Table Extraction Tools

1. Tabula: Done
2. Camelot: Done
3. Pdfplumber: Done
4. PyMuPDF:
5. Img2table: In progress
6. TATR (Table Transformer):
7. Extractable (Leverages TATR):
8. Gmft (Leverages TATR):
9. Sudhanshu1304/table-transformer:
10. Baskar-forever/TableExtractor-Advanced:
11. LLMWhisperer:
12. Opendatalab/PDF-Extract-Kit:
13. TableNet:
14. Amazon Textract:
15. Google Document AI:
16. IBM Watson Discovery:
17. Microsoft Form Recognizer:
18. Unstructured IO:
19. Devashish Prasad’s GitHub CascadeTabNet: Uses Mask RCNN and ResNeXt. **Pretrained Weights**: Included via Model Zoo section; you can download checkpoint files directly (e.g., for ICDAR/TableBank).
20. YOLOv8s: Hugging Face model foduucom/table-detection-and-extraction

from ultralytics import YOLO

model = YOLO('foduucom/table-detection-and-extraction')

results = model.predict('page\_image.png')

for res in results:

for bbox, cls in zip(res.boxes.xyxy, res.boxes.cls):

print('Table at', bbox, 'type', cls)

1. Table Transformer Detection based model: microsoft/table-transformer-detection  
   trained on PubTables1M