University of BRISTOL

BRISTOL VISION INSTITUTE

14TH VISION RESEARCHERS COLLOQUIUM

Friday, 19 June 2026

Life Sciences Building,
Bristol University

KEYNOTE

Prof Nicholas Roberts

Programme

09:30	Registration opens
10:00	**Welcome/ Registration / Tea and Coffee**
10:15	**Welcome to BVI** by Laszlo Talas, Bristol Vision Institute
10:20	**History of BVI** by Prof David Bull, Director of MyWorld, Bristol University
10:30	**Keynote:** Prof Nicholas Roberts, Head of School of Biological Sciences, Bristol University- **The Ecology of Vision**
11:30	**Tea/ Coffee and poster Presentations**
12:00	**Luca Courtney**- University of Bristol: **Dynamic Backgrounds Enhance the Confusion Effect**
12:15	**Taysa-Ja Newman -** University of Sussex: **Are the benefits of nature innate?: The role of familiarity and image properties in infant scanning efficiency for different environments**
12:30	**Craig Thompson** - Cardiff University: **Detecting tiny hands in the wild: Automated analysis of headcam videos from young children with and without Down syndrome**
12:45	**Shoaib Nabil-** University of Sussex: **The Role of Chromatic Information in Predicting and Perceiving Urban Inequality and Wealth**
13:00	**Lunch and Poster Presentations**
14:00	**Alfie Prettyman** - University of Bristol: **Movement patterns of layer hens in large groups as an indicator their welfare**
14:15	**Isaac Henderson**- University of Bristol: **A Novel, Child-Friendly Paradigm for Studying Visual Search Development**
14:30	**Saeid Moradipour**- University of Bristol: **Measuring the visual presence of gambling marketing in sports broadcasts**
14:45	**Zhe Gong**- University of Sussex: **Colour as a lived and context-dependent visual experience in autistic adults**
15:00	**Coffee and Poster Presentations**
15:30	**Ruirui Lin** - University of Bristol: **Low-light video enhancement under extreme noise**
15:45	**Cátia Oliveira** - Cardiff University: **The TinyExplorer Ecosystem: Open tools for studying infants' audio-visual experiences**
16:00	**Edoardo Tozzi**- University of Bristol: **Adolescent Vulnerability in Visual Persuasion: EEG Alpha and Beta Desynchronization During Exposure to Sports and eSports Gambling Ads**
16:15	**Ian Pennock**- University of Sussex: **Improving encoding model predictions across the visual cortex by incorporating color and eye's optics**
16:30	**Close-** Jasmina Stevanov, Bristol Vision Institute
16:35	**Informal drinks reception** (Sky lounge)

Keynote

The Ecology of Vision

Professor Nicholas Roberts

Head of School of Biological Sciences, University of Bristol

Abstract

From stomatopods to spiders, and from fish to moths, the Ecology of Vision explores how animals see and use visual information. Bristol has played a leading role in this field for almost 50 years, beginning with John Lythgoe's seminal book The Ecology of Vision, and continuing through the discoveries of Innes Cuthill and Julian Partridge on how ultraviolet light influences bird behaviour. More recently, we have shown a number of new ways in how animals use the polarization of light as a form of visual information, an ability normally hidden from human vision.

In this talk, I will highlight some recent discoveries that continue this Bristol tradition. Over the last few years, we have discovered how marine animals have evolved novel ways to control the flow of light to create near-perfect open-water camouflage. I will also show how information theory can provide intuitive insight into the form and function of visual systems, particularly for multi-task-based behaviours. Finally, I will showcase how new methods and human population models are helping us understand the global effects of light pollution at scales beyond individual behaviour.

Biography

Nick Roberts is a Professor of Sensory Ecology at the University of Bristol and is currently the Head of the School of Biological Sciences. His research spans many aspects of the Ecology of Vision, including visual physiology, biological optics and understanding the impacts of light pollution. He is particularly interested in how animals exploit optical physics for visually guided behaviours.

After undertaking an undergraduate degree in Astrophysics, Nick's interest in animal vision and biology developed during his PhD in the optics of soft matter systems. After his PhD, he held a number of research fellowships that enabled him to work in Canada at the University of Victoria and Queens University, before moving to the School of Biological Sciences in Bristol in 2009.

