Institution: Durham University		
Unit of Assessment: 11 - Computer Science and Informatics		
Title of case study: Real-time Scene Understanding for On-Vehicle Automotive Sensing		
Period when the underpinning research was undertaken: 2014-2019		
Details of staff conducting the underpinning research from the submitting unit:		
Name(s): Toby Breckon	**Role(s) (e.g. job title):** Professor	**Period(s) employed by submitting HEI:** 2013 - present
Period when the claimed impact occurred: 2014 - 2020		
Is this case study continued from a case study submitted in 2014? N		

1. Summary of the impact

Our development of novel algorithms for automatic scene understanding for both on-road and off-road vehicles has enabled informed technology roadmapping and commercial product development for both driver assistance systems and vehicle autonomy ("driver-less cars"). This research, carried out between 2014 and 2019 at Durham University, has contributed to:

- the realisation and large-scale evaluation of the "Human-like Guidance" navigation concept for driver/vehicle interaction at **Renault**
- the design, test and evaluation process of several different sensor and algorithm options for on-road and off-road vehicle sensing systems at **Jaguar Land Rover (JLR)**
- the design, test and evaluation of a novel vehicle localisation (positioning) sensor that is now in commercial production with **Machines with Vision**
- the design and development of multi-camera 3D scene mapping sensor systems at **ZF Race Engineering - Conekt Engineering Services** (formerly TRW Conekt)

This directly informed the research and development at two of Europe's leading automotive manufacturers (France/UK: Renault / Jaguar Land Rover - annual combined R&D budget: ~GBP4.5 billion), enabled the supply of 3D scene mapping systems to the UK MoD (ZF) and supported the translation of vehicle localisation technology into rail, helping to protect 4.3 billion passenger journeys annually over ~57,000 km of track (Germany/UK: Machines with Vision).

2. Underpinning research

This research relates to the use of automated image understanding techniques for on-board vehicle sensing pertaining to both assistive driver technology (known as Advanced Driver Assistance Systems (ADAS)) and vehicle autonomy (i.e. "driverless cars"). Within this context we addressed the two key algorithmic tasks within on-vehicle scene understanding - *"Where am I?"* (known as the task of *localisation)* and *"What is around me?"* (known as the task of *semantic scene understanding).* The key challenge is to be able to address these tasks accurately, efficiently (i.e. in real-time relative to the vehicle speed) and robustly under varying environmental (weather) conditions for applications in both ADAS and vehicle autonomy.

In contrast to the prevailing trend of using an increasingly complex (and costly) array of sensors to support automotive sensing tasks, the Durham research team specifically targeted the use of low-cost camera sensors and research algorithmic approaches required to offer efficient and robust sensing under varying environmental conditions.

The key findings of this research at Durham were that:

- real-time object detection and classification from a single on-board camera can be extended to simultaneously provide high-level sub-categorical attributes (orientation, state, colour) using a single end-to-end convolutional neural network (CNN) architecture (processing at up to 10 fps) and this outperforms traditional hand-crafted feature-driven approaches for such tasks [R1].

