Recovering Origin–Destination Flows from Bus CCTV: Early Results from Nairobi and Kigali

Nthenya Kyatha, Jay Taneja
College of Information & Computer Science, UMass Amherst
Corresponding email: mkyatha@umass.edu

Motivation: Why OD from CCTV?

Public transport in sub-Saharan Africa (SSA) is dominated by buses and matatus that routinely operate over capacity.

Agencies still rely on manual OD surveys that are:

- infrequent and expensive
- hard to scale
- blind to illegal roadside stops and detailed in-vehicle flows.

Existing automated approaches (AFC, APCs, RFID, mobile phone data) are:

- hardware-heavy, high-maintenance
- often not deployed at scale in SSA
- usually provide aggregate counts, not passenger-level OD.

Key idea: Use CCTV that already exists for security + telematics to recover passenger-level OD flows with minimal extra hardware.

Contribution

1. Baseline 4-stage CCTV + telematics pipeline for OD inference tailored to SSA buses.
2. Deployment on real buses in Nairobi and Kigali using only existing CCTV + telematics—no extra sensors.
3. Quantitative evaluation on annotated short clips:
   - High accuracy under low-density, well-lit conditions
   - Sharp degradation under overcrowding, modality shifts, posture changes, and non-standard door use.
4. Simple, domain-informed fixes (door-state gating + flexible stop definition) that substantially boost exit-door accuracy without changing the detector or Re-ID backbone.

Takeaway: We can already recover useful OD matrices from messy SSA CCTV, but door context and telematics matter more than just better CNNs.

UrbanAI Context

Setting: Buses in Nairobi (Kenya) and Kigali (Rwanda) with onboard CCTV and telematics.

Typical challenges:

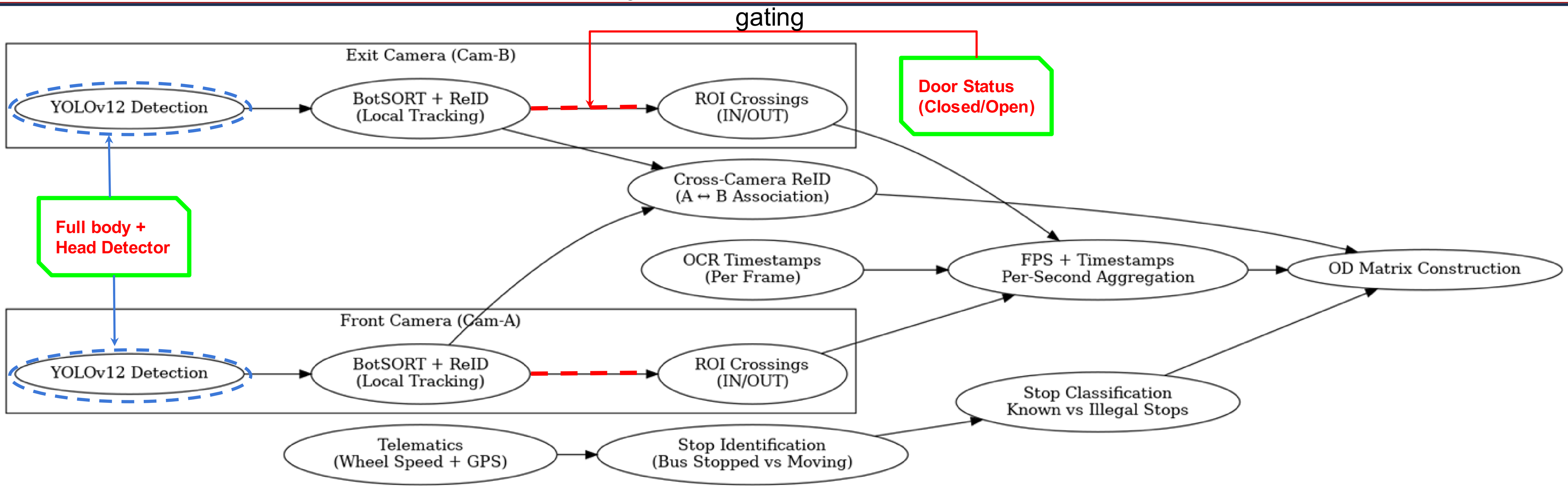
- Overcrowded doorways
- Low-resolution, sometimes monochrome CCTV
- Non-standard door usage
- Illegal stops in traffic, away from official bus stops.

