-S over the base. Visual-domain and serialization differences are possible explanations, but distinguishing them requires further controlled experiments. These cross-paper comparisons contextualize the scores; they do not establish a controlled ranking of systems.

Rare-geometry detail. The GT-defined rare-span predicates overlap. The counts and source-macro scores for each subset appear in Table 5.

Table 5: Rare-geometry breakdown. Each metric cell is Base→TABFIT (percent); N counts unique tables satisfying the GT-only predicate. The selected route has 167 HiTab records with unresolved provenance.

GT-only subset	N	TEDS-S	Exact rect. F1
Rare rowspan-only	6,317	84.61→89.59	43.37→55.70
Rare colspan-only	4,945	87.32→90.86	48.03→50.20
Rare 2D span	10,464	83.69→88.03	31.58→39.25
Interior anchor	8,410	84.98→89.13	50.29→58.83
Maximum span ≥ 4	15,992	87.97→91.52	51.32→57.76

A.2 ABLATION OVER LOWER-SCORING ROUTES

Table 6 reports alternative routes described as provenance-clean in the study's audit. They vary trainable capacity, the MMSci increment, rare-focused continuation, preference tuning, and reward profile. Because several factors differ across branches, these are configuration comparisons rather than a factorial ablation of the selected route. None exceeds the selected TABFIT checkpoint's public macro 93.5498/94.8095 or rare-span TEDS-S 91.1266.

Table 6: Supplementary route ablation. Public columns are equal-source macro averages over four clean tests; rare columns use the frozen 17,959-example public rare-span slice. All numbers are percentages.

Route	Public TEDS	Public TEDS-S	Rare TEDS-S	Rare rect. F1	PT-1M TEDS-S	TV TEDS-S
Base	89.2518	91.5627	87.6986	50.0157	93.2523	90.9021
Full SFT	89.6325	91.8010	88.2815	51.2830	93.6605	90.9336
LoRA+aligner SFT (Strict-LoRA)	90.5404	92.3335	89.0826	52.0160	94.5768	90.7899
LoRA+aligner SFT, +MMSci	90.5485	92.3303	88.9402	51.2929	94.3988	90.7953
Full SFT → rare SFT	89.6585	91.8123	88.3048	51.3363	93.6855	90.9453
→ DPO	89.6511	91.8090	88.3214	51.4819	93.6761	90.9503
→ GRPO-Q0	89.6682	91.8292	88.3060	51.2986	93.6843	90.9726
→ GRPO-Q1	89.6681	91.8267	88.2968	51.3700	93.6863	90.9337
→ GRPO-Q2	89.6652	91.8260	88.2888	51.4127	93.6871	**90.9859**

The strongest audited-data ablation is Strict-LoRA. It improves public macro TEDS/TEDS-S by 1.2886/0.7708 points and rare-span TEDS-S/exact F1 by 1.3840/2.0003 points. The 21k MMSci variant changes public TEDS/TEDS-S by only $+0.0081/-0.0032$ relative to Strict-LoRA. On the weaker full-SFT branch, rare-focused SFT, DPO, and short GRPO continuations produce only small changes. Q2 places additional weight on soft merged-rectangle overlap. Relative to its DPO parent, it changes public macro TEDS-S by only $+0.0170$ points; on the public rare-span slice it improves exact rectangle F1 over the base by 1.3970 points but TEDS-S by only 0.5903 points, below the pre-specified 1-point criterion. These results show limited gains from changing the reward profile on this branch; they do not isolate the effect of the selected route's filtering.

A.3 STAGE EVOLUTION ON THE FULL-SFT BRANCH

Figure 6 traces the full-SFT ablation branch in Table 6. Its public macro TEDS-S evolves from 91.5627 (base) to 91.8010 (full SFT), 91.8123 (rare-focused SFT), 91.8090 (DPO), and 91.8267 (GRPO-Q1). DPO slightly reduces the public score while increasing rare-span TEDS-S; GRPO-Q1 reverses part of that trade-off. These are checkpoint scores, not a per-step training curve. This branch differs from the selected TABFIT route in SFT initialization and training data, so its stage deltas cannot be attributed to the selected route.

A.4 SELF-SELECTED VALIDATION SETS

The self-selected sets are reported separately from the public cohorts. The 10k set is balanced across source and complexity; the 500-table ControlledRender rare-dev set is independently seeded and hash-disjoint from the 2,000 rendered training examples. They are used for diagnosis and regression checks, not for the headline public-test claim.

