

AI-powered System to Automate and Analyze Data from Various Groundwater Sources

AI and Computer Vision help transform traditional manual water data collection into a reliable, scalable, and real-time digital process, supporting India's water security mission effectively.

April 2025



Established in 2005 by philanthropist Rohini Nilekani, Arghyam is a public charitable foundation committed to strengthening the ecosystem's capacity to achieve water security in India. The organization operates in the water sector, focusing on ensuring safe and sustainable water for all and emphasizes a community-led approach, designing solutions for scale, and generating trusted data. Notably, Arghyam has developed the India Water Portal, a comprehensive knowledge resource on water, and has been instrumental in promoting participatory groundwater management. Its collaborative efforts with civil society, government, and market players position Arghyam as a significant influencer in India's water sector.

Headquartered in Bengaluru, Karnataka, Arghyam has a pan India presence, having supported over 140 projects across 22 states, reaching millions of individuals.

Problem Statement

Arghyam, in its mission to enable water security across India, is playing an advisory role in implementation of the Government of India's (GoI) Jal Jeevan Mission (JJM), under which GoI is committed to providing potable water to every household in the country. One of the challenges of JJM implementation is the lack of evidence of service delivery on the ground, and therefore, a lack of confidence in the delivery outcomes of JJM.

Arghyam decided to solve this problem using the latest advancements in technology such as AI and ML, and experimented with the idea of capturing the Bulk Flow-Meter (BFM) readings, at regular intervals, of every water scheme implemented as part of JJM in the state of Assam (JJM, Assam). As part of the experiment, a pump operator would read the BFM reading and send it across to a published number at regular intervals. By cross-referencing the reading with the number of households in the service area of the JJM, Assam scheme, the amount of water delivery can be deduced, thus helping determine the delivery outcome. However, this manual process is prone to human error, irregularity, and delays in data capture, and lacks evidence, especially given the scale and diversity of human factors involved.

To ensure a high level of data quality, and real-time monitoring, Arghyam needed an automated solution to reliably capture the BFM readings at regular intervals. Therefore, Arghyam decided to capture photographs of BFMs and extract readings using various AI/ML techniques. This is when Arghyam engaged Sahaj Software to build a robust AI/ML-powered software platform that could interpret these BFM photographs, at scale, with minimal human intervention.

Challenges Prior to Sahaj Software's Involvement

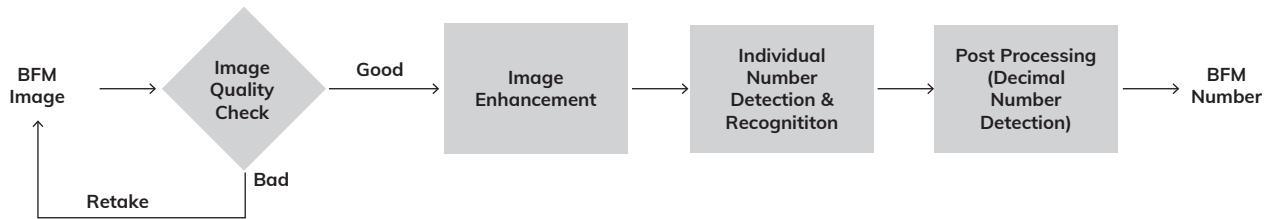
- **Unstructured Image Data:** Flow-meter images varied in angle, lighting, and quality due to field conditions.
- **Manual Interpretation:** Field staff had to interpret and transcribe readings, often leading to mistakes.
- **No Real Time Feedback:** Errors were caught too late, leading to the loss of critical water usage data.
- **Lack of Scalability:** Manual processes are hard to maintain as the initiative is expected to be rolled out across all states.

Engagement Summary

Sahaj designed and delivered an AI-driven pipeline capable of automated flow-meter digit recognition. The solution was lightweight, adaptable, and capable of being integrated with Arghyam's digital workflow. Delivered in under 10 weeks, the project laid the technical foundation for a more automated, trustworthy, and scalable water monitoring system.