Goal: Enable data-driven route design, equity analysis, and e-bus planning using vision + onboard sensors.

Data & Evaluation

- 11 CCTV segments, 3–8 minutes each.
- Conditions evaluated include low, medium and peak crowding, color vs monochrome segments and partial occlusions.
- Ground truth data contains per-door entry/exit counts and derived OD matrices per clip.

System Overview



Two–stream pipeline with new and enhanced modules highlighted. **Baseline (black):** Same logic for both cameras but works in low-crowd conditions. **Reality:** Exit door is a choke point → heavy occlusion, queues, posture changes.

Baseline Pipeline

- For each camera, detect and track passengers with BotSORT + ReID to get per-camera tracklets.
- Link tracklets from front camera (Cam-A) and exit door camera (Cam-B).
- OCR to extract on frame timestamp & aggregate ROI-based IN/OUT events into per-second counts.
- Use telematics data (GPS, wheel speed) to detect and classify stops.
- **New:** Use **head detection** for IN/OUT events during **overcrowding** and **activate** counter **ONLY** when door is **open**.
- For each passenger trajectory, identify **boarding stop** and **alighting stop** → increment cell in OD matrix.

Results

Entrance Video	Ground Truth		Baseline		Door State		Hybrid Det		All Together	
	IN	OUT	IN	OUT	IN	OUT	IN	OUT	IN	OUT
9:03 am	8	0	7	0	7	0	7	0	7	0
12:28 pm	40	1	39	0	39	0	39	0	39	0
7:45 pm	0	3	0	2	0	2	0	2	0	2
8:42 pm	1	0	1	0	1	0	1	0	1	0
9:11 pm	45	3	42	2	42	2	42	2	42	2

Exit Video	Ground Truth		Baseline		Door State		Hybrid Det		All Together	
	IN	OUT	IN	OUT	IN	OUT	IN	OUT	IN	OUT
9:03 am	0	0	0	1	0	0	0	1	0	0
9:34 am	3	1	3	2	3	1	3	2	3	1
12:28 pm	0	0	1	1	0	0	1	1	0	0
7:45 pm	0	0	0	0	0	0	0	0	0	0
8:42 pm	0	4	0	4	0	3	0	4	0	1
9:11 pm	0	0	0	0	0	0	0	0	0	0

- Hybrid detector does not improve exit accuracy on current segments vs baseline.
- **Takeaway**: Simple domain cues (door state, near-zero speed with open doors) fix more than just adding another detector.

