PRESS RELEASE

BharatGen launches Patram 1.5, an Indian vision-language model for documents and images

The 7-billion-parameter open-weight model works with scanned, stamped, signed, handwritten and mixed-language documents, and is now available on Hugging Face and the BharatGen Playground.

Key highlights

- Patram 1.5 reads a document as an image and reasons over it, instead of relying only on text extracted from the page; it also understands photographs and diagrams
- Introduces OCR 3.0: users ask for only the information they need, in the format their own systems use
- Trained to say “I don’t know” when a field is blank or a page is blurred, instead of guessing
- Evaluated on 49 tests, alongside 20 other AI models that read images and seven OCR systems

New Delhi, 26 November 2026: BharatGen, the AI company backed by the Department of Science and Technology (DST) and the IndiaAI Mission, MeitY, today launched Patram 1.5, a vision-language foundation model built in India, for India.

Banks, insurers, hospitals and government offices in India handle large volumes of scanned, stamped, handwritten and mixed-language documents every day, from KYC submissions and bank statements to medical reports and government circulars. Patram 1.5 has been designed for such documents, including those with images, photographs, tables, charts, forms, signatures, seals and poor-quality scans.

Most document systems today rely on Optical Character Recognition (OCR), which converts a page into plain text. To answer questions about that text, the output is often passed to a large language model (LLM), which can reason over words but cannot see the page itself. Anything lost in the OCR step, such as the layout of a table, a stamp or a handwritten note, never reaches the LLM. A vision-language model (VLM) such as Patram 1.5 brings the two together: it looks at the page as an image and reads and reasons over it in a single step, so these details remain part of what it understands.

Patram 1.5 also introduces what BharatGen calls OCR 3.0. Earlier OCR returns all the text on a page in a fixed way, or in formats decided by the system, usually through a pipeline of separate tools. With OCR 3.0, this is built into the model itself: users describe in plain language only the information they need, and the model returns it in the format and structure their own systems use. A bank processing statements, for instance, can ask for just the customer ID and branch, as JSON with its own field names.

The model has also been trained to recognise when it cannot answer. If a field is blank or a figure is illegible, it says what it cannot read and why, instead of guessing; where a reasonable guess is possible, it shares the guess and flags it for a person to check. If a

question is ambiguous, such as “the date” on an electricity bill that shows several dates, it asks which one the user means. This matters in financial services, healthcare and governance, where one wrong entry can affect a decision.

BharatGen evaluated Patram 1.5 on 49 tests: 34 public benchmarks and 15 tasks from PatramBench, its own benchmark built from 8,773 pages of Indian documents, which will be released on Hugging Face in the coming days. The evaluation covered 20 other AI models that read images, both specialist and general-purpose, of up to 763 billion parameters, and seven specialist OCR systems. Results for every model and test are available here.

On the BharatGen Playground, people can work with Patram 1.5 through Chat: upload a document or an image, then ask questions about it or tell the model what to pull out, in plain language. The Playground also offers ready-to-use solutions for common document tasks (Parse, Extract, Classify, Split and Section, with Digitize, Translate and Edit coming soon), which can be combined into workflows, such as sorting a set of KYC documents and cross-checking them before a person signs off, with templates for finance, healthcare and legal use.

Speaking on the launch, **Rishi Bal, Chief Executive Officer, BharatGen**, said, “A model alone can only take you so far. The BharatGen Playground gives the world a full stack for document intelligence, with models, solutions and workflows working together.”

Prof. Ravi Kiran Sarvadevabhatla, Consortium Member, BharatGen, and Professor at IIIT Hyderabad, said, “With Patram 1.5 and BharatGen Playground, we chart a new capability frontier for sovereign vision models and document understanding workflows.”

Availability

Patram 1.5 is available as open weights on Hugging Face. At 7 billion parameters, it is compact enough for banks, hospitals and government departments to run on their own infrastructure, so sensitive documents need not leave their systems. It can also be tried on the BharatGen Playground. Patram 1.5 is the latest in BharatGen’s Patram family, which began with Patram 1 in June 2025. More details are available here.

Technical details

- **Type:** vision-language foundation model, 7 billion parameters
- **Reads:** documents as images (scans, photos of pages, forms), as well as photographs and diagrams; several images in one request
- **Returns:** text, including Markdown, HTML and JSON
- **Languages:** English, and English mixed with Indian languages
- **Context length:** up to 32K tokens
- **Benchmark:** PatramBench, 11,828 test items from 8,773 pages of Indian documents (to be released on Hugging Face in the coming days)
- **Access:** open weights on Hugging Face; BharatGen Playground

About BharatGen

BharatGen builds open-weight foundation AI models from scratch for India across text, speech and vision, with its text models covering all 22 Scheduled Indian languages. The models are available on Hugging Face and AIKosh, so developers, startups, researchers and enterprises can adapt them and run them on their own infrastructure. BharatGen also

works directly with government and industry partners to deploy them in governance, financial services, education, healthcare and culture. Established in October 2024, BharatGen is a Section 8 not-for-profit AI company backed by the Department of Science and Technology (DST) and the IndiaAI Mission under MeitY. It is anchored at IIT Bombay and draws on a consortium of nine of India's top academic institutions. Learn more at bharatgen.com.

For media queries, please contact:

Kiwi Network
Shivani Bali
pr@kiwinetwork.in
+91 99718 43743