Presentations

Presentation 1

Luca Courtney- University of Bristol

Dynamic Backgrounds Enhance the Confusion Effect

Motion is ubiquitous in natural environments; examples include changes in illumination, wind-blown foliage, and flowing water. Such background motion has previously been shown to impede predation by masking the movement of a prey target. How dynamic backgrounds interact with movement of prey in groups remains unclear. We investigated whether dynamic backgrounds contribute to the 'confusion effect': reduced predation success resulting from difficulty in discriminating a target from its conspecifics ('distractors'). In computer-based experiments, human surrogate 'predators' tracked a single target among distractors, moving unpredictably on backgrounds that drifted unidirectionally. Tracking was poorer when backgrounds and objects moved at the same speed, consistent with temporary masking of prey movement and subsequent difficulty re-locating the target amongst distractors. A subsequent experiment investigated whether spatial matching between objects and their background further constrains this masking effect. Our findings identify a role for dynamic backgrounds in impeding successful predation of group-living prey.

Presentation 2

Taysa-Ja Newman - University of Sussex

Are the benefits of nature innate?: The role of familiarity and image properties in infant scanning efficiency for different environments

Nature is suggested to be inherently engaging, evoking 'soft fascination' (Kaplan & Kaplan, 1989), an idea often related to efficient processing of scenes (Franěk & Petružálek, 2025). Adults display varied scanning efficiency for scenes that differ in fascination (Berto et al., 2008). Processing efficiency across development may be influenced by familiarity (Fantz, 1964), but fascination in infancy and its relation to familiarity or image properties has not been explored. To investigate this, 40 infants (2–12 months) were eye-tracked viewing nature and urban images rated by caregivers for familiarity. Thirty adults rated images on fascination, naturalness, and restorativeness. Infants exhibited soft fascination (lower scanning efficiency) for images higher in naturalness, fascination, and restoration (highest $p = .012$, lowest $BF = 5.406$). Familiarity did not predict scanning efficiency ($p = .605$). Image properties underlying infant scanning and adult ratings were not identical, suggesting different visual properties predict soft fascination across development.

Presentation 3

Craig Thompson - Cardiff University

Detecting tiny hands in the wild: Automated analysis of headcam videos from young children with and without Down syndrome

Hand and body movements play a crucial role in creating a diverse visual catalogue for learning, particularly in children, where learning and understanding are constructed through visual and embodied experiences. Head-mounted cameras provide a unique insight into these behaviours in naturalistic environments; however, manually annotating hand-object interactions in egocentric video is extremely time-consuming and difficult to scale. This project expands an existing hand-object model by incorporating data from a large dataset of children with and without Down syndrome, while adding features relevant to understanding developmental behaviour. Current work demonstrates an improved hand detection F1 score of 86.5% compared with 77.9% in the baseline model, with hand ownership classification and in-hand object identification both above 83% across challenging egocentric headcam footage from children aged 2–60 months. This work demonstrates how computer vision can reduce the annotation burden and enable large-scale analysis of children's everyday visual and motor experiences.

Presentation 4

Shoaib Nabil- University of Sussex

The Role of Chromatic Information in Predicting and Perceiving Urban Inequality and Wealth

Economic inequality harms social cohesion (Gobel and Carvacho, 2024), but the environmental features signalling inequality are yet to be understood.
We captured >3000 colour-calibrated images from highstreets in London and extracted street-level economic inequality (Gini) and wealth (median house price) (Suss, 2023). Our image analysis showed that the low-level chromatic statistics of scenes can predict objective inequality and wealth.
We tested whether removing the chromatic signal would affect subjective judgements of the socio-economic conditions by presenting virtual tours in colour or greyscale. We found removal of colour does not affect such judgements. Moreover, participants (N = 80) were successful in predicting the overall wealth, but not inequality. Ratings on affect, engagement and greenness were significantly correlated with objective economic measures.
We conclude that although image properties can predict economic status, people may not be sensitive to these signals when making socio-economic judgements based on the appearance of a street.

Presentation 5

Alfie Prettyman - University of Bristol: Movement patterns of layer hens in large groups as an indicator their welfare

"Domestic hens experience welfare challenges not seen in wild populations, including feather pecking, smothering, restricted movement, and chronic stress. I have used neural-network based analysis of flock movement to quantify how hens distribute and interact in groups, and to identify environmental conditions associated with tight clustering, smothering events and injurious feather pecking.

