ChatGPT

Surya OCR vs. Sarvam AI Vision – Competitive Analysis

Executive Summary

- **Accuracy**: Surya achieves ~97% similarity vs Tesseract (88%) on multilingual OCR [1]; Sarvam reports ~95–96% word accuracy on Hindi and other Indic languages [2] (and 84.3% on an English OCR benchmark [3]).
- **Languages & Layout**: Surya supports 90+ languages (global coverage) with built-in layout & table detection [4]; Sarvam targets 23 languages (English + 22 Indian) [5] and also handles document layouts.
- **Performance**: Surya (GPU-accelerated) processes a page in ~0.09s (line detection) [6]; Sarvam's speed isn't published (its 3B-parameter model likely slower per page). Sarvam is cloud/SaaS with on-prem option; Surya is open-source (PyTorch) and self-hosted.
- **Community & Credibility**: Surya is open-source (GPL), with 19.3k GitHub stars [7] and published benchmarks. Sarvam is proprietary (ISO/SOC2 certified [8]) with $41M funding [9] and government partnerships [10].
- **Trade-offs**: Surya excels in flexibility (no vendor lock-in, broad language) but requires in-house management. Sarvam offers a turnkey, enterprise-grade solution with strong Indic performance, at presumably higher cost and less transparency.

1. Technical Performance

OCR Accuracy (CER/WER)

- **Surya**: Benchmarked against Tesseract on multilingual text, Surya achieved an *average similarity* of **0.97** vs **0.88** for Tesseract [1]. This implies Surya's OCR outputs are ~97% identical to ground truth on tested data. (No explicit CER/WER given, but similarity ~0.97 corresponds to ~3% WER.)
- **Sarvam**: Reports very high accuracy on Indian languages. For example, Sarvam Vision attains **95.91%** word accuracy on Hindi and **93.42%** on Tamil [2] (meaning WER ~4% and ~6.6%). On an English OCR benchmark (olmOCR-Bench), Sarvam scored **84.3%** accuracy [3], outperforming Google Gemini 3 Pro (80.2%) and GPT-5.2 (69.8%).

Complex Layouts (Tables, Multi-Column, Forms)

- **Surya**: Includes layout analysis and table detection. On PubLayNet (a scholarly layout dataset), Surya achieved precision/recall of ~0.85–0.93 on images, tables, text blocks, titles [11]. Reading-order accuracy is reported ~88% (on a noisy test) [12]. Table recognition on IBM's FinTabNet gave nearly perfect row detection (1.0) and 0.9863 column intersection [13].
- **Sarvam**: Sarvam's documentation indicates built-in layout analysis (it outputs HTML/Markdown preserving document structure [14]). No quantitative metrics are published, but the business press

highlights “layout-aware extraction” [15]. Sarvam’s focus on layout (forms/APAR) suggests strong performance, but no direct scores are available.

Indic Languages Performance

- **Surya**: Supports **90+ languages** globally [4], including major Indic languages. Surya examples include Hindi text detection [16]. However, no specific accuracy numbers per language are published. As an open toolkit, accuracy will depend on its underlying models (likely trained on multi-lingual data).
- **Sarvam**: Specializes in Indian languages (22 official languages). Their reported word accuracies are very high for many: Hindi 95.91%, Bengali 92.61%, Tamil 93.42%, Telugu 87.70%, Marathi 93.13%, Malayalam 91.60% [2]. This region-specific tuning likely outperforms generic OCR in those scripts. (One outlier is Kashmiri at 55.93%, suggesting some limitations.)

Speed & Throughput

- **Surya**: On GPU (NVIDIA A10), Surya processes ~10 pages/sec for text line detection (0.0945s/page) [6]. Layout analysis is ~0.13s per image [11]. (Tesseract on CPU was 0.29s/page, so Surya is ~3× faster with a GPU.)
- **Sarvam**: No official speed numbers. Given the large model (3B params [17]) and cloud API, we estimate slower per-page inference than Surya’s GPU. Sarvam emphasizes accuracy over raw speed. Pricing for Surya’s hosted API promises “consistent speed” with no spikes [18], but details are proprietary.

Model Size & Hardware

- **Surya**: Requires PyTorch; can run on CPU or GPU. (Installation instructions assume GPU for best performance.) Benchmarks used an A10 GPU (48GB VRAM) [19]. The toolkit auto-downloads model weights (~GB scale) on first run.
- **Sarvam**: Uses a **3-billion parameter Vision-Language Model** [17]. Likely requires powerful GPUs or cloud clusters. Sarvam offers both cloud service and an enterprise on-prem option [20], implying heavy infrastructure. Exact model size not public, but 3B parameters suggests tens of GB memory footprint.

