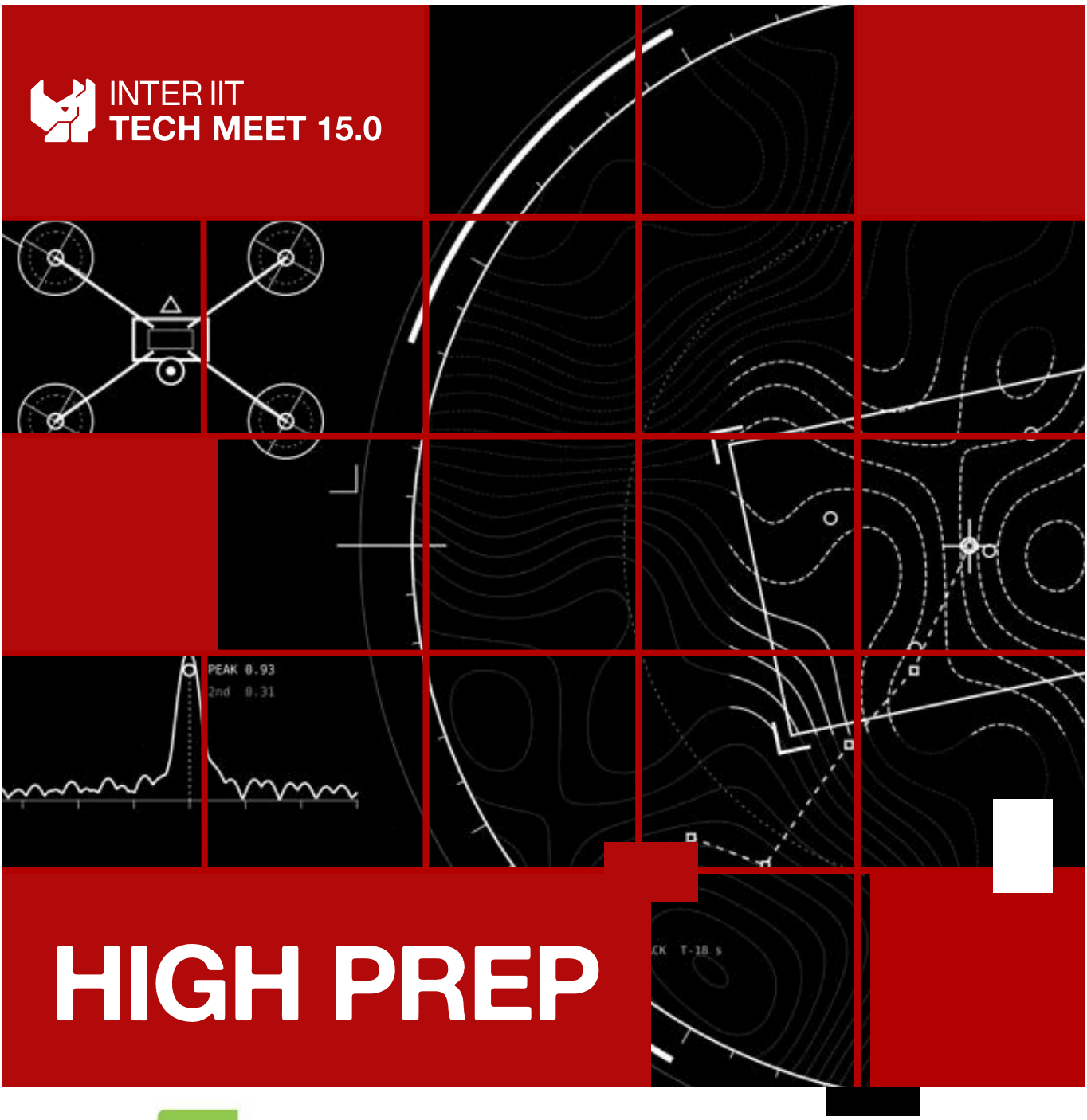
INTER IIT
TECH MEET 15.0



HIGH PREP

ideaForge®
Create. Inspire

UAV Visual Geo-Localisation Challenge

About ideaForge

At the forefront of cutting-edge drone technology, with a focus on developing superior solutions for intelligence, surveillance, and reconnaissance (ISR) purposes, ideaForge has its origins nearly two decades ago. The company is built with a mission to develop technically superior UAVs for various missions, utilizing modern advancements in aerospace and robotics. From being a humble start-up, now ideaForge Technology has recently offered its IPO and become a listed company in NSE and BSE.