Deep learning can reliably quantify flock distribution and movement in commercial settings, including large groups where individual identification is not feasible. Scaling these methods to commercial farms requires adapting models for tightly packed birds and low-light conditions, where bounding box detection is unreliable. Current work demonstrates that segmentation-based and group-level approaches are practical for extracting movement and space-use metrics from large flocks. These methods could easily be adapted for looking into enrichment use over time."

Presentation 6

Isaac Henderson- University of Bristol

A Novel, Child-Friendly Paradigm for Studying Visual Search Development

Visual search ability develops throughout childhood in neurotypical children. However, appropriate experimental paradigms to assess atypical development in naturalistic search are needed. Previous research uses simplified search paradigms, which, while maintaining high levels of experimental control, suffer from poor engagement, and reduced ecological validity and accessibility for atypical individuals, rendering these paradigms less informative for clinical practice. We developed a novel visual search task utilizing a set of 20 narrative-based, naturalistic scenes with child-friendly visual features. Semantic and feature cues were manipulated, and objective measures of visual clutter across search stimuli were defined. Eye-tracking data revealed significant increases in fixation counts and time to target fixation, and visual clutter was positively correlated with increases in these measures. Findings using our task will have important clinical implications, with comparisons between adult, neurotypical and atypical participants uncovering new insights into the developmental trajectories of individuals with a range of cognitive deficits.

Presentation 7

Saeid Moradipour- University of Bristol

Measuring the visual presence of gambling marketing in sports broadcasts

Gambling brands appear throughout broadcasted sports, on billboards, jerseys, and all over the playing field, but measuring their extent and frequency relies on trained coders who record each presence manually, taking dozens of hours. My PhD project asks whether computer vision can take on a meaningful share of this work without losing what makes content analysis valid: clear definitions, traceable decisions, and outputs that map to a defined codebook. I am building a two-stage pipeline that combines a YOLO object detector trained on an annotated sample with an OCR layer that detects text-based brands that the model was not trained on. My presentation will cover the approach, the trade-offs between automation and accuracy, and what this type of measurement can and cannot detect. Feedback is welcome.

Presentation 8

Zhe Gong- University of Sussex

Colour as a lived and context-dependent visual experience in autistic adults

Colour has been studied in autism research in terms of preference (e.g. Grandgeorge & Masataka, 2016), sensitivity and discrimination (e.g. Franklin et al., 2010), and adaptation (e.g. Maule et al., 2018), but less is known about how autistic adults experience colour in everyday environments. This doctoral research examines colour as a lived and context-dependent aspect of autistic adults' visual and sensory experience, with implications for sensory-inclusive design. The project uses a multi-method design, including an online questionnaire, photovoice, interviews, and participatory workshops. Initial findings suggest that colour is linked not only to liking or disliking particular hues, but also to visual comfort, overwhelm, mood, regulation, accessibility, and environmental usability. Participants' accounts also indicate strong individual variation, challenging one-size-fits-all assumptions in colour-related design guidance. This presentation will introduce the project framework and early findings, focusing on how everyday colour experience can inform vision science, autism research, and inclusive environmental design.

Presentation 9

Ruirui Lin - University of Bristol

Low-light video enhancement under extreme noise

Low-light video enhancement (LLVE) is challenging due to noise, low contrast, and color degradation. While learning-based methods enable fast inference, they often fail under heavy real-world noise because they do not sufficiently exploit long-term temporal cues. We propose DWTA-Net, a novel deep-learning recurrent LLVE framework with a recurrent design. DWTA-Net adopts an integrated two-stage architecture: Stage I restores local structure and color via multi-frame alignment for temporally consistent Mamba-based enhancement, while Stage II performs recurrent refinement using a novel dynamic weight-based temporal aggregation guided by optical flow, functioning as a recurrent denoiser that adapts to motion. We further introduce a texture-adaptive loss that preserves fine details in textured regions while suppressing noise in homogeneous areas. Experiments on real-world low-light footage show that DWTA-Net achieves stronger noise suppression and fewer artifacts, delivering superior visual quality compared with state-of-the-art methods.