Solution Developed by Sahaj

To address the challenges faced in manual BFM readings and data inaccuracies, Sahaj developed a comprehensive and modular Computer Vision (CV)-based solution. The goal was to achieve 97% accuracy in reading extraction on good-quality images, and to build a scalable, open-source backend pipeline to support rural water management efforts, starting with deployment in Assam.



Overall Technical Workflow of the Solution

Key Components of the Solution

Sahaj engineered a set of specialized models and modules that work together to form a robust CV pipeline:

- 1. Image Quality Classification Model:** Automatically evaluates image clarity and tags low-quality images, prompting retakes during data collection. This ensures only reliable images are sent for reading extraction.
- 2. Image Enhancement Module:** Applies preprocessing techniques (e.g., deblurring, contrast enhancement) to improve image quality, increasing the chances of successful digit recognition.
- 3. Broad Water Meter Detection Model:** Locates and crops the meter region within the image using bounding box detection trained on annotated datasets.
- 4. Individual Digit Recognition Model:** Identifies each digit within the meter region using deep learning based techniques, fine-tuned for meter-specific fonts and lighting variations.
- 5. Color Classification Model:** Differentiates red digits from black ones. Since red digits typically represent the decimal part, this helps infer the correct structure of the reading.

This step-by-step architecture was intentionally built to handle a wide range of edge cases from poor lighting and reflections to oriented meter images. The result is a highly accurate, low-latency solution that automates a previously manual, error-prone task, enabling scalable and dependable water monitoring.

Challenges in Executing the Project

Despite a well-defined solution, executing the project in real-world conditions brought several technical and data-related challenges that required thoughtful problem-solving and iterative refinement.

- **Inconsistent Image Quality**

Field images varied widely in resolution, lighting, orientation, and clarity—impacting digit extraction accuracy. We built an image quality classification model to filter unreadable images and ensure only viable inputs reached downstream models.

- **Limitations of Rectangular Bounding Boxes**

Standard rectangular annotations failed on tilted meters, capturing irrelevant parts of the dial. Switching to Oriented Bounding Boxes (OBBs) improved cropping precision, digit isolation, and model performance, especially for rotated images.

- **Red Digit Identification**

Detecting the red-colored decimal digit was critical but difficult due to lighting variation and background noise. A pixel-comparison method (last digit vs. previous digits) outperformed basic color thresholding. We also trained a deep learning classifier on 2,000 annotated samples, which confirmed the algorithm's accuracy.

Key Achievements and Results

High accuracy

The solution achieved a reading extraction accuracy of ~97%, exceeding the initial target of 90%. This level of precision significantly reduces the need for manual verification and establishes trust in automated reading.

Faster Latency

The end-to-end pipeline—submission of an image to flow-meter reading—operates with a latency of just 1–1.5 seconds per image, making it practical for real-time deployment in the field. In fact, if the system leverages GPU, the inference time comes down to as low as 600 ms. These results validate both the technical robustness and deployment readiness of the system.

Road Ahead

With the successful implementation of automated flow-meter reading recognition, Arghyam is now taking a bold step toward building a larger digital public infrastructure called Jal Soochak. This upcoming platform, envisioned as a WhatsApp/mobile app-based system, aims to provide deep visibility into the functioning of drinking water schemes across India under the JJM. Despite thousands of crores being allocated by the Gol in each planning cycle for JJM, a critical

gap persists: lack of accurate, on-the-ground data to inform where to invest. Jal Soochak aims to bridge this gap, enabling data-driven planning and targeted interventions..

This data-driven platform could become a transformational decision-support tool for policymakers, helping optimize resource allocation, monitor impact, and rapidly identify problem areas.

To shape this vision further, Sahaj and Arghyam conducted a Discovery Exercise to collaboratively design the broader architecture, user experience, and data flows of Jal Soochak. The goal is to create a scalable, inclusive, and actionable platform that empowers communities and governments alike to ensure safe and sustainable water access for all.

" We truly appreciated the professionalism, clear communication, and thoroughness your team demonstrated throughout the project. What stood out most was the exceptional ownership taken at every stage, that reflected in well thought out decisions and careful implementation. The dedication and strong work ethic at Sahaj have clearly driven the excellent results we achieved together."

Gurudutt Ramchandra
Manager Solutions, Arghyam