2. Credibility & Adoption

Open Source & Community (Surya)

- **GitHub Stats**: Surya is widely used—**19.3k stars** and **1.3k forks** on GitHub [7]. This indicates strong community interest. It has 952 commits and 135 open issues, showing active development.
- **Benchmarks & Reports**: Surya’s README includes benchmarks vs Tesseract and Google [21] [22]. There are blog posts (e.g. on Medium/DEV) showcasing Surya’s features [“Using SuryaOCR with Docling”] (not official, but user content). No peer-reviewed papers specifically for Surya as of 2025.
- **Usage in Production**: No public enterprise case studies; Surya is a component (Datalab’s platform uses it). Enterprises may be cautiously adopting it via the Datalab API or self-hosting due to GPL. Community reviews (GitHub issues, forums) praise its accuracy but note GPU needs.

Sarvam AI (Proprietary)

- **Company Backing**: Sarvam AI (India) raised **$41M** in funding [9] and has partnerships with Odisha and Tamil Nadu for sovereign AI hubs [10]. It's a startup (founded 2023 [9]) but high-profile.
- **Certifications**: Sarvam claims ISO certification and SOC 2 Type II compliance [8], which builds trust for enterprise use.
- **Benchmarks**: Only the company's own benchmarks are public (as in press releases [3] [2]). No third-party validation available. No GitHub (closed), so no public code review or star-rating.
- **Adoption**: Being new, Sarvam has no widely published case studies yet, but the government associations and tech press coverage (Business Standard, TechCrunch) lend credibility.

3. Deployment & Cost

Feature	Surya OCR (Open Source)	Sarvam Vision (Proprietary)
License	GPL (open-source) [23]; free to use, though commercial use may require license.	Closed-source; SaaS or on-prem enterprise license (pricing undisclosed).
Pricing	Free to install and run. (Datalab offers a paid hosted API separately.)	Custom enterprise pricing (likely higher); no free tier publicly.
Deployment Options	Self-host on any GPU/CPU machine; cloud or on-prem. Datalab also offers a managed API.	Cloud-based SaaS; on-prem option for enterprises (hybrid cloud) [20].
Integration	Python CLI/API; Docker. Requires PyTorch environment.	RESTful API (JSON/HTML output) [14]; SDKs for various languages.
Data Privacy	Fully in control (data stays on your servers).	Can keep data local with on-prem offering. Claims "data residency."
Support & Community	Community support via GitHub/ Discord [24]; commercial support via Datalab API.	Commercial support from Sarvam; marketing claims SLA (SOC2).
Feature Set	OCR + line detection + layout analysis + table recog. (LaTeX OCR included.) [25]	OCR + layout/text+ table extraction, with language models (Sarvam Vision). Full AI platform features (proofreading, summarization coming) [15].

Language & Layout: Surya's universal vision vs Sarvam's Indic focus. Surya handles many scripts out-of-box; Sarvam optimized for Indian context (English + 22 Indic scripts). For an Indian government use-case, Sarvam's specialized training may yield higher raw accuracy on native scripts [2], whereas Surya offers broad coverage and flexibility.

Cost Considerations: Surya is free open-source, reducing licensing cost (only compute/HW cost). Sarvam, being proprietary, likely incurs usage fees (either per-page SaaS or subscription). Surya's cost model shifts complexity to setup/maintenance; Sarvam's shifts cost to licensing but offers a turnkey service.

4. Competitive Positioning

Unique Strengths

- **Surya OCR**: Open-source and extensible; supports a vast range of languages; includes advanced layout/table detection. Community-verified (19k stars) [7]. Fast on GPU hardware. No vendor lock-in – ideal for tech-savvy organizations that want flexibility.
- **Sarvam AI**: Purpose-built for Indian language documents and forms. Offers end-to-end document intelligence (includes reasoning, proofreading). Comes with enterprise security certifications and government-focused marketing. Outsources model maintenance – easier for enterprises that want an off-the-shelf solution.

Trade-offs

- **Accuracy vs Control**: Sarvam likely has edge accuracy on Indian languages (due to specialized training) [2], but Surya can be fine-tuned by users. Surya's open models allow custom domain adaptation.
- **Cost & Deployment**: Surya avoids license fees and can run on internal servers for data sovereignty. Sarvam is a managed service (simpler deployment, but costs more and ties to vendor).
- **Speed & Scalability**: Surya can leverage powerful GPUs for high throughput (as seen by ~0.094s/page) [6]. Sarvam's large model might have higher latency unless heavily scaled; however, it could be optimized in cloud.

Enterprise Risks/Concerns

- **Using Surya (OSS)**: Risks include: needing in-house ML expertise; potential GPL licensing obligations; less formal support (though Datalab has an API offering). For highly regulated use-cases (government), open-source may raise concerns about guaranteed support and accountability.
- **Using Sarvam (Proprietary)**: Risks include: vendor lock-in and higher cost. Being a young startup, long-term viability/performance is not proven beyond initial results. Enterprises will want contract guarantees (SLAs).