Presentation 10

Cátia Oliveira - Cardiff University

The TinyExplorer Ecosystem: Open tools for studying infants' audio-visual experiences

"Studying infants' everyday learning environments remains difficult because existing methods often fail to capture the dynamic, multimodal, and socially embedded nature of children's experiences in naturalistic settings. Recent advances in technology, including more user-friendly hardware and AI-powered software, are creating new opportunities to study infants' audio and visual experiences in rich, scalable ways utilising head-mounted cameras (headcams). However, the widespread use of headcams in research is still limited by two major barriers: (a) developing safe, lightweight, child-friendly recording systems, and (b) the time-intensive process of manual audio and video annotation.

The TinyExplorer Ecosystem is an open-source platform designed to address these challenges by combining wearable egocentric recording hardware with locally run AI-powered analysis tools. The TinyExplorer Detection App currently integrates automated face detection and is being expanded to include hand detection, hand-object interaction analysis, and audio processing. Together, these tools support scalable, reproducible research on real-world learning environments. "

Presentation 11

Edoardo Tozzi- University of Bristol

Adolescent Vulnerability in Visual Persuasion: EEG Alpha and Beta Desynchronization During Exposure to Sports and eSports Gambling Ads

Adolescents are highly exposed to subtle gambling advertisements embedded in dynamic sports and esports content on social media, exploiting developmental differences in visual attention and persuasive processing. This consumer neuroscience study examined age-related differences using electroencephalography (EEG) and galvanic skin response (GSR) while adolescents (15–17 years) and adults (28–43 years) viewed gambling advertisements within a documentary context.
Key findings revealed stronger event-related alpha desynchronization (8–13 Hz) in adolescents, particularly over parietal-occipital regions during visually rich gameplay segments. This reflects heightened and more sustained visual attentional engagement and cortical resource allocation compared to adults, who showed weaker, more variable alpha responses.
By linking vision science with developmental neuroscience and marketing, this research demonstrates how immersive visual narratives in digital media exploit adolescents' immature attentional control. It provides novel neurophysiological evidence to inform media literacy, regulation, and cross-disciplinary interventions protecting youth from engaging visual marketing communication.

Presentation 12

Ian Pennock- University of Sussex

Improving encoding model predictions across the visual cortex by incorporating color and eye's optics

Most encoding models of fMRI responses to complex stimuli have greyscale RGB values as inputs, and the eye's optics are not modelled. To build biologically realistic encoding models of the visual cortex it is important to include these factors. Here we incorporated models of color and the eye's optics to a Gabor wavelet pyramid-based encoding model to predict BOLD responses in the visual cortex in the large 7T fMRI Natural Scenes Dataset (NSD). The NSD includes 8 participants who each saw 9,000-10,000 unique natural scenes. First, we built a luminance (L+M)-only encoding model as our baseline. To the baseline we added a model of responses in the retinogeniculate color channels, which improved encoding accuracy on average by 21% across V1-V4. We modelled the eye's optics using the ISETBio toolbox, which conferred small improvements of about 3% to prediction accuracy for BOLD responses. We conclude that accounting for color and the eye's optics in encoding models of V1-V4 BOLD responses to natural scenes improves prediction accuracy.

Posters

Salvador Alban, Cardiff University

Leveraging Perception Encoder for Crime detection in CCTV Footage

This project introduces a Weakly Supervised Video Anomaly Detection (WSVAD) framework evaluated on the UCF-Crime dataset. Because frame-level annotations are unavailable during training, the model employs a Multiple Instance Learning (MIL) strategy, learning to distinguish between normal and anomalous video bags using only video-level labels. Utilizing a pre-trained CLIP model, the system generates robust anomaly and normalcy memory embeddings from text prompts describing various scene types. During training, a Transformer-based temporal adapter processes 32-segment video features, refining them for comparison against these text-anchored memory banks via Top-K similarity retrieval. The network optimizes a combined ranking, sparsity, and temporal smoothness loss to accurately localize temporal events. By anchoring visual features to explicitly defined textual representations of crime and normal behaviour, the model achieves highly competitive frame-level Area Under the Curve (AUC) scores for anomaly detection.

