

# Synthetic Tabular Data Diversity and Size Effects on Multimodal Model Generalization

Assignee Research

June 9, 2026

## Abstract

This report synthesises findings from 15 peer-reviewed papers addressing the following research question: What is the effect of synthetic tabular data diversity versus dataset size on the cross-domain generalization performance of multimodal models integrating text and table inputs. 18 claims were extracted from source literature; 2 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 4.2/10. This report is a machine-generated literature synthesis and does not constitute original research.

## 1 Introduction

This paper examines: Table Detection in the Wild: A Novel Diverse Table Detection Dataset and Method. Research question: What is the effect of synthetic tabular data diversity versus dataset size on the cross-domain generalization performance of multimodal models integrating text and table inputs?.

## 2 Methodology

Systematic literature search across multiple databases yielded 15 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 4.2/10.

## 3 Results

15 papers retrieved. 18 claims extracted; 2 independently verified. Quality review score: 4.2/10.

## 4 Limitations

This report is a machine-generated literature synthesis and does not constitute original research. Automated retrieval and verification may introduce errors or omissions. Review scores reflect automated assessment, not human peer review. Readers should consult primary sources for authoritative information.

## 5 Extracted Claims

Claim	Verified	Confidence
The STDW dataset contains 2345 scanned PDFs with 5102 tables.	×	0.03
The STDW dataset contains 4945 searchable documents with 7329 tables.	×	0.04
The STDW dataset has 7294 samples with 900.72 diversity.	×	0.07
The STDW dataset is manually annotated.	×	0.04
The STDW dataset includes both searchable PDFs and scanned documents.	×	0.04
The RetinaNet method achieves an IOU of 0.5 and an AP of 0.78 on the STDW dataset.	×	0.05
The Selective Search method achieves an IOU of 0.5 and an AP of 0.61 on the STDW dataset.	×	0.04
The base learning rate used in the methodology is 0.0004.	×	0.02
The optimizer used in the methodology is SGD.	×	0.01
The momentum used in the methodology is 0.9.	×	0.01
The batch size used in the methodology is 1.	×	0.01
The buffer size used in the methodology is 20 * batch size.	×	0.01
The number of epochs used in the methodology is 30.	×	0.02
The learning rate schedule used in the methodology is Constant.	×	0.02
The loss function used in the methodology is Focal loss.	×	0.06
The feature extractor used in the methodology is Resnet-50.	×	0.02
The STDW dataset is available at <a href="https://www.kaggle.com/datasets/mrinalim/stdw-dataset">https://www.kaggle.com/datasets/mrinalim/stdw-dataset</a> .	✓	0.27
The STDW dataset is available at <a href="https://huggingface.co/datasets/n3011/STDW">https://huggingface.co/datasets/n3011/STDW</a> .	✓	0.24

## References

- <http://arxiv.org/abs/2406.04295v2>
- <http://arxiv.org/abs/2403.10075v2>
- <http://arxiv.org/abs/2209.09207v2>