
Supplementary Material– When Semantics Mislead Vision: Mitigating Large Multimodal Models Hallucinations in Scene Text Spotting and Understanding

Anonymous Author(s)

Affiliation

Address

email

1

2 In this Supplementary Material, we provide additional resources to support the evaluation and under-
3 standing of TextHalu-Bench. Specifically, we include our testing script **eval.py**, which demonstrates
4 how to evaluate MLLMs on our benchmark using example data. To facilitate quick testing and clarity,
5 we also provide five annotated image-text pairs for each of the five scenario categories—*Business*,
6 *Industry*, *Transportation*, *Public Facilities*, and *Daily Life*. These samples are carefully selected to
7 illustrate the typical challenges within each domain and help users better understand the evaluation
8 format.