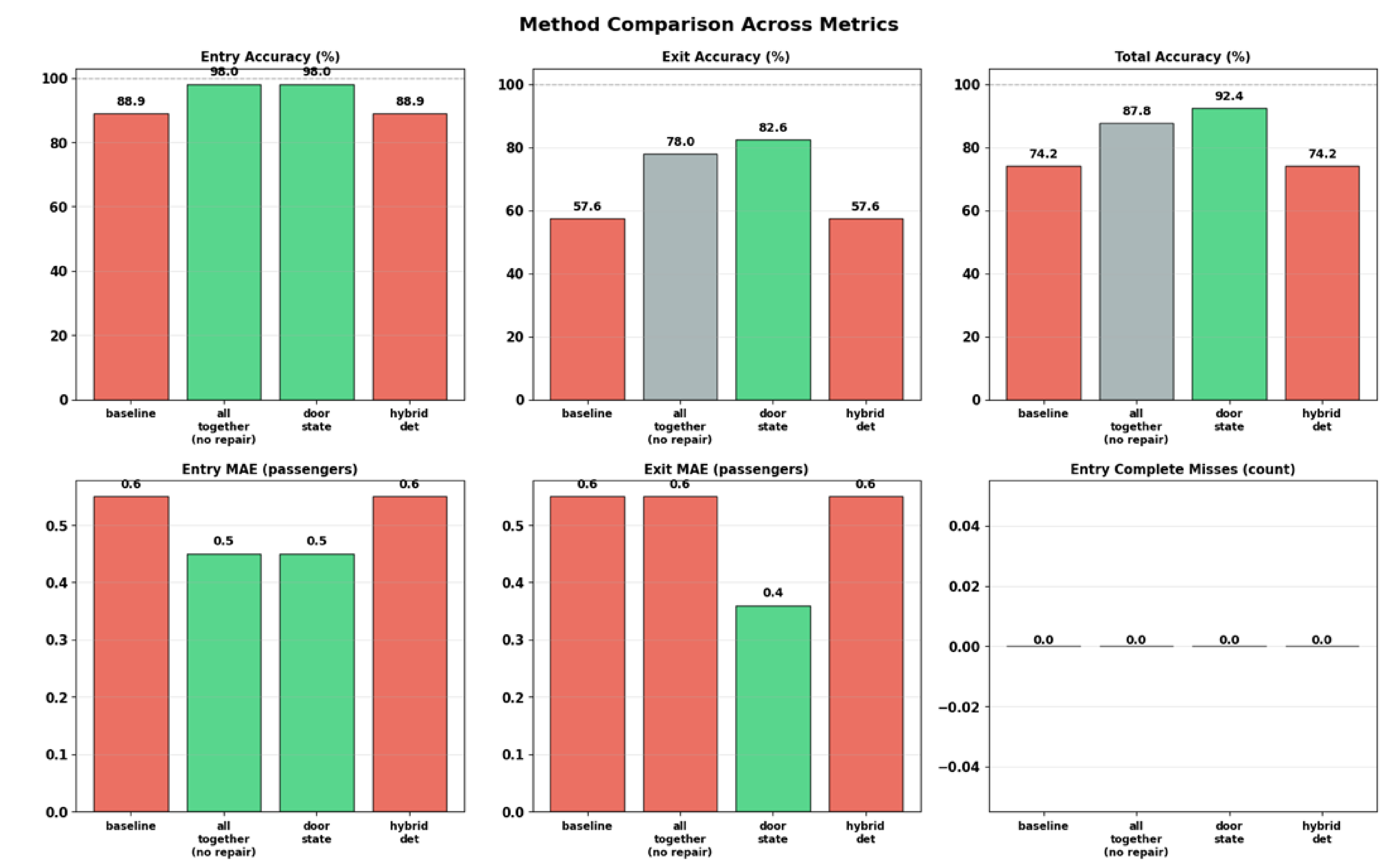


Table 1: Baseline pipeline accuracy under different operating conditions.

Condition	Detection	Tracking	OD Matrix Accuracy
Ideal (low density, good lighting)	40/42 correct, +4 FP	Stable IDs	Matches manual tallies
Overcrowding (peak hour)	Missed detections	Frequent ID switches	∼40% undercount
Poor lighting / shadows	Degraded confidence	Unstable IDs	Diverges at exits
Posture changes	Robust detections	Fragmented tracklets	Incomplete trips
Color ↔ B/W switching	Robust	Inconsistent features	∼17% degradation
Non-standard door use	Robust	Ambiguous direction	Noisy OD matrix

Deployment-Specific Challenges

- During overcrowding & occlusion at doors, only heads visible; tracklets fragment.
- Due to posture changes (standing, sitting, leaning) appearance embeddings drift; Re-ID continuity breaks.
- Modality shifts (color ↔ monochrome / IR) causes a drop in recall; embeddings unstable.
- Visual similarity (uniforms, similar clothing) cause identity switches; ambiguous matches.
- Boarding + alighting both doors cause a mismatch between direction of motion and door semantics.

Key Takeaways

- CCTV-only OD is feasible in low-resource settings when fused with telematics.
- Behavioral and operational context (door state, stop logic) is as important as model choice.
- Simple rules (door gating, near-zero speed with open doors) deliver big gains in exit-door accuracy.
- Failure under extreme crowding and modality shifts highlight the need for robust, deployment-aware Re-ID embeddings and trajectory-aware identity repair and motion-informed gating.

Next Steps

- Build stronger cross-camera, motion-aware gating + cost-optimal assignment for crowded doors.
- Controlled robustness studies by varying crowding and posture to quantify ID stability.
- Scale to full-day deployments and analyze equity impacts.
- Streamline the counter to be queue aware so that people waiting outside are not counted until they truly cross into or out of the vehicle.

Authors thank BasiGo for granting access to CCTV footage and operational data from buses in Nairobi and Kigali, UMass Amherst for institutional support, computing resources, and mentoring and WiML for providing travel support.