2021 AIMLAC, Bangor University projects

AI to engender Fast Visualization Ideation Design

1st supervisor: Professor Jonathan C. Roberts / School Computer Science Electronic Engineering
2nd supervisor: Dr Panagiotis (Panos) Ritsos / School Computer Science Electronic Engineering

Ideating new creative visualizations takes much time; often design sketching is used and many alternative ideas are sketched. While recently researchers have been developing tools to help users create visualizations, they focus on developing a final visualization design (Roberts et al 2020), rather than capturing or helping the user in the ideation process. Tools like Tableau's show-me, D3, Lyra, Keshif and the Improvise environment enable different visualisations to be crafted, but it is less easy to create or manage new ideations. Dashboards can be crafted in many ways, using D3.js, high-charts.com, datahero.com and in tools such as Tableau, SAS or Power BI, but again do not help the user analyse, collaborate, or share their ideas of the ideation process. Researchers have started to make smart design interfaces (e.g., Chen et al 2020), which help designers create new data-visualizations, recommend design ideas, encourage good design, suggest ideas and return to previous ideas (perhaps forgotten). This research will investigate current visualization design strategies such as the Five Design-Sheet method (Roberts et al. 2016), multiple views (Roberts et al 2019), deep learning, ideation techniques, recommender systems, and version control, to make the ideation process smart. The work will investigate how to mix smart guidance lead by AI recommendations with version control, and incorporate learnt behavior into visualization ideation process, which will encourage novel designs to be explored and captured.

- Chen X, Zeng W, Al-Maneea HMA, **Roberts** JC, Chang R. 2020, Composition and Configuration Patterns in Multiple-View Visualizations. IEEE Transactions on visualization and computer graphics.
- **Roberts**, J.C., C Headleand, and P.D. Ritsos. "Sketching Designs Using the Five Design-Sheet Methodology". IEEE Trans on visualisation and Comp. graphics. 2015, 22(1). 419-428. https://doi.org/10.1109/TVCG.2015.2467271
- **Roberts**, JC, Al-Maneea, HMA, Butcher, P, Lew, R, Rees, G, Sharma, N & Frankenberg-Garcia, A 2019, 'Multiple Views: different meanings and collocated words', *Computer Graphics Forum*, vol. 38, no. 3, pp. 79-93. https://doi.org/10.1111/cgf.13673

Edge-based object recognition for immersive analytics in Web-based XR

1st supervisor: Dr Panagiotis (Panos) Ritsos / School Computer Science Electronic Engineering
2nd supervisor: Prof Jonathan C. Roberts / School Computer Science Electronic Engineering

Project description: We are Increasingly being immersed in a technology-mediated world, where the omni-presence of data introduces increased needs in mechanisms facilitating *in-situ* cognition, reasoning and sensemaking [2, 3]. In parallel, edge computing, facilitated by future networks, such as 5G, is transforming the way data is being processed and delivered from millions of devices around the world, bringing computing and analytics close to where the data is created [1]. Building on these synergies, this project will investigate the use of edge-based object recognition using distributed neural networks (DNN), as a mechanism for *in-situ* registration and data processing for mobile, Web-based Immersive Analytics (IA) in Extended Reality (XR). Object-recognition can provide accurate and real-time registration [1], yet its practical application still faces important challenges. Current object-recognition

systems are either self-contained, or cloud-based, yet face low latency and poor user experience respectively. Deep Learning, and DNNs, can provide effective solutions for object detection, and ameliorate these challenges [1]. In addition, they have the potential to provide adaptive MR interfaces, and multimodal sensing capabilities [3], useful for advanced IA experiences [2].

[1] P. Ren, X. Qiao, Y. Huang, L. Liu, S. Dustdar and J. Chen, "Edge-Assisted Distributed DNN Collaborative Computing Approach for Mobile Web Augmented Reality in 5G Networks," in *IEEE Network*, vol. 34, no. 2, pp. 254-261, March/April 2020, doi: 10.1109/MNET.011.1900305.

[2] P. W. S. Butcher, N. W. John and **P.D. Ritsos**, "VRIA: A Web-based Framework for Creating Immersive Analytics Experiences," in *IEEE Transactions on Visualization and Computer Graphics (Early Access)*, 2020 doi: 10.1109/TVCG.2020.2965109.

[3] J. C. Roberts, **P.D. Ritsos**, S. K. Badam, D. Brodbeck, J. Kennedy and N. Elmqvist, "Visualization beyond the Desktop--the Next Big Thing," in *IEEE Computer Graphics and Applications*, vol. 34, no. 6, pp. 26-34, Nov.-Dec. 2014, doi: 10.1109/MCG.2014.82.