Kane Colston, University of Bristol

Continuous Wildlife Monitoring in Remote Settings Using Low-Cost Camera Systems and GPUs for CNN-based Species Identification

Observation of wildlife in situ has long been a central aim of behavioural science and conservation, yet conventional camera traps often suffer from trigger bias, data loss, and limited applicability for species, particularly those lacking consistent infrared signatures. To address these issues, we developed a low-cost (<$150), Raspberry Pi-based remotely powered unit capable of continuous recording in remote environments. Field tests in La Pampa, Argentina, captured 445 hours of footage across six carcass sites, documenting scavenger arrivals and stay durations. We trained and validated a YOLOv8x object detection model to identify multiple species, achieving a mean average precision (mAP@50–95) of 0.746 and an F1 score of 0.906, with low error compared to manual counts. A lightweight YOLOv8n model deployed on a low-cost NVIDIA GPU performed comparably. This approach offers a scalable, affordable alternative to traditional camera trapping, enabling continuous monitoring and efficient AI-assisted analysis of wildlife behaviour in remote settings.

Clare Davis, University of Sussex

Mapping Real-World Aversive Visual Experiences in Autism

"Aversive visual experiences (AVEs) are frequently described by autistic people, yet experimental research often relies on laboratory measures that may not translate to real environments. To inform a more ecologically valid approach, we held a participatory workshop with autistic adults exploring lived visual experiences and evaluating existing visual sensitivity scales. Participants described current scales as reductive and lacking real-world relevance, while highlighting the importance of clutter, colour, movement, predictability and cumulative sensory load.

These findings informed development of the ongoing Visual Experience Study, which uses experience sampling and participant photography to investigate real-world AVEs in autistic and non-autistic adults. Reports of AVEs across everyday environments will allow us to map where and when they occur. Photographs of everyday scenes, rated for visual aversiveness, will be analysed to identify image statistics (e.g. spectral slope, colour contrast) that predict aversive responses across environments. This work will support a systematic approach to autism-inclusive visual design. "

Tadeáš Dvořák, Masaryk University

From Perceptual Fluency to Visual Discomfort: Cortical and Oculomotor Responses to Natural and Built Environments

The human visual system is evolutionarily optimized for the scale-invariant statistical properties of natural scenes. Conversely, built environments often deviate from these statistics, exhibiting unnatural distributions of spatial frequencies that induce visual discomfort. To investigate the neurophysiological basis of this discomfort, 67 participants viewed natural and urban scenes during a simultaneous fMRI, EEG, and eye-tracking measurement. We computationally extracted low-level image statistics from each stimulus, including luminance contrast and local chromatic differences. Our results reveal that unnatural image statistics underlie distinct oculomotor and cortical responses. High local chromatic differences in urban scenes increased saccadic exploration and metabolic demand in the early visual cortex. Furthermore, high luminance contrast significantly correlated with decreased occipital alpha oscillations, indicating reduced perceptual fluency and increased visual discomfort. Ultimately, these findings demonstrate how specific unnatural image statistics impose a measurable metabolic load on the visual cortex, contributing to visual discomfort.

Jago Hoope, University of Bristol

Digital Medication. Observer Support Systems: Biometric-Guided Perceptual Tools for Cognitive Support

This poster presents a suite of experimental systems exploring how observers construct stable perception from incomplete or degraded information. By combining behavioural measures, perceptual tasks, memory reconstruction tools, and biometric monitoring, the platform investigates how attention, uncertainty, and cognitive state influence perception and decision-making. Preliminary findings suggest that perceptual performance often transitions through distributed zones of uncertainty rather than a single collapse threshold. The long-term aim is to develop adaptive digital environments that respond to an individual's cognitive state in real time, supporting attention, memory, and cognitive wellbeing through personalized, data-driven interventions.

Zhuodong Jiang, University of Bristol

RUSplatting: Robust 3D Gaussian Splatting for Sparse-View Underwater Scene Reconstruction

Reconstructing high-fidelity underwater scenes remains a challenging task due to light absorption, scattering, and limited visibility inherent in aquatic environments. This paper presents an enhanced Gaussian Splatting-based framework that improves both the visual quality and geometric accuracy of deep underwater rendering. We propose decoupled learning for RGB channels, guided by the physics of underwater attenuation, to enable more accurate colour restoration. To address sparse-view limitations and improve view consistency, we introduce a frame interpolation strategy with a novel adaptive weighting scheme. Additionally, we introduce a new loss function aimed at reducing noise while preserving edges, which is essential for deep-sea content. We also release a newly collected dataset, Submerged3D, captured specifically in deep-sea environments. Experimental results demonstrate that our framework consistently outperforms state-of-the-art methods with PSNR gains up to 1.90dB, delivering superior perceptual quality and robustness, and offering promising directions for marine robotics and underwater visual analytics.