Industry Context (2025–2026)

Open-source OCR has matured; tools like Surya provide quality rivaling cloud APIs in many tasks. Proprietary services still win on ease-of-use and support. Many enterprises use hybrid approaches (e.g., run open OCR on-prem for sensitive data, use cloud for less sensitive). The trend is toward "open model + custom solution" vs "one-size-fits-all" SaaS.

5. Validation & Proof Points

- **Surya Benchmarks**: Surya's claimed metrics are documented on GitHub (cited above). No independent benchmark (e.g. academic paper) found, but its GitHub charts provide transparency on test methodology [21] [26].
- **Sarvam Claims**: Only Sarvam's PR and blog provide numbers [2] [3]. No third-party audits are available yet.
- **Case Studies**: Surya is newer and community-driven; no public case study citations. Sarvam lists government partnerships (Odisha/Tamil Nadu) [10] as references.
- **User Feedback**: Surya has many users (GitHub issues/comments). Complaints include occasional OCR errors and GPU resource needs. Sarvam has media praise; no public user forum reviews (the field is niche).
- **Red Flags**: For Surya – reliance on a small core team (datalab-to), GPL license (some companies avoid GPL). For Sarvam – being unproven at scale (yet to show large deployments publicly).

6. Indian Government Context

- **Document Types**: Both tools can process forms and reports. Sarvam likely excels on APAR-like forms (given focus). Surya can OCR mixed scripts and layouts, but may need custom training/tuning for specific forms.
- **Mixed English-Hindi**: Surya explicitly includes Hindi examples [16] and claims 90+ languages, so it should handle code-mixed text. Sarvam was built for Indian context, so it should handle bilingual text robustly (though no public test on mixed layouts is cited).
- **India-Specific OCR**: Besides Sarvam, other India-focused OCR (e.g. Google India, Academic projects) exist, but Sarvam is the main recent product. Surya, being global, is not India-specific but its multilingual support is beneficial.
- **Data Privacy**: Surya (self-hosted) allows data to stay on-premise. Sarvam offers on-prem deployment but is primarily cloud/SaaS. For sovereign data, an open-source (Surya) solution may align better with "digital sovereignty" policies, whereas Sarvam markets itself as an "Indian sovereign AI" provider (partnerships with states) [10].

Recommendations for Positioning

- **Highlight Surya's strengths**: Emphasize its high accuracy (97% vs Tesseract [1]), broad language support, and layout capabilities. Stress community validation (19k stars [7]) and cost advantage (no license fees).
- **Leverage Sarvam's cues**: Acknowledge Sarvam's focus on Indic languages if relevant, but point out that Surya also supports Hindi, Tamil, etc. by name. Example: Surya's README shows Hindi OCR images [16], showing capability.
- **Address concerns proactively**: Prepare to answer about support (mention Datalab's managed API option) and licensing (GPL usage terms).
- **Proof Points**: Use Surya's benchmark charts [21] [11] to show objective performance. For Sarvam, cite press claims [3] but note independent validation is pending.
- **Competitive Edge**: Position Surya as "open AI tool crafted by leading ML community" vs Sarvam as "proprietary startup solution". Tailor pitch: government clients will appreciate open-source transparency and on-prem control; enterprises may value Sarvam's ease-of-deployment.

Sources: Surya's GitHub readme/benchmarks [21] [26] [7] ; Sarvam's blog and news [2] [3] [5] [10] .

---

[1] [4] [6] [7] [11] [12] [13] [16] [18] [19] [21] [22] [23] [24] [25] [26] GitHub - datalab-to/surya: OCR, layout analysis, reading order, table recognition in 90+ languages
https://github.com/datalab-to/surya

[2] Sarvam Vision | Sarvam AI
https://www.sarvam.ai/blogs/Sarvam-vision/

[3] [15] [17] India's Sarvam AI reportedly beats ChatGPT, Gemini in key benchmark tests | Tech News - Business Standard
https://www.business-standard.com/technology/tech-news/sarvam-ai-document-ocr-indic-language-benchmarks-performance-126021001082_1.html

[5] [14] Sarvam Vision | Sarvam API Docs
https://docs.sarvam.ai/api-reference-docs/getting-started/models/sarvam-vision

[8] Introducing Indus | Sarvam AI
https://www.sarvam.ai/blogs/introducing-indus

[9] India's Sarvam launches Indus AI chat app as competition heats up | TechCrunch
https://techcrunch.com/2026/02/20/indias-sarvam-launches-indus-ai-chat-app-as-competition-heats-up/

[10] Sarvam Announces Sovereign AI Partnerships with Indian States | Sarvam AI
https://www.sarvam.ai/blogs/partnerships-with-indian-states/

[20] Sarvam Vision Information - AI Document Intelligence Guide for Indian Languages
https://sarvamvision.com/