ideaForge is a vertically integrated company with an in-house product development centre, allowing us to design, develop, engineer, and manufacture indigenous unmanned aerial vehicles (UAVs). ideaForge was founded in 2007, and over time became the pioneer and pre-eminent market leader in the Indian unmanned aircraft systems ("UAS") market. ideaForge had the largest operational deployment of indigenous UAVs across India, with an ideaForge manufactured drone taking off every four minutes on average for surveillance and mapping. As of today, their customers have completed over 500,000 flights using our UAVs.

ideaForge UAVs are equipped with industry-leading specifications and capabilities, comparable to those of other established global players in the UAV industry. Their in-house design, development, engineering, and manufacturing capabilities have enabled us to develop better products in response to changing customer demands, thereby improving customer experience with our products. The company serves domestic and international customers in the defense and civil sectors, primarily for surveillance, mapping, and surveying applications. Our customers include the Indian Armed Forces, homeland security authorities comprising certain central armed police forces, state police departments, disaster management forces, forest departments, government along with private entities, enterprise sectors and many more.

The company is closely working with the Indian Armed Forces for the deployment of their drones for important security and surveillance missions. ideaForge UAVs are deployed across Indian borders, assisting the Indian Army in surveillance and secure the nation.

Furthermore, several Police Departments, Government and private authorities have deployed the ideaForge UAVs for a variety of surveillance, mapping, and surveying operations such as anti-terror, anti-Naxal operations, illegal activity prevention, crime control, traffic monitoring, mine area planning, volumetric estimations, land survey and many others.

Background and objective

UAVs for outdoor wide area surveillance depend on GNSS for position and navigation. When GNSS is jammed, spoofed or unavailable, matching the onboard camera view to stored satellite imagery is one of the few ways to recover an absolute position fix. The hardest case, and the focus of this challenge, is thermal imagery matched against daytime satellite maps, by day and by night, with seasonal change, low-texture terrain, and a tight compute budget to run on the UAV using edge processing.

In this challenge, teams will build a system that locates a single thermal (LWIR) UAV image on a satellite map, running entirely on an embedded computer.

The goal is a robust, self-aware localizer: accurate where a match is possible, and willing to decline where it is not. A confident wrong answer is treated as worse than no answer.

Problem Statement

Each test case is a single thermal UAV image and an approximate location. The final evaluation will use a reference region of around 100 km x 100 km. The approximate location is a single WGS84 point; the true position is always within 5 km of it, with the offset sampled uniformly. The objective is to calculate the geo-coordinate (WGS84 latitude and longitude) of the UAV's position when each image was captured. Not every image will resolve into a high-confidence match, so an image may also return a “no match” status. The scoring assigns higher points to more accurate positions and negative points to very large errors.

Metadata and camera

- UAV altitude above the take-off point.
- UAV attitude (yaw, pitch, roll) in degrees, as the body-to-world rotation from FRD body axes to NED, in aerospace ZYX order: R = Rz(yaw) · Ry(pitch) · Rx(roll).
- The camera is hard-mounted looking straight down (nadir at zero roll and pitch) and rotated −90° about its optical axis, Rz(−90°): image right points to the UAV's left and image down to the UAV's front.
- Camera details: resolution 320 × 256, nominal horizontal FOV of 92°, intrinsics, distortion coefficients and image format. Only the nominal mount is given; the exact boresight calibration and lever arm are not supplied. Images are provided as captured (not undistorted).
- Ground truth is the UAV's GPS position logged during the flight.