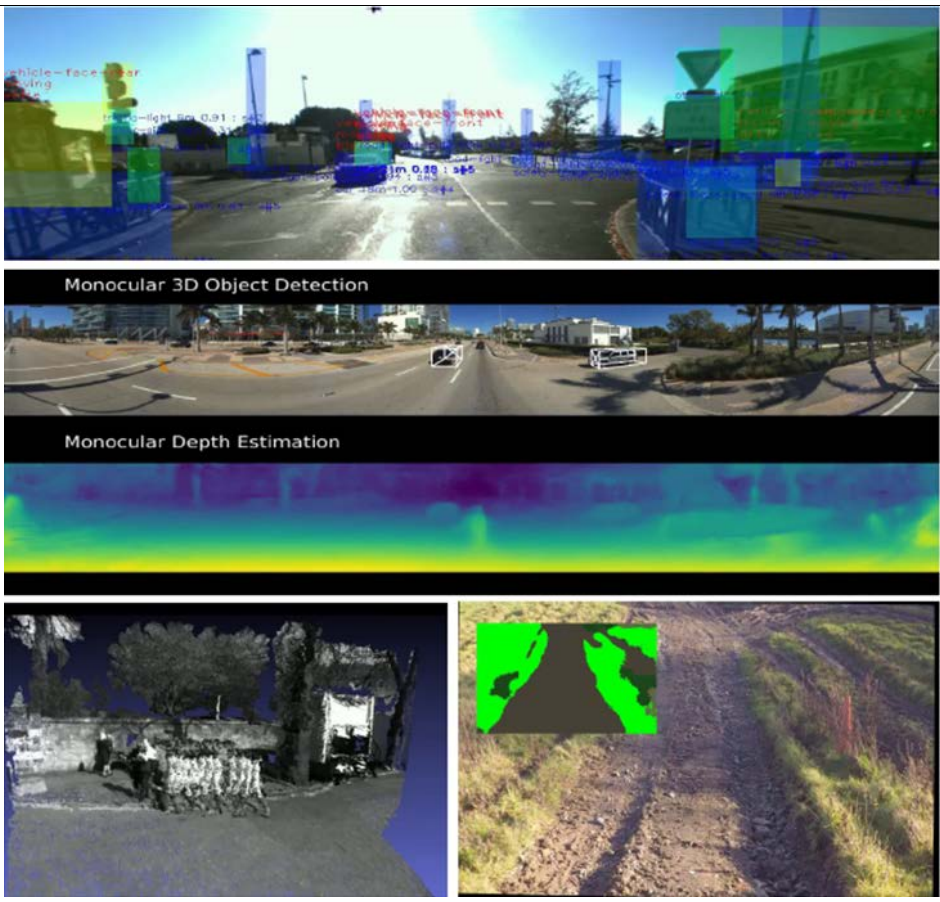

Figure 1: Algorithm output for object detection (top, [R1]), monocular 360° depth and 3D object detection (middle, [R6]}, stereo scene mapping (bottom left, [R2]) and off-road semantic scene understanding (bottom right, [R2, R4]).

- semantic scene segmentation via CNN, where each image pixel is labelled by its object category, can be successfully extended from the urban to off-road driving environment via the use of transfer learning over a very limited dataset and outperform traditional hand-crafted feature-driven approaches [R2]. Using stereo depth information for semantic understanding [R4] offers only a marginal improvement over a single monocular colour camera [R2].
- robust 3D scene mapping, of an urban driving environment, can be performed in the presence of dynamic scene objects (vehicles, pedestrians) based on the combined use of stereo visual odometry and structure from motion with an optimally calibrated real-time stereo depth estimation approach [R3].
- "end-to-end" autonomous driving (i.e. monocular image input → speed + direction output) can be extended to the off-road environment based on the use of stereo visual odometry as a methodology for prior recovery of human off-road driving behaviour in order to use this as an input to an "end-to-end" CNN approach [R5].
- real-time object detection and monocular depth estimation (from a single camera) can be extended to use 360° panoramic imagery, from low-cost consumer-grade spherical cameras, to offer both 3D object detection and ranging in addition to high granularity 3D scene depth around the entire vehicle as it transits an urban environment [R6]. In related work, the quality of the resulting 360° monocular depth recovered via [R6] is shown to be significantly superior to that obtained via the use of a conventional stereo vision pipeline with use of consumer-grade spherical (360°) cameras.
- the placement, mounting and inter-synchronisation of multiple low-cost sensors with other on-board auxiliary equipment can be readily optimised for the on/off road to offer effective vehicle situational awareness to support broader ADAS and autonomy functionality overcoming issues such as vibration and fouling [R1-R6].

3. References to the research

[R1]. On the Performance of Extended Real-Time Object Detection and Attribute Estimation within Urban Scene Understanding (K.N. Ismail, T.P. Breckon), In Proc. Int. Conf. on Machine Learning Applications, IEEE, pp. 641-646, 2019. [10.1109/ICMLA.2019.00117]

[R2]. From On-Road to Off: Transfer Learning within a Deep Convolutional Neural Network for Segmentation and Classification of Off-Road Scenes (C.J. Holder, T.P. Breckon, X. Wei), In Proc. Euro. Conf. on Computer Vision Wks, Springer, pp. 149-162, 2016. [DOI:10.1007/978-3-319-46604-0_11]

[R3]. Generalized Dynamic Object Removal for Dense Stereo Vision Based Scene Mapping using Synthesised Optical Flow (O.K. Hamilton, T.P. Breckon), In Proc. International Conference on Image Processing, IEEE, pp. 3439-3443, 2016. [http://dx.doi.org/10.1109/ICIP.2016.7532998]

[R4]. Encoding Stereoscopic Depth Features for Scene Understanding in Off-Road Environments (C.J. Holder, T.P. Breckon), In Proc. International Conference Image Analysis and Recognition, Springer, pp. 427-434, 2018. [https://doi.org/10.1007/978-3-319-93000-8_48]

[R5]. Learning to Drive: Using Visual Odometry to Bootstrap Deep Learning for Off-Road Path Prediction (C.J. Holder, T.P. Breckon), In Proc. Intelligent Vehicles Symp., IEEE, 2018. [10.1109/IVS.2018.8500526]

[R6]. Eliminating the Dreaded Blind Spot: Adapting 3D Object Detection and Monocular Depth Estimation to 360° Panoramic Imagery (G. Payen de La Garanderie, A. Atapour-Abarghouei, T.P. Breckon), In Proc. Euro. Conf. on Computer Vision, Springer, 2018. [https://doi.org/10.1007/978-3-030-01261-8_48]

Papers have been peer reviewed as part of the publication process and show clear originality and rigour. Publication time-line: some research publications delayed for commercial reasons.

