On Vocabulary Reliance in Scene Text Recognition

What is the problem?

The issue of "vocabulary reliance" in scene text recognition, where models perform poorly on out-of-vocabulary words despite high accuracy on in-vocabulary words.

What has been done earlier?

Focus on Accuracy: Previous methods improved benchmark accuracy using deep learning and synthetic data but often relied on memorized vocabulary.

Generalization Issues: These models performed well on in-vocabulary words but struggled with out-of-vocabulary words, leading to poor generalization.

Prerna Sahu,B421036

On Vocabulary Reliance in Scene Text Recognition

What are the remaining challenges?

Improving Generalization: One of the main challenges is enhancing the models' ability to accurately recognize out-of-vocabulary words, which are not present in the training data.
Reducing Vocabulary Reliance: Addressing the over-dependence on training vocabulary to ensure that models can handle a wider variety of words and text appearances in real-world scenarios.

What novel solution proposed by the authors to solve the problem?

Analytical Framework: A framework for systematically studying and addressing vocabulary reliance.
Mutual Learning Strategy: Collaborative learning between attention-based and segmentation-based models to improve generalization.
Mixed Dataset Training: Use of combined lexicon-based and random text data to reduce overfitting and enhance generalization.

Prerna Sahu,B421036