Financial Table Extraction Pipeline

End-to-End Architecture with Table Transformer

INPUT

Financial

Document

(PNG/PDF)

Image

STEP 1

Table Detection

Table Transformer

microsoft/table-transformer-detection

Time: 15.4s   mAP 95%

BBox

STEP 2

Structure Recognition

Table Transformer

microsoft/table-transformer-tsr

Time: 1.9s   TEDS 89%

Grid

STEP 3

OCR Extraction

EasyOCR

GPU-accelerated

Time: 7.2s   Conf 88%

Cells

STEP 4

Numeric Normal.

Rule-based

Pattern matching

Time: <1ms   94.9%

Values

STEP 5

Semantic Class.

Rule-based

Position + content

Time: <1ms   ~95%

Types

STEP 6

Table QA

Fuzzy Lookup

Row/Col matching

Time: 14ms   30.0%

Answer

Detected Table

Structure

OCR Grid

BBox [x1,y1,x2,y2]   36 rows x 7 cols   107 text cells

conf: 99.99%   grid[row][col]

Pipeline Summary

Total Time: 24.5 seconds

Detection Model: Table Transformer

OCR Engine: EasyOCR

Final QA Accuracy: 30.0%

Test Type: FAIR (no GT peek)

Error Cascade

Step 1: Detection   cell bboxes   **99.99%**

Step 2: TSR   **~95%**

Step 3: OCR   **54% (CER)**   ← Bottleneck

Step 4: Numeric   **94.9%**

Step 5: Semantic   **~95%**

Step 6: QA   **30.0%**

QA Example

Question:

"Property, plant and equipment"

for "Group 2025"

Lookup: row=5, col=1

Expected: 1,511,947

Predicted: 1,511,947 [OK]

OUTPUT

Answer

"1,511,947"