4. Details of the impact

Background: With the global car market comprising ~80+ million vehicle sales per annum, the global autonomous vehicle development market is estimated to grow from USD54.23 billion (2019) to USD556.67 billion in 2026 (source: Allied Market Research). The UK Industrial Strategy has seen over GBP200 million invested in UK Connected Autonomous Vehicle (CAV) research to date (UK Department for Transport), including ~GBP1 million on projects involving Durham University, and rising industrial research budgets annually from major automotive manufacturers and component suppliers (2019: Renault (~EUR4 billion), Jaguar Land Rover (~GBP2 billion), ZF (~EUR2.6 billion)).

Commercial Impact: Our underpinning research has enabled and informed the development, evaluation and safety testing of CAV/ADAS sensing in the UK and beyond:

- **Renault** (France) used [R1,R3] to support the development of their *"Human Like Guidance" (HLG)* ADAS / CAV concept whereby the driver and vehicle communicate verbally based on a shared visual understanding of the road environment (i.e. with the vehicle acting as if it were a human co-pilot / e.g. vehicle → driver: "*Turn left before the red car on the right, just after the railings"*).

  In collaboration with Durham, Renault used real-time object detection techniques from [R1] to *"understand the performance characteristics of current state-of-the-art scene understanding techniques when applied to the extended requirements of HLG, including the ability to robustly detect a wide-range of road-side and urban environment objects and their attributes."* [E1] *"This informed in the down-selection of both algorithms and on-vehicle sensing hardware for HLG evaluation within Renault"* allowing Renault *"to demonstrate the initial HLG concept via a simulation based prototype, that included a real-time scene understanding module for object detection supplied by Durham, culminating in a successful presentation / demonstration to Renault-Nissan executives (2017). This resulted in strong positive feedback on the value to the end-user and an executive directive to accelerate development of the HLG concept within Renault." -* [redacted], Renault [E1]

  Subsequently, the Durham team co-developed a real-time software component to perform extended real-time object detection and attribute estimation [R1], coupled with low-cost dense stereo based object range estimation [R3] (Figure 1, top). This was supplied to Renault, under commercial contract with Durham, and installed onto on-vehicle GPU processing hardware with an integrated all-weather stereo camera rig, GPS and inertial measurement unit (Figure 2 – bottom middle/right). This enabled Renault to *"construct and demonstrate an integrated on-vehicle HLG prototype, incorporating a scene understanding module for real-time object detection and range estimation supplied as a combined hardware and software solution by Durham (2018). This resulted in on-vehicle realisation of the HLG concept within Renault"* [E1].

Figure 2: Exemplar on-vehicle sensor rigs used to support data collection and research under [R1-R6] in collaboration with Machines with Vision (left top/bottom), ZF Race Engineering (top middle), Jaguar Land Rover (top right) and Renault (bottom middle/right) – on test at Durham University.

This on-vehicle HLG prototype, constructed in collaboration with Durham, enabled Renault to *"perform extended proof of concept evaluation of the HLG concept using an integrated on-vehicle HLG prototype with 60 different test drivers on open roads around a common pre-defined evaluation route in Versailles. This enabled extensive on-road experimentation whereby the test drivers where guided via HLG voice navigation commands that were automatically derived from a list of visible scene objects (type, position, attributes) provided in real-time from the Durham scene understanding module (2018/19)."* [E1]

From this collaboration with Durham, *"The resulting experimentation and analysis of the HLG concept, enabled by our collaborative research work, has directly informed the vehicle design and development process within Renault with HLG now being considered within the production design (post-research) phase of Renault vehicles. From the HLG prototyping exercise with 60 different test drivers on public roads, Renault has identified and ranked several uses cases with their associated customer value. In this way, Renault has built a road map for the integration of theses uses cases in accordance with this associated customer value. In addition, Renault has registered 2 patents on vocal 'human like' guidance sentences [patent: FR3078565A1, 2018]"* in support of *"autonomous driving technologies to be available in 15 Renault models by 2022 as part of our current 'Drive the Future' strategic plan"* [E1].

- **Jaguar Land Rover's** (UK) development and evaluation of CAV in *"collaboration with ... Durham [based on R2, R4, R6], … resulted in research that has informed our ability to evaluate 3D depth mapping ..., scene segmentation and ability for accurately describing scene features. This enabled us to utilise this .. in .. investigating more feature-rich ADAS (Advanced Driver Assistant Systems), and to push forward vehicle autonomy."* - [redacted], Jaguar Land Rover [E2]

  This has enabled JLR to *"increase our understanding on ... the usefulness of vision systems on vehicles ... and the impact that they can have in autonomy."* [E2] *"The impact of this research … for driver assistance systems and vehicle autonomy, in both the off-road [R2, R4] and on-road [R6] environment has notably informed our internal research and development process for the range of .. options that may feature in production vehicles [patent: WO2018007079A1, 2016] ... and will have a large, positive impact ... towards autonomy."* (Fig. 2, top/right) [E2]

- **Machines with Vision's** (UK, Germany) development of a novel all-weather CAV localisation sensor in *"collaboration … [with Durham] has enabled … on-road testing of early low-TRL versions of our localisation sensor [patent: US20190265038A1, 2016] ... benchmarking against visual stereo odometry"* and facilitated development of *"a refined higher-TRL [Technology Readiness Level] version of RoadLoc, our automotive-specific localisation sensor, in collaboration with both Durham and Jaguar Land Rover"* [E3]. Based on the body of automotive research [R1-R6] underpinned by prior on-vehicle test and evaluation (see Figures 1 / 2 + [E3]), *"the research and industrial experience of the Durham team within the automotive sector, has directly contributed to the successful award and … delivery of ... research contracts by Machines with Vision."* (additional income to company: GBP436,000 – [E3])

  "Today, Machines with Vision… [has] a turnover of GBP547,789 (2020) and projected GBP1 million of revenue in 2021 ..., representing a significant growth from ... founding [in] 2016. … Without our research collaboration … [with] Durham we would never have had many of the insights and connections with the automotive sector, nor the ability to independently test and validate our sensor designs." - [redacted] Machines with Vision [E3].

  "Our ability to provide technical early-stage evidence of both on-vehicle testing and proof-of-concept benchmarking ..., in addition to the guidance on … sensor packaging / mounting / interfacing received from your team in the early days of our … localisation sensor, has been instrumental in contributing to ..." - approximately GBP600,000+ of additional company income [E3] and full commercialisation of a rail variant of the CAV localisation sensor**.** This will be fitted onto *"all **DB (Deutsche Bahn, German Railways)** measurement trains … by the end of 2019"* [E4]. This task was completed and now provides localisation in areas of poor GPS coverage (e.g. tunnels, cuttings etc.) and improved 2cm localisation accuracy over the entire network at increased train operating speeds [E4]. *"DB measurement trains"* are specifically equipped track survey inspection trains that are used to routinely monitor track condition in order to ensure railway safety across the ~41,000km of track in Germany that carries ~2.6 billion passengers annually (2019, source: DB). This is in addition to ongoing work *"with **Network Rail (UK)** on ... its measurement train (responsible for .. safety on all the UK's high-speed lines)"* that helps protect an additional ~1.7 billion passengers across ~16,000km of track in the UK.
- **ZF Race Engineering - Conekt Engineering Services** (UK) collaborated with Durham to use stereo-based 3D scene mapping from [R3] to *"develop and demonstrate 3D content generation from multiple cameras"* and to *"design and develop the 3D scene reconstruction from key-frame based photogrammetry"*- [redacted] (ZF), [E5]*. "This enabled the realisation of high resolution real-time 360° 3D information using low-cost camera sensing … for use in ... platform autonomy, ... path sensing and object avoidance"* [E5] using data obtained from the multi-camera rig constructed at Durham (Figure 2, top middle). *"This technical work, coupled with the 3D computer vision research expertise at Durham, contributed to the successful delivery of [2 projects to the]* ***Defence Science and Technology Laboratory [UK Ministry of Defence]"*** [E5] resulting in GBP164,000 of additional income to the company (2015/2016) and contributing to growth in the company patent portfolio in this area [patent: US20190182467A1, 2017].

<u>Summary:</u> Research from Durham [R1-R6] has enabled research and development at leading automotive manufacturers including patentable technology (France/UK: Renault, Jaguar Land Rover, ZF), enabled commercial sensor systems to be supplied to UK MoD (ZF) and supported technology translation into rail which now helps to protect ~57,000 km of track and ~4.3 billion passenger journeys annually (UK/Germany: Machines with Vision).

5. Sources to corroborate the impact

[E1]. Testimonial – Groupe Renault, [redacted] February, 2020.
[E2]. Testimonial – Jaguar Land Rover, [redacted] August 2019.
[E3]. Testimonial – Machines With Vision Ltd – [redacted], December 2020.
[E4]. Machines With Vision | Deutsche Bahn (DB) – website article (accessed, 14th Oct. 2019)
[E5]. Testimonial - ZF Race Eng. - [redacted] Aug. 2020