The Next Frontier in Embodied AI: Autonomous Vehicles

Engineering IB Paper 8 - Autonomous Driving | April 2022

Dr Alex Kendall, Co-Founder & CEO
1. Driving Intelligence
2. Sensors
3. Offboard software
4. Safety
Part 1: Driving Intelligence
2007 DARPA Urban Challenge

Commercial self driving car efforts

HD Maps
(brittle / slow to build / expensive to maintain)

LiDAR Sensors
(expensive / short lifespan)

Hand-Designed Rules
(rigid / clunky)
A PARADIGM SHIFT

Pioneering the next-generation AV architecture

- Solving self-driving with data
- End-to-end deep learning
- Lean sensors & compute
Deep learning has achieved superhuman performance in comparably complex settings to autonomous driving which are more accessible.
First wave of AI: virtual
Next wave of AI: physical
Autonomous driving is an embodied intelligence problem
Learning to drive in London

Autonomous driving from monocular cameras and end-to-end deep learning. No HD mapping, unnecessary sensors or hand-coded rules. Just pure intelligence.
AV 2.0: learning to drive with end-to-end deep learning and computer vision

**Driving Input**, $10^6$ dimensions

- Cameras (6 @ 25 Hz)
- GNSS
- Goal conditioning from standard sat-nav Map
- Vehicle State
- + others where required

**Driving Output**, $10^1$ dimensions

- Motion Plan
- Vehicle Controls

**Representation signal**

Learning signal for optimisation

Decoded human-interpretable intermediate representations

Semantics, geometry, motion prediction.
We first trained a driving model using only training data collected in London, UK. We then tested this model in five other UK cities, exposing it to diverse driving scenarios over a period of two months.

Unlocking new markets faster

MULTI-CITY GENERALISATION TEST
Quantifying driving domain differences between cities

Road Features Detected on New City Test Routes (indexed to a normal London routes)

- Traffic light
- Pedestrian crossing
- Bus lane
- Cycle lane

Road Density Detected on New City Test Routes (indexed to a normal London routes)

- Car detected
- Bus detected
- Pedestrian detected
- Cyclist detected
Learning to drive in London

First time driving in Coventry
Learning to drive in London

First time driving in Manchester
Part 2: Sensor Design on an Autonomous Vehicle
Sensing

The dominant sensor modalities used in robotics are:

Proprioceptive (internal state)
- Actuators (i.e., motor speed, position)
- Inertial Measurement Unit (IMU)

Exteroceptive (external state)
- Global Navigation Satellite System (GNSS)
- RADAR
- LiDAR (a.k.a. laser sensors)
- Cameras
Sensing: Actuators

Steering motor position (Electric Power Assisted Steering, EPAS)

Drive torque (motor current)

Wheel speed

Rotary speed/position encoders measure the motion of teeth past a sensor

Wheel speed

Pulse period
Sensing: Inertial Measurement

Microelectromechanical system (MEMS)
• Acceleration sensing (3D)
• Angular velocity sensing (3D)

IMUs are extremely useful, but suffer from drift over time
Sensing: Global Navigation Satellite System (GNSS)

Pros:
Global 2.5D positioning: [x, y, θ]

Cons:
~1-10m accuracy
Consumer-grade limited to ~5Hz
Requires 4+ satellites for a fix
Urban canyons hugely degrade GNSS performance: multipath effects + blocked signal
Sensing: Cameras

Typical camera:
~1-8MP, ~8-14 bit colour depth (Red, Green, Blue), 30-200Hz
Tradeoff: frame rate vs resolution – limited by serial data rate and compute

Monocular cameras

Optical flow, visual odometry, and localisation

Object detection, tracking, segmentation

Stereo Cameras

Depth sensing from a pair of images
Sensing: Radar

Pros:
Depth sensing robust to weather, lighting conditions, ~200m+ range, can ‘see through’

Cons:
Noisy, multipath effects

E.g., Continental ARS441

E.g., Waymo imaging radar visualisation
Sensing: LIDAR

Pros:
- Depth sensing robust to lighting conditions, very accurate with low noise
- 100-300m+ range pointcloud: [x, y, z]
- 0.3-10M points/second at 5-20Hz

Cons:
- Degraded by rain, snow
- Expensive (though improving)

E.g., Velodyne 3D lidar
Question: how much data does this autonomous vehicle collect per second?
Question: how much data does this autonomous vehicle collect per second?

