A Deep Learning Approach to Water Point Detection and Mapping Using Street Level Imagery

Neil Patel
Massachusetts Institute of Technology

Abstract

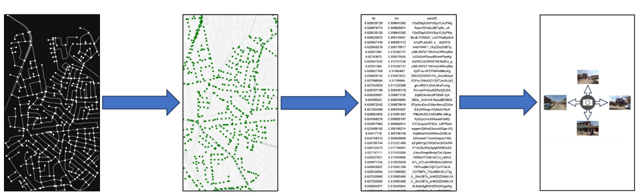
Households in developing countries often rely on alternative shared water sources that exist outside of the datasets of public service providers. This poses a significant challenge to accurately measuring the number of households outside the public service system that use a safe and accessible water source. This paper proposes a novel deep learning approach that utilizes a convolutional neural network to detect water points in street-level imagery from Google Street View (GSV). Using a case study of the Agege local government area in Lagos, Nigeria, the model detected 36 previously unregistered water points with 94.7% precision.

Context

- ~44% of Lagos households rely on off-premises water sources, often financed by non-governmental entities such as NGOs, politicians, or philanthropists[1]
- The Geo-Referenced Infrastructure and Demographic Data for Development (GRID3) Nigeria Lagos Public Water Points Dataset is the primary existing geospatial dataset of communal water points, but covers just 16% of the estimated publicly owned facilities in Lagos State[2]
- Lagos has deep GSV coverage, allowing water points to be geo-referenced through deep learning on street-level imagery

Data Collection

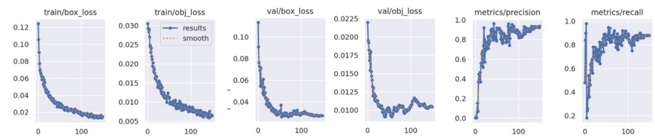
- Downloaded street network of Agege local government area (LGA) from OpenStreetMap
- Randomly generated 1,028 request points at 50m intervals across the street network
- Extracted metadata on date, latitude, longitude, and panoramic ID for 776 request points accessible from GSV
- Downloaded four 600 × 400-pixel images at 0-, 90-, 180-, and 270-degree angles to fully capture the surrounding area



Download Street Network | Generate Image Request Points | Extract Metadata | Download Images

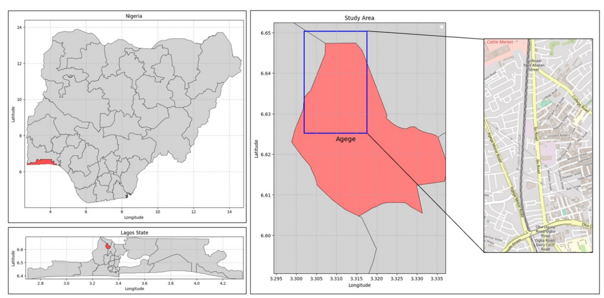
Model Training

- YOLOv5s model trained on 215 GSV images, randomly assigned to training (70%), validation (20%), testing (10%)
- *Epochs/Batch Size:* 150 epochs with batch size of 16
- *Optimizer:* Stochastic Gradient Descent
- *Loss Function:* YOLOv5 loss function (combination of objectness loss, classification loss, and box loss)
- *Performance:* 92.8% precision, 88.0% recall, 90.3% F1-score



Results: Agege Local Government Area

- Model deployed using a case study of Agege LGA with a minimum 80% confidence threshold to reduce false positives
- 38 water facilities detected at 37 sites, with 2 false positives ➔ 36 true facilities geo-referenced with 94.7% precision
- GRID3 dataset did not register any points in study area



GRID3 Dataset | Google Street View Detections

Scalability & Cost-Effectiveness

- Covering major cities requires the ability to analyze 1M+ GSV images in a cost and time-efficient manner
- Google Street View API provides $200/month credit, enough for 28,571 images/month free-of-charge (~371 km)
- Cloud-hosted GPUs allow rapid inference speed without requiring investments in computing hardware

Hardware	Cost	Inference Time (s)	Inference Speed (f/s)
NVIDIA A100-SXM4 GPU	Premium	4,829	0.63
NVIDIA Tesla V100 SXM2 GPU	Premium	2,133	1.42
NVIDIA Tesla T4 GPU	Free	2,073	1.47
Google TPU v2	Free	1,589	1.91

Limitations

- Cannot distinguish between "public" and "private" based on visual features alone – requires survey or manual inspection
- GSV imagery is only accessible from street network – model can only detect water points visible from the nearest road
- Coordinate locations reference the image capture site, which may not match the exact location of the water point

Conclusions

- Model successfully geo-references 36 previously-unregistered water points with 94.7% precision
- Access to free cloud-hosted GPUs enables rapid inference at scale at minimal financial cost
- Requires complementary data collection to refine coordinate estimates and evaluate contextual factors, including ownership, accessibility, and affordability

References

[1] Federal Ministry of Water Resources (FMWR), Government of Nigeria, National Bureau of Statistics (NBS) and UNICEF. 2022 Water, Sanitation and Hygiene: National Outcome Routine Mapping (WASHNORM) 2021: A Report of Findings. Abuja, Nigeria: FCT

[2] GRID3 Nigeria. 2022 Lagos Public Water Points. [Dataset]. GRID3 Nigeria. https://grid3.gov.ng/dataset/lagos-public-water-points/

Links

GitHub | Model | Contact