The synthetic rare-dev base is already near saturation, making a one-point improvement mathematically impossible. It therefore serves only as a non-regression check; the 17,959-example public GT-only slice supports the main rare-structure claim.

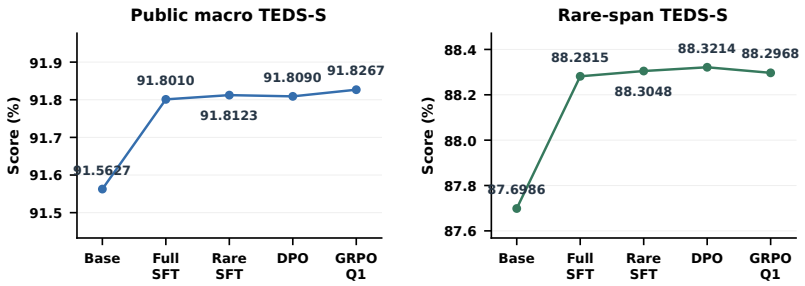


Figure 6: Full-SFT ablation trajectory: TEDS-S on fixed public and rare-span cohorts. Points denote successive checkpoints of this branch, not of the selected TABFIT route.

Table 7: Self-selected validation results (percent). †Unresolved provenance for 167 HiTab training records.

Checkpoint	eval10k TEDS	eval10k TEDS-S	rare-dev TEDS	rare-dev TEDS-S
Base	91.1809	93.2018	99.8930	99.8930
Strict-LoRA	92.0292	93.6012	99.9159	99.9159
TABFIT†	**93.7422**	**94.9333**	**100.0000**	**100.0000**

A.5 PER-SAMPLE BEHAVIOR ACROSS TRAINING STAGES

Two smaller examples from the fixed eval10k diagnostic set show residual errors after SFT and DPO. They are selected illustrations, not a prevalence estimate. The panels show the saved predictions of four sweep checkpoints; their lineage relative to the selected TABFIT route is unaudited. Scores are per-sample TEDS/TEDS-S in percent.

Case A — continuing merge correction. On a 12×11 SynthTabNet validation table, TEDS-S progresses from 87.80 (base) to 89.92 (SFT), 91.45 (DPO), and 91.60 (GRPO). TEDS follows 83.65→88.94→90.46→90.62. Exact anchor mismatches are not monotonic (17, 12, 9, 10), showing why tree scores alone cannot establish exact merge recovery.

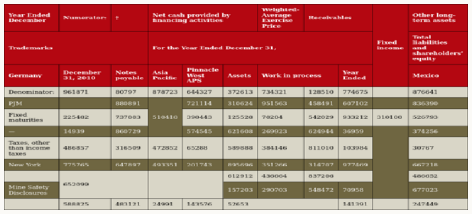
Case B — residual error after a DPO plateau. On PubTables-1M::val::PMC5732969_table_2 (ground truth 8×9), the base output has a 9×10 grid. SFT and DPO retain that grid and score 83.18/84.44 after a base score of 78.95/80.00. GRPO yields a 9×9 grid and 90.26/91.57; mismatching anchors fall from 25 at base to 22 after SFT/DPO and 14 after GRPO. The saved outputs, scores, and image/GT hashes are in `evidence/qualitative_cases.jsonl`.

Figure 7 visualizes Case A and Figure 8 visualizes Case B.

Case A — continued merge recovery

full crop · full predicted tables · unaudited sweep checkpoints · eval10k

Input image (full crop)
GT grid 12×11 · complex



SynthTabNet::val::image_035546_1634635960.878361

Ground truth
12×11 · full GT

ground-truth topology and values

Base
12×11 · 83.65/87.80

17 mismatching anchors

SFT
12×11 · 88.94/89.92

12 mismatching anchors

DPO
12×11 · 90.46/91.45

9 mismatching anchors

GRPO
12×11 · 90.62/91.60

10 mismatching anchors

Figure 7: Case A: 12×11 SynthTabNet table (`image_035546_1634635960.878361`) from fixed eval10k. Saved Base/SFT/DPO/GRPO predictions show rising TEDS-S (87.80/89.92/91.45/91.60), while exact anchor mismatches are 17/12/9/10. Values are clipped only for display. Checkpoint lineage is unaudited.