6x forward facing cameras, 4x cameras per side: ~10MP images, 15Hz

= 150M pixels / sec / camera = 8,100 MB/s → 810 MB/s after compression

+ 6x 3D LiDAR: ~128 vertical * 3600 horizontal * 10Hz = 5M points / sec / lidar = 332 MB/s

+ 3x Radar: 256 channel * 512 depth * 30Hz = 4M points / sec / radar = 94 MB / s

Vehicle state (speed, location, etc) = minimal

= 1.2 GB / s ! (or 30+ PB / day !)
Computer Vision

3D geometric and semantic perception from surround vehicle monocular cameras
Part 3: Offboard Software

A large part of the complexity in developing an AV is off board the vehicle. AVs require substantial software platforms and tools:

- Data storage and log replay
- Machine learning training systems
- Validation and verification
- Fleet management
- Simulation
Simulation for Autonomous Driving

- Simulation is crucial to scale autonomy.
- Simulation offers limitless, realistic (visual, behavioural, and kinematic), unbiased, diverse training and evaluation samples.
- Simulation provides statistically significant insights at 1/10th the cost, 10x the speed, and 10x the repeatability compared to real-world testing.
- Simulation offers probably the only route to cost-effective deployment.
Visual Diversity
Part 4: Safety
Important Safety Concepts

- **Operational Design Domain (ODD):** the operating environment within which the autonomous vehicle can perform safely, described by the static elements, dynamic elements and environmental elements of the scene.

- **Safety Case:** a structured argument, supported by evidence, intended to justify that a system is acceptably safe for a specific application in a specific operational design domain.

- **Functional Safety:** Ensuring electronic failures will not cause unacceptable risks to human life.

- **Safety of the Intended Functionality (SOTIF):** Ensuring autonomous vehicle behaviour is absent of unacceptable risk.
## Safety != perfect: Case study on human performance

<table>
<thead>
<tr>
<th>Metric</th>
<th>Method of calculation</th>
<th>UK Human-Level Performance</th>
<th>Notes / Method of estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intervention rate (km / intervention)</td>
<td>total autonomous distance driven divided by number of valid interventions (ignoring platform-fault and end-of-run success interventions)</td>
<td>95,400 km / intervention</td>
<td>13.1% of vehicles make motor insurance claims per year (<a href="source">source, 2016</a>). Vehicles travel on average 12,500 km per year (<a href="source">source, 2016</a>). Therefore we can estimate distance per accident for UK drivers as 95,400 km / accident.</td>
</tr>
<tr>
<td>Speed Limit Compliance (%)</td>
<td>% of time the vehicle drives in excess of the speed limit provided by our map API</td>
<td>97.993 %</td>
<td>calculated from 100 hours of expert human driving data</td>
</tr>
<tr>
<td>Lane Following (km / int)</td>
<td>total autonomous distance driven divided by number of valid, non-junction interventions</td>
<td>221,400 km / intervention</td>
<td>43.09 % of total accidents occur during lane following. Assuming the distance driven by a car is negligible in junctions, this value is overall Intervention rate divided by 43.1% (<a href="source">source</a>).</td>
</tr>
<tr>
<td>Unprotected Intersections (%)</td>
<td>% of unsignalised intersections successfully navigated without any valid intervention within 10m before or after the junction</td>
<td>99.999238 % (every 130,000 intersections)</td>
<td>From the information about roundabouts below, and an estimate that roundabouts are ~40% safer than junctions (<a href="source">source</a>), we estimate that humans are 99.999238% successful at junctions.</td>
</tr>
<tr>
<td>Protected Intersection (%)</td>
<td>% of traffic lights successfully navigated without any valid intervention within 10m before or after the junction</td>
<td>99.999543 % (every 220,000 intersections)</td>
<td>There are ~25,000 roundabouts in the UK (<a href="source">source</a>) and 397,025 kms of road network (<a href="source">source</a>) in GB. Assuming the number of roundabout in the UK is approximately equal to GB (NI only has 2% of UK’s roads (<a href="source">source</a>)), there are 0.0629 roundabouts / km in UK. On average, vehicles encounter 2,533M roundabouts per year, considering 40,234M kms are driven per year (<a href="source">source</a>). Therefore, humans have a 99.999543% success at roundabouts, considering total UK roundabout accidents is 11,571 per year (<a href="source">source</a>).</td>
</tr>
<tr>
<td>Roundabouts (%)</td>
<td>% of roundabouts successfully navigated without any valid intervention within 10m before or after the junction</td>
<td>99.999543 % (every 220,000 intersections)</td>
<td></td>
</tr>
</tbody>
</table>
Swiss cheese model of Safety

Hazards

Losses prevented

Loss not prevented
Understanding epistemic model uncertainty: When and what we don’t know

(a) Input Image  (b) Ground Truth  (c) Semantic Segmentation  (d) Aleatoric Uncertainty  (e) Epistemic Uncertainty

Measuring Uncertainty for Autonomous Driving

Real world closed-loop testing

High uncertainty, **no intervention**

Intervention during **high uncertainty**

Average uncertainty around interventions
Conclusions

- AVs are the space race of our generation that promises to save millions of lives, transform cities, and make mobility ever more accessible.
- Embodied autonomy - taking AI out of the lab into the physical world - will make this possible and is the next major frontier in artificial intelligence.
- AVs are an incredibly rich and fascinating source of hard technical challenges across on-board robotics/AI and off-board software/tooling.
Interested in tackling the space race of our generation with Wayve?
careers@wayve.ai
Further reading:


