

1. Mobile App (Expo/React Native)

Technology	Purpose
React Native + Expo	Frontend framework for cross-platform app development.
react-native-vision-camera	High-performance camera library for real-time video capture.
react-native-reanimated / react-native-canvas	Render bounding box overlays on the video stream.
Axios / Fetch API	Send frames to the backend API and receive overlay data.
sharp.js / expo-image-manipulator	Compress/resize frames (e.g., WebP/JPEG) before sending.
Zustand / React Context API	State management for overlay data and network status.
react-native-websocket	Optional: Persistent connections for real-time streaming.

2. Backend API (Python)

Technology	Purpose
FastAPI	High-performance Python framework for building the API (async support).
Uvicorn / Gunicorn	ASGI server to run FastAPI with concurrency.
OpenCV	Process frames (resize, decode, normalize) for AI models.
Redis	Cache frequent predictions or temporary frame data.
Celery / RabbitMQ	Optional: Queue management for scaling inference tasks.
PyTorch / TensorFlow	ML frameworks for car detection and brand classification.
Ultralytics YOLOv8	Pre-trained object detection model (cars).
MobileNetV3 / ResNet-18	Lightweight models for car brand classification.

Technology	Purpose
ONNX Runtime / TensorRT	Optimize inference speed (GPU/CPU).

3. AI/ML Pipeline

Technology	Purpose
Roboflow / Label Studio	Annotate and preprocess car brand datasets.
Kaggle / Stanford Cars Dataset	Source labeled car brand datasets.
TensorFlow Lite / PyTorch Mobile	Convert models to edge-friendly formats (future on-device inference).
OpenVINO	Optimize models for Intel CPUs (server-side).
Albumentations	Data augmentation for model training.
Weights & Biases / MLflow	Track model training experiments.

4. Networking & Optimization

Technology	Purpose
WebSocket / HTTP/2	Reduce connection overhead for frequent requests.
Protocol Buffers (Protobuf)	Efficient serialization for overlay data (smaller payloads).
NGINX	Load balancing and reverse proxy for the API.
AWS CloudFront / Cloudflare	CDN to reduce latency for global users.
FFmpeg	Optional: Dynamic video transcoding (server-side).

5. DevOps & Deployment

Technology	Purpose
Docker	Containerize the API and AI model for scalability.
AWS EC2 / Google Cloud	Host the backend API and AI models.
Firebase / AWS Amplify	Optional: Mobile app backend services (auth, storage).
GitHub Actions / GitLab CI	CI/CD pipelines for automated testing/deployment.
Prometheus + Grafana	Monitor server latency and API performance.
Sentry	Error tracking for mobile and backend.

6. Testing & Debugging

Technology	Purpose
Postman / Insomnia	Test API endpoints during development.
Detox / Appium	End-to-end testing for the mobile app.
pytest	Unit/integration tests for the Python backend.
Charles Proxy	Inspect network traffic between app and API.

Key MVP Priorities

1. Mobile App:

- `react-native-vision-camera` for frame capture.
- Compress frames to 720p/WebP before sending.

2. Backend:

- FastAPI + YOLOv8 (car detection) + MobileNet (brand classification).
- Use ONNX/TensorRT for model acceleration.

3. Networking:

- WebSocket for real-time communication.
- Deploy API on a cloud region closest to your users.