George Lewis, University of Bristol

An open-source thermal camera platform to monitor wildlife: a pilot study in thermoregulatory behaviour of European adders (Vipera berus)

Infrared thermography offers a passive, non-invasive way to detect and monitor animals regardless of light conditions. However, thermal cameras have historically been expensive and limited to handheld use, making long-term automated deployment difficult. We present a low-cost, modular sensing platform built around a Raspberry Pi, equipped with a thermal camera core, 3D-printed enclosures, and an open-source library for recording radiometric footage.

To test the camera, we monitored thermoregulatory behaviour of European adders in the Mendips. Adders were filmed basking on corrugated roofing sheets, with simultaneous RGB footage enabling species and sex identification. The system ran continuously for several weeks on exchangeable USB power banks, reliably

capturing radiometric data for quantitative analysis.
This pilot shows that low-cost, automated thermal monitoring in the field is achievable, with broader applications in behavioural ecology and non-invasive health monitoring.

Yixuan Li , University of Bristol

Q-Tacit: Image Quality Assessment via Latent Visual Reasoning

Vision-Language Model (VLM)-based image quality assessment (IQA) has been significantly advanced by incorporating Chain-of-Thought (CoT) reasoning. Recent work has refined image quality reasoning by applying reinforcement learning (RL) and leveraging active visual tools. However, such strategies are typically language-centric, with visual information being treated as static preconditions. Quality-related visual cues often cannot be abstracted into text in extenso due to the gap between discrete textual tokens and quality perception space, which in turn restricts the reasoning effectiveness for visually intensive IQA tasks. In this paper, we revisit this by asking the question, 'Is natural language the ideal space for quality reasoning?' and, as a consequence, we propose our method, a new paradigm that elicits VLMs to reason beyond natural language in the latent quality space. Our approach follows a synergistic two-stage process: (i) injecting structural visual quality priors into the latent space, and (ii) calibrating latent reasoning trajectories to improve quality assessment ability. Extensive experiments demonstrate that our method can effectively perform quality reasoning with significantly fewer tokens than previous reasoning-based methods, while achieving strong overall performance. This paper validates the proposition that language is not the only compact representation suitable for visual quality, opening possibilities for further exploration of effective latent reasoning paradigms for IQA. Source code will be released to support future research."

Joanne Lin, University of Bristol

ELVIS: Enhance Low-Light for Video Instance Segmentation in the Dark

Video instance segmentation (VIS) for low-light content remains highly challenging for both humans and machines alike, due to noise, blur and other adverse conditions. The lack of large-scale annotated datasets and the limitations of current synthetic pipelines, particularly in modeling temporal degradations, further hinder progress. Moreover, existing VIS methods are not robust to the degradations found in low-light videos and, consequently, perform poorly even after finetuning. In this paper, we introduce ELVIS (Enhance Low-Light for Video Instance Segmentation), a framework that enables domain adaptation of state-of-the-art VIS models to low-light scenarios. ELVIS is comprised of an unsupervised synthetic low-light video pipeline that models both spatial and temporal degradations, a calibration-free degradation profile estimation network (VDP-Net) and an enhancement decoder head that disentangles degradations from content features. ELVIS improves performances by up to +3.7AP on the synthetic low-light YouTube-VIS 2019 dataset and beats two-stage baselines by at least +2.8AP on real low-light videos.

Xiaolan Liu, University of Bristol

Embodied Agents for Human Intent and Behaviour Modelling

This work presents an early-stage embodied simulation platform for modelling human behaviour in smart home environments. Unlike conventional agent systems that follow mechanical plan-and-complete task sequences, our demo explores a more realistic closed loop between an AI agent, a behaviour controller, and a Unity-based visual environment. In a simulated home-working scenario, a virtual character, Alice, observes the environment, generates thoughts and behaviour descriptions, performs simple actions in Unity, and receives environmental feedback that informs subsequent decisions. The current demo includes interactive objects such as a computer, phone, sofa, and game, supporting behaviours related to focused work, distraction, rest, and task recovery. Although the embodied smart home simulation is still under development, this work demonstrates a pathway towards vision-aware human-centric digital environments. Future work will extend object states, sensor feedback, and text-to-motion generation to enable more natural embodied interaction.