FLOOD-AI: Using Artificial Intelligence to Investigate the Impact of Land Management Decisions on River Flood Risk

1st supervisor: Dr Sopan Patil / School of Natural Sciences
2nd supervisor: Dr Panagiotis Ritsos / School of Computer Science and Electronic Engineering

Project description: Hydrological models are essential tools for simulating streamflow in river basins and are widely used for understanding, and forecasting, a river's flood response to storm events. However, appropriate application of hydrological models requires a priori calibration of parameters using historical measured streamflow data. Previous research has shown that the relationship between hydrological model parameters and physical river basin properties (e.g., topography, soils, land use) is too complex to characterize using traditional statistical models. This limits the ability to determine how parameter values will modify if land use change alters the physical structure of a river basin. Recent advances in Artificial Intelligence (AI), specifically in Deep Learning (DL), have resulted in the ability to provide efficient high-dimensional interpolators that can handle data of multiple dimensions and heterogeneous information, such as those encountered in hydrological modelling. In this project, our goal is to develop AI techniques that can help improve the ability of hydrological models to predict the impact of land use change on river flood risk. Specifically, we propose a novel use of AI and information visualization to interactively relate hydrological model parameters to the physical properties of river basins. Our approach will involve development of DL techniques to extract high level abstractions in the hydrological model and physical river basin data, which can be used to test the impact of land management decisions on river flood risk. This abstraction will be made available to end-users via an interactive visualization interface to facilitate the flood risk investigation of multiple scenarios of land management changes (e.g., increase in urbanization by 10%). Our training dataset will include data from >1000 river basins across the UK, and the coupled AI-hydrological modelling workflow will be streamlined to operate on Supercomputing Wales High-Performance Computing (HPC) framework.

(1) Patil, S. and M. Stieglitz, "Modelling daily streamflow at ungauged catchments: What information is necessary?", Hydrological Processes, 28(3), 1159-1169, 2014.
(2) Patil, S. D., Y. Gu, F. S. A. Dias, M. Stieglitz, and G. Turk, "Predicting the spectral information of future land cover using machine learning", International Journal of Remote Sensing, 38(20), 5592-5607, 2017.
(3) S. Rizou, K. Kenda, D. Kofinas, N. M. Mellios, P. Pergar, P. D. Ritsos, J. Vardakas, K. Kalaboukas, C. Laspidou, M. Senožetnik, and A. Spyropoulou, "Water4Cities: An ICT platform enabling Holistic Surface

Water and Groundwater Management for Sustainable Cities," in Proceedings of 3rd EWaS International Conference, Lefkada, Greece, 2018.

Predicting the "Relative" Coastal Weather and Conditions.

1st supervisor: Peter Robins / School of Ocean Science
2nd supervisor: Matt Lewis / School of Ocean Science

Project description:
Using AI and ANN (artificial neural networks) to develop a novel Met Ocean prediction tool, with end users training the algorithm's assessment of a user's impression of condition and risk. Outcome is a new "app" that users download and use to make decisions (with constant confirmation of "forecast accuracy"); using coastal forecasts and the user's assessment of the quality through ANN and AI to determine relative forecast based on user's opinion of previous forecasts. Reduced levels of risk and maintenance costs in coastal activities can thus be estimated. Impact will include integration with two world-leading companies in the Operation and Maintenance realm of offshore energies: Turbine transfers (www.turbinetransfers.co.uk/) and MetOcean solutions (www.metocean.co.nz/). Current Met-Ocean condition forecasts have focused on the physical conditions for users of the coastal environment; for example, wave height and period, wind strength and direction. In the coastal zone, recreational users and industry currently have to make decisions using this forecast information, alongside their own experience and skills level, including their appetite for risk (e.g. a two water-sport enthusiasts may have completely different views of the quality of conditions due to many other factors; and this has led to wide spread perception of forecast inaccuracy at the coastal zone). Indeed, offshore wind farm maintenance operatives have different boats and levels of acceptable risk, typically based on wave steepness and wave-tide interaction processes that are not included in forecast models due to the computational burden. Such human interpretation of forecast data often falls within classical heuristic risk traps settings. Current physical Met-Ocean forecast model products have recognized this change in user requirements, but are still processed based forward stepping physical models – with a new focused on uncertainty quantification to improve the user's perception of accuracy. This novel application of