

Text benchmark. Planned for a later phase.

We are currently testing different OCR models, including Surya and Sarvam. We will complete the Text benchmark later, after comparing the readers and settling the evaluation approach.

MODELS UNDER COMPARISON

Surya OCR · Sarvam · other OCR readers

Additional candidates include PaddleOCR / PaddleOCR-VL and Tesseract.

NOW

Test the readers

Run the same images through different models and compare how accurately they read Indian-language text.

NEXT

Review the results

Check errors, difficult scripts and unclear labels with human review. Choose a consistent evaluation setup.

LATER

Complete the benchmark

Finalise the Text benchmark and publish comparable results once the model evaluation is ready.

Current status

Model comparison is in progress. The complete Text benchmark is deferred; no final model ranking is being presented.