Jiaqi Ma, Cardiff University

Through the Looking Glass: A Dual Perspective on Weakly-Supervised Few-Shot Segmentation

Meta-learning aims to uniformly sample homologous support-query pairs, characterized by the same categories and similar attributes, and extract useful inductive biases through identical network architectures. However, this identical network design results in over-semantic homogenization. To address this, we propose a novel homologous but heterogeneous network. By treating support-query pairs as dual perspectives, we introduce heterogeneous visual aggregation (HA) modules to enhance complementarity while preserving semantic commonality. To further reduce semantic noise and amplify the uniqueness of heterogeneous semantics, we design a heterogeneous transfer (HT) module. Finally, we propose heterogeneous CLIP (HC) textual information to enhance the generalization capability of multimodal models. In the weakly-supervised few-shot semantic segmentation (WFSS) task, with only 1/24 of the parameters of existing state-of-the-art models, TLG achieves a 13.2\% improvement on Pascal-5^{i} and a 7.9\% improvement on COCO-20^{i}. To the best of our knowledge, TLG is also the first weakly-supervised (image-level) model that outperforms fully supervised (pixel-level) models under the same backbone architectures.

Cameron Mackey, University of Bristol

Colouration and Patterns across Eubrachyura

There is huge diversity in the colouration and patterning of crabs, with some brandishing vibrant colours and complex patterns. These traits are thought to be involved with background matching, disruptive coloration, aposematism and intraspecific signalling. While there have been studies into individual groups and species, there has not been a wide-scale analysis of patterns and colouration. This study aims to better understand what factors influence the evolution of crab appearances through ecological correlates of habitat, morphology, and taxonomy. I explore associations between carapace coloration and substrate, red coloration and water depth, and conspicuousness and toxicity across more than 7000 species of Eubrachyura to highlight the multiple functions of crab coloration.

Peter Moore, University of Bristol

Developing earlier detection of bovine respiratory disease in dairy calves through body temperature changes using machine learning

Background: Bovine respiratory disease (BRD) is a highly prevalent syndrome in cattle worldwide. Temperature, coughing and behavioural changes are the primary clinical signs of BRD. Wearable technology is gaining popularity, yet can become expensive and unsustainable. It also tends to measure a single modality (e.g. movement). Contactless monitoring of calves for subtle clinical signs with configurable machine learning systems may provide an alternative to wearable technology.
Objective(s): To build a network of low-cost, open-source and customisable monitors (including thermal and RGB video, along with microphones) to detect potential cases of BRD earlier than current manual methods.
Methods: Monitors, comprised of a Raspberry Pi 4 computer collected data from calf pens. YOLO object detection models were trained to identify calf heads in RGB and thermal videos. BRD treatment was used as a proxy measurement of disease occurrence. Temperatures of head regions were analysed retrospectively from that date for changes that indicated when the febrile response began.
Results: Initial analyses show the set-up to be successful in constant monitoring of calves in pens. The YOLO models were able to repeatedly detect calf heads and calf head regions, allowing for the creation of a temperature profile for each calf that could be correlated with clinical disease.

Crispian Morris, University of Bristol

PocketDVDNet: Realtime Video Denoising for Real Camera Noise

Live video denoising under realistic, multi-component sensor noise remains challenging for applications such as autofocus, autonomous driving, and surveillance. We propose PocketDVDNet, a lightweight video denoiser developed using our model compression framework that combines sparsity-guided structured pruning, a physics-informed noise model, and knowledge distillation to achieve high-quality restoration with reduced resource demands. Starting from a reference model, we induce sparsity, apply targeted channel pruning, and retrain a teacher on realistic multi-component noise. The student network learns implicit noise handling, eliminating the need for explicit noise-map inputs. PocketDVDNet reduces the original model size by 74\% while improving denoising quality and processing 5-frame patches in real-time. These results demonstrate that aggressive compression, combined with domain-adapted distillation, can reconcile performance and efficiency for practical, real-time video denoising.

Caitlin Naylor, University of Bath

Co-developing a protocol for addressing body perception disturbances in CRPS using a novel VR tool to personalise hand appearance'

Individuals with Complex Regional Pain Syndrome (CRPS) experience body perception disturbances, where the appearance and feeling of their affected limb becomes distorted. Despite high heterogeneity in disturbances, clinical interventions typically adopt a one-size-fits-all approach. We co-developed a virtual reality (VR) tool that is tailored to individual body perceptions, enabling users to externalise and manipulate their own visual representations of the hand. Using a Patient and Public Involvement approach, four individuals with upper-limb CRPS created three embodied visual models: a typical hand, their perceived distorted hand, and a desired hand appearance. The tool enabled real-time visual personalisation, capturing perceived changes in size, shape, and structure of the limb. Embodiment of these virtual models altered body perception, pain and motor control, with highly individualised responses. These findings emphasise that targeting subjective visual distortions through personalised virtual embodiment offers a promising approach for assessing and reshaping disrupted body perception in CRPS.

