

Technical Memo: SHIRO AI

Restaurant Intelligence Platform

Cameron Kuperman Alex Tabaku Ben Tang Harshith Guduru

January 18, 2026

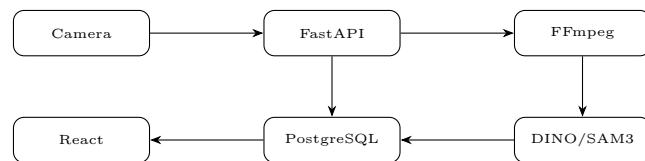
1 Executive Summary

Problem. Restaurant operations rely on manual table monitoring, subjective waiter assignments, and tedious scheduling—leading to inconsistent experiences and unfair tip distribution.

Solution. SHIRO AI automates three core challenges: (1) **Table Classification**—CCTV analysis classifies tables as *clean*/*occupied*/*dirty*; (2) **Waiter Routing**—fairness-first algorithm balances workload and tips; (3) **AI Scheduling**—constraint-aware engine with demand forecasting.

The system achieves 92%+ accuracy at 1 FPS, guarantees no waiter receives <50% of average tips, and targets Gini <0.25 for fair hours distribution.

2 System Architecture



Stack: FastAPI + SQLAlchemy (async REST), FFmpeg + OpenCV (video), PyTorch + HuggingFace (ML), React + Zustand (frontend).

3 ML Classification Pipeline

DINOv3 Classifier: Frozen ViT-B/16 backbone (768-dim) → Attention Pooling (CLS + weighted patches) → MLP (512→128→3 classes).

Technical Novelties:

1. **Group-Based Splitting:** GroupShuffleSplit by session+timestamp+table prevents CCTV frame leakage
2. **Attention Pooling:** Learns discriminative patches (plates, people) via Linear→Tanh→Softmax
3. **Imbalanced Handling:** Focal Loss $(1 - p)^\gamma$, class weights, Mixup ($\alpha=0.2$)

SAM3 Zero-Shot: Text-prompted segmentation—person mask >10% ⇒ occupied; plate mask >0.5% ⇒ dirty; else clean.

Temporal Smoothing: N-frame consensus (default: 5) reduces jitter by requiring consecutive agreement.

4 Data Annotation Pipeline

Custom workflow for rapid CCTV labeling:

1. **SAM3 Segmentation:** Generates table masks → computes *minimal rotated bounding boxes*
2. **Frame Extraction:** FFmpeg splits videos into per-table crops at configurable FPS
3. **Annotation GUI:** Keyboard-driven interface for fast clean/occupied/dirty labeling

Rotated bbox stored as center+size+angle or 4-corner coordinates for angled camera support.

5 Review Scraper

Yelp: Selenium + BeautifulSoup with undetected-chromedriver (anti-bot), pagination, “Read more” expansion.

Google: Selenium + API for location-based queries and star ratings.

Pipeline: Scraper → JSON → `bulk_ingest()` → PostgreSQL (dedupe by ID) → LLM sentiment analysis.

6 Waiter Routing Algorithm

Fairness-first scoring balances efficiency with equity:

$$\text{priority} = \underbrace{(\text{eff} \times 1.0)}_{\text{perf}} - \underbrace{\left(\frac{\text{tables}}{\max} \times 3.0 \right)}_{\text{load}} - \underbrace{\left(\frac{\text{tips}}{\text{total}} \times 2.0 \right)}_{\text{tips}} - \text{recency}$$

Underserved Override: Waiter with <50% avg covers *and* <50% avg tips bypasses recency penalty—guarantees no one is systematically disadvantaged.

7 AI Scheduling Engine

Score-and-rank with: (1) **Demand Forecaster**—weighted history with 0.85^{weeks} decay; (2) **Constraint Validator**—hard (availability, max hours) and soft (preferences); (3) **Fairness Calculator**—targets Gini <0.25; (4) **LLM Reasoning**—human-readable explanations.

8 API Reference

Method	Endpoint	Purpose
POST	/api/v1/videos/upload	Upload video (max 100MB)
POST	/api/v1/videos/{id}/process	Run ML classification
POST	/routing/recommend	Get waiter recommendation
POST	/schedules/run	Trigger AI scheduling
POST	/reviews/ingest	Bulk import reviews

Conclusion

SHIRO AI integrates computer vision, fairness-optimized routing, constraint-aware scheduling, and review analytics into a unified restaurant intelligence platform—improving operational efficiency, staff satisfaction, and customer insights.