Evaluation Data. The evaluation set consists only of thermal (LWIR) images: 16-bit, non-radiometric, from a nadir-mounted camera that is not gimbal-stabilised. Images are independent single frames, captured by day and by night at altitudes from 100 m to 3000 m above take-off. Some frames may turn out to be hard or impossible to match, for example over open water, cloud, haze or featureless terrain. No unmatchable frames are added artificially: the test set is real flight imagery as captured. Frames are not labelled as matchable or unmatchable, because a frame one method cannot match may be matched by another; every frame is scored the same way, and declining is the right choice whenever a method is not confident.

Reference Data. The satellite reference is Airbus optical imagery with 6 spectral bands, not orthorectified. No thermal reference imagery, no elevation model and no labelled training set are supplied. Teams may use public datasets, such as elevation models, Sentinel-2 or OpenStreetMap, if they declare them and their terms of service permit this use.

System assumptions

Model training, data pre-processing and development testing may happen on any system available to the teams. However, in the final evaluation all solutions must run on the fixed target hardware which would be an NVIDIA Jetson Orin NX 16 GB with 512 GB NVMe SSD, in 15 W power mode.

Competition phases and dataset

The competition runs over 8 weeks in three phases, ending with a blind evaluation on a region teams have not seen.

1. Problem Definition and Prototype Development (weeks 1–2)
The complete details of the problem which will complement this high level problem specification, including the exact input and output formats, metadata schema, satellite datasets to be used, hardware and software configurations of the embedded compute module to be used in evaluation, exact scoring constants, etc. will be communicated by the end of Week 1. The reference datasets and a development set of thermal images from real flights will be provided by the end of Week 2. Teams may use public datasets and create synthetic training data.

2. Mid-point review (end of week 4)
Teams must submit both a brief design note on their solution approach and a working container that runs on the target hardware and software configuration. This will ensure that solutions and the input output interfaces fit the platform early on and avoid rework at the final evaluation stage.

3. Final evaluation (end of week 8)
For the final evaluation, the reference matching region will be provided to the teams 24–48 hours before the evaluation deadline. Teams can use this advance preparation period to pre-process data or generate any artefacts which need to be loaded on the compute platform for matching during the evaluation. The blind evaluation dataset will be provided on the day of the final evaluation. Depending on the logistics feasibility of the final evaluation phase, teams may be asked to submit the container of their evaluation to be run on the organizer's hardware or asked to run the container on their own hardware conforming to the same specs, in the presence of the organizers and generate the results artefacts.

Rules of engagement

Solutions must be fully automatic, self-contained and reproducible on the target hardware.

Submission

- One Docker image built on the target hardware and specified base image, with all models and code/scripts inside it.
- A single entry point that takes the reference data folder, the input folder and an output path. The container processes the images one at a time, in the listed order.
- Output in the specified format: one row per image with name, status, latitude, longitude (if matched) and an optional confidence score. The confidence score is not used in scoring. A missing or malformed row counts as “no match”.
- A technical report describing full details of the method, data and licences used, any internal benchmarking results on public or test datasets and known limitations.
- Extra files carry no penalty. Deployment on a real UAV is not required.

Allowed

- Any method or model, including pretrained models and datasets whose licence permits non-commercial research use, declaring the public pre-trained models and datasets used in the final report.
- Synthetic training data, including EO-to-thermal translation and views generated from the supplied satellite reference.
- Teams may preprocess the final evaluation region's reference data on any hardware of their choice, within the advance window, including training or adapting models on it. Preprocessing must be fully automatic, using code, weights and configuration submitted before the region is released. No manual annotation, tuning or code changes are permitted after release. Organisers may re-run the submitted pipeline to verify the results.

Not allowed

- Network access during evaluation, or downloading any data at run time.
- Attempts to reverse engineer ground truth from file names, ordering, timestamps, etc.
- Using information from one test image when answering another: each image is answered on its own.
- Using any reference or training imagery of the evaluation region other than what organisers supply and declared public datasets.

