How does the computational efficiency of FERDERnet scale with increasing input resolution compared to other re

Assignee Research

June 10, 2026

Abstract

The flourishing blossom of deep learning has witnessed the rapid development of text recognition in recent years. However, the existing text recognition methods are mainly proposed for English texts. As another widely-spoken language, Chinese text recognition (CTR) in all ways has extensive application markets. Based on our observations, we attribute the scarce attention on CTR to the lack of reasonable dataset construction standards, unified evaluation protocols, and results of the existing baselines. To fill this gap, we manually collect CTR datasets from publicly available competitions, pro

1 Introduction

This paper examines: Benchmarking Chinese Text Recognition: Datasets, Baselines, and an Empirical Study. Research question: How does the computational efficiency of FERDERnet scale with increasing input resolution compared to other real-time on-road emotion recognition models?.

2 Methodology

Systematic literature search across multiple databases yielded 6 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 3.3/10.

3 Results

6 papers retrieved. 17 claims extracted; 0 independently verified. Quality review score: 3.3/10.

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 experiments were conducted on an NVIDIA RTX 2080Ti GPU with 11GB memory.	×	0.01
The input image was resized into 32 $\times$ 256 for all experiments.	×	0.02
The combined alphabet for all experiments consists of 7,934 characters.	×	0.01
No data augmentation and pre-training strategies were used for the baselines.	×	0.03
The evaluation protocols follow the ICDAR2019 ReCTS Competition rules.	×	0.08
The PRAB method improves the average accuracy of MORAN, SAR, and TransOCR by 1.83%, 3.35%, and 2.33% respectively.	×	0.02
CRNN has 12.4M parameters and 751.0 FPS.	×	0.00
ASTER has 27.2M parameters and 107.3 FPS.	×	0.00
MORAN has 28.5M parameters and 301.5 FPS.	×	0.00
SAR has 27.8M parameters and 93.1 FPS.	×	0.00
SEED has 36.1M parameters and 106.6 FPS.	×	0.00
MASTER has 62.8M parameters and 16.3 FPS.	×	0.02
ABINet has 53.1M parameters and 92.1 FPS.	×	0.02
TransOCR has 83.9M parameters and 164.6 FPS.	×	0.00
The decoder of SEED shows superiority in recognizing low-quality text images.	×	0.02
MASTER utilizes the self-attention mechanism to learn a more powerful and robust representation for distorted text image	×	0.03
ABINet is an autonomous, bidirectional, and iterative method for scene text recognition.	×	0.09

References

- http://arxiv.org/abs/2306.09372v1
- http://arxiv.org/abs/2112.15093v2
- http://arxiv.org/abs/2512.14002v1