Tianhao Peng, University of Bristol

Inverse Generative Modelling for Neural Image Compression

Learned image codecs face a long-standing trade-off: distortion-optimized designs preserve pixel fidelity but blur textures at low bitrates, while generative designs
recover perceptual sharpness at the cost of pixel fidelity, large model size, slow decoding, and a separate model per bitrate. We propose PerQ, a learned image codec that resolves this trade-off through a structural observation: the gap to a perception-optimal reconstruction lies within
a rate-determined ball around the backbone's output, given by a transport bound on quantization error. PerQ instantiates
this as a lightweight, rate-conditioned flow-matching compensator attached to a frozen multi-rate distortion-optimized backbone. From a single trained checkpoint spanning a wide bitrate range and decoding faster than existing generative codecs, PerQ leads all single-checkpoint codecs on perceptual metrics (LPIPS, FID, DISTS) and matches the strongest distortion-optimized codec on PSNR — a combination not previously demonstrated.

Yijia Song, University of Bristol

Diffusion-based InSAR Phase Unwrapping for Large-scale and Complex Remote Sensing Applications

Phase unwrapping is a fundamental problem in InSAR data processing, supporting geophysical applications such as deformation monitoring and hazard assessment. Its reliability is limited by noise and decorrelation in radar acquisitions, which makes accurate reconstruction of the deformation signal challenging. We propose a denoising diffusion probabilistic model (DDPM)-based framework for InSAR phase unwrapping, UnwrapDiff, in which the output of the traditional minimum cost flow algorithm (SNAPHU) is incorporated as conditional guidance. To evaluate robustness, we construct a synthetic dataset that incorporates atmospheric effects and diverse noise patterns, representative of realistic InSAR observations. Experiments

show that the proposed model leverages the conditional prior while reducing the effect of diverse noise patterns, achieving on average a 10.11% reduction in NRMSE compared to SNAPHU. It also achieves better reconstruction quality in difficult cases such as dyke intrusions.

Siyue Teng, University of Bristol

High-Fidelity Video Compression Based on Invertible Neural Transform and Implicit Conditioning

Learning-based video compression has recently achieved competitive rate-distortion performance compared with conventional video codecs. However, most existing methods rely on non-invertible analysis-synthesis transforms, where reconstruction quality is affected by both quantization and transform approximation errors. This becomes particularly limiting at higher quality settings, where quantization errors are small and transform-induced distortion becomes dominant. To address this, we propose an invertible neural network-based video codec for wide-range and high-fidelity compression. The key idea is to preserve an invertible main transform path before quantization, while introducing content-adaptive context through a compact implicit conditioning field. This separates strongly correlated video content from harder-to-model fine details, allowing different components to specialize in complementary reconstruction tasks for more efficient compression. To further improve compressibility, we introduce a scheduled masking strategy that progressively concentrates informative content into fewer latent channels for more effective entropy coding.

Katherine Ward, Cardiff University

Revealing cortical organisation: visual acuity and eccentricity in infantile nystagmus

"Early in life, the visual cortex (V1) adapts to the type and quality of visual input it receives. In infantile nystagmus (IN), abnormal eye movements disrupt the stable formation of images on the retina, suggesting radical changes in V1 organisation.
This study aimed to investigate V1 organisation in IN by deriving visual acuity (VA) as a function of retinal eccentricity and calculating 'E2' values – the retinal eccentricity at which VA doubles relative to the fovea – to model cortical magnification.
VA thresholds were measured psychophysically for 7 participants with IN and 5 controls using 2-alternative forced-choice Gabor stimuli at 0°, 5°, 10°, 15° and 20° eccentricity. Linear regression lines were fitted to estimate E2 for each participant.
E2 values were significantly larger in those with IN than controls (P=0.006), indicating reduced cortical magnification. This V1 adaptation may limit the visual potential of individuals outside of neuroplastic periods, supporting earlier intervention for any therapeutic approaches."