Verification.
Organisers may inspect the code, execution logs of any top-ranked submission and re-run it. A violation leads to disqualification.

Data handling.
All supplied imagery is for this competition only, may not be redistributed, and must be deleted afterwards. Derived indexes and model weights may be kept.

Evaluation Dataset Scoring Methodology

The evaluation score on the blind dataset is based on accuracy and availability. Compute time, storage and preprocessing time are not scored; they are used to break ties.

Position Error: Error is the geodesic distance (WGS84) in metres between the predicted and true UAV position. The accuracy a system can reach depends on altitude, because each image pixel covers more ground the higher the UAV flies. Each error is therefore compared with a reference error σ that grows with the image's ground sample distance (GSD), the ground size of one pixel at the image centre looking straight down:

$$\sigma = \sqrt{(\ 8^2 + (3 \times GSD)^2\ )}\ \text{metres}, \quad GSD = h / f$$

where h is the altitude above take-off (with a minimum of 100 m) and f is the focal length in pixels from the supplied intrinsics. The 8 m floor covers ground-truth and reference-map errors that teams cannot control.

Prediction Accuracy	Points
Error up to 1σ	+10
1σ – 2σ	+6
2σ – 4σ	+3
4σ – 8σ	0
Over 8σ	−20
No match returned	0

For example, at 1000 m σ is about 17 m: errors up to 17 m earn +10, and errors above about 138 m earn −20. A wrong answer costs twice what a perfect one earns, so a reliable reject step is rewarded. An error exactly on a boundary takes the better range.

Accuracy Score:
The accuracy score is the mean points over all test images, normalised by the maximum of 10 points per image:

accuracy score = (total points) / (10 × number of test images)

A perfect submission scores 1. Images returned as “no match” score 0, so availability is built into the score: returning “no match” on harder images, such as higher altitudes, lowers it, and answering more images only helps when those answers are accurate. Because σ scales with GSD, a point earned at 100 m is worth the same as a point earned at 3000 m. Availability, points per answer and accuracy at different altitudes will be reported for every team.

Time limit and latency:
Each image must be processed within 30 seconds on the target hardware; an image not answered within 30 seconds counts as “no match”. Latency, the total time taken to process the entire final blind evaluation dataset (excluding the reference preprocessing done in the advance window), is not part of the score and is used only as a tie-breaker.

Final Evaluation Dataset Score:
The final evaluation dataset score is the accuracy score scaled to its 60% weight: final evaluation dataset score = 60 × accuracy score, out of 60. A negative accuracy score counts as 0. Differences within the statistical uncertainty of the score are treated as ties, broken in this order: (1) lower compute time, the total latency of the blind run; (2) lower storage used on the SSD by the Docker image and reference artefacts; (3) lower preprocessing time, the wall-clock time of the automatic reference preprocessing in the advance window, taken from its submitted logs.

Final Contest Scores

The final score for the entire problem duration will be composed of the Mid-term and Final presentations and the final evaluation dataset score

Component	Weight
Mid-point review	20% (based on clarity of solution approach, sample container complying to input output interfaces)
Final presentation	20% (based on explanation of solution approach, presentation and response to jury questions)
Final evaluation dataset score	60% (based on the final evaluation score for accuracy and availability; compute time, storage and preprocessing time break ties)

Note: *The exact and full problem specifications, including input and output formats, metadata schema, satellite dataset details, the hardware and software platform specifications, and the final scoring constants, will be declared by the end of Week 1.*

Disclaimer:
Since this is a high-prep, research oriented problem, not all aspects of the constraints, assumptions, criteria etc. may be fully detailed in the limited problem statement text - certain specifics may be missing from the problem specification which may be clarified over the Discord group communications or during any review. In case of any such findings, the organizers reserve the right to add these clarifications to the problem statement with the stipulation that all changes or clarifications will be communicated at the same time to all teams even if the query is raised by one team and that these clarifications do not materially alter the nature of the problem itself. The final decision in this regard will rest with the organizers themselves and cannot be challenged by any of the teams.