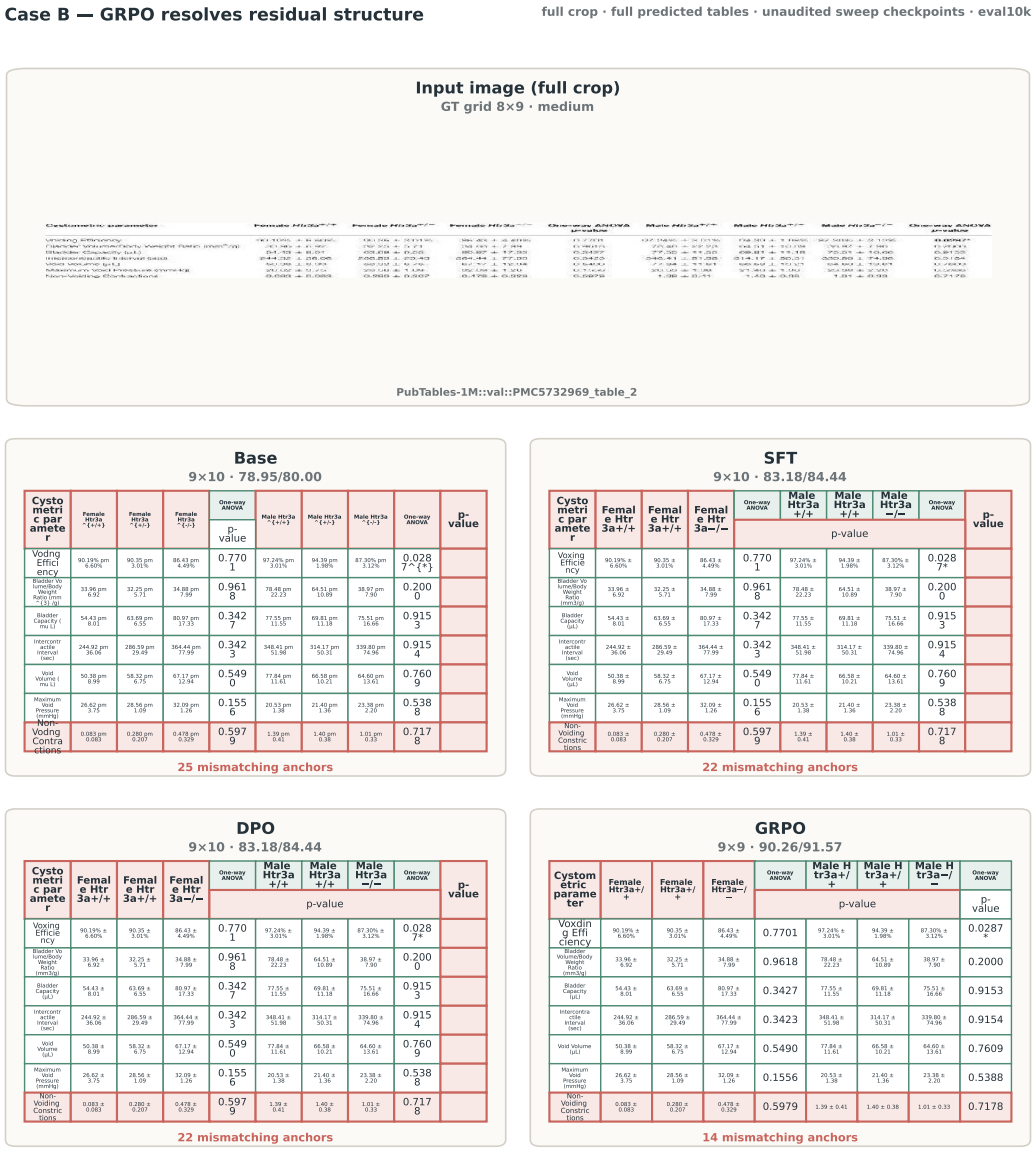


Figure 8: Case B: 8×9 PubTables-1M table (`PMC5732969_table_2`) from fixed eval10k. TEDS-S rises from 80.00 at base to 84.44 at SFT, remains 84.44 at DPO, and reaches 91.57 at GRPO. This is a genuine DPO plateau followed by a further GRPO gain; the sweep-checkpoint lineage is unaudited.