Teaching Machines to See India, and Getting the World to Notice



Why a centre, and why now

Computer vision has become the eyes of modern AI. It reads satellite images for crop health and flood extent, screens medical scans, and lets drones and robots act in the physical world. Most of the models doing this work, though, are trained on data from elsewhere and built in labs abroad. They often stumble on Indian terrain and Indian sensors, and India has had little say in how they are designed.



Talent was never the shortage. Indian students fill vision labs across the world. What the country lacked was a place at home with the funding, compute and ambition to do frontier work and publish it at top-tier conferences including CVPR, ICCV, ECCV and NeurIPS, the venues where the field decides its direction. Indian groups are still a small presence there.

HOW IT BEGAN

VIMAGE was set up to close that gap. Through its CSR programme, HRTI Private Limited partnered with IIT Bombay to establish the IITB–HRTI Centre of Excellence in Vision, Multimodal AI and Geo-Intelligence, with Prof. Biplab Banerjee of CSRE and C-MInDS as Principal Investigator.

Instead of building one lab around one group, the Centre opened an institute-wide call for proposals. Faculty from many departments answered it, and VIMAGE today supports more than 20 projects across vision, multimodal learning and geospatial AI. The reasoning was practical. Good vision problems come from medicine, engineering and the earth sciences as often as from computer science, and the money should follow the problems.

A FIRST YEAR OF RESULTS

The work has travelled fast. In 2026 alone, 29 papers carrying the Centre's affiliation have been published, among them main-track papers at CVPR and ECCV, further work at WACV, BMVC and ACL Findings, and a dozen journal articles.

Geo-intelligence is where the Centre has gone furthest. Its GeoMeld dataset aligns about 2.5 million optical, radar and land-cover samples and attaches captions that an agentic pipeline writes and then checks against physical measurements. The GeoMeld-FM foundation model is trained on it, and the whole resource, presented at the CVPR 2026 MORSE workshop, is open to anyone.
GeoDisaster, to be presented at BMVC 2026, asks whether AI agents can carry out real disaster workflows such as flood mapping, damage assessment and safe routing. Agents are scored against answers computed by GIS tools, not answers written by another language model. Multi-sensor datasets and agent benchmarks are now appearing worldwide, but few groups ground both language and agents in measurable physical evidence this way.

The team is also building uncertainty-aware models that can tell when they should not be trusted, a property that matters as much in a hospital as in satellite monitoring.

BUILDING A COMMUNITY

Papers are only part of the story. The Centre's first national workshop on Recent Trends in Computer Vision Technologies brought more than 160 researchers, founders and policymakers to campus. A Winter School follows in November 2026, and a Distinguished Lecture Series will bring leading vision researchers from around the world to Powai.

A growing team of postdoctoral fellows, PhD scholars and research engineers now works under the Centre, and its links with government agencies and industry keep the research tied to problems people actually need solved. A dedicated VIMAGE space is being readied on campus, with shared GPU infrastructure for large vision workloads. Most Indian labs still borrow that kind of compute one project at a time.

Bridging Research & Policy

"In two years I want VIMAGE to be a regular name at CVPR, ICCV and NeurIPS, etc. with open models that Indian agencies and startups are actually using." Prof. Biplab Banerjee, Principal Investigator

WHAT COMES NEXT

The near-term plan follows from that. Open the new space and compute facility, make the Winter School and lecture series annual fixtures, and release the GeoMeld-FM models and agent benchmarks for others to build on.

The longer goal is bigger. Prof. Banerjee wants a trustworthy geospatial AI stack for India, built and maintained in the country, and a centre that international researchers choose to visit and strong Indian PhDs choose to stay for.

WHY IT MATTERS BEYOND POWAI

For India, VIMAGE is a working example of how industry CSR can fund frontier science without tying it to a single product. Its outputs are open, its grants cross departments, and its research is tested against national needs in agriculture, disaster response and health.



For the global field, it adds data and models from a part of the world that today's training sets barely see, and that makes everyone's systems more robust.

The aim was never another lab. It is a place where India's best vision research is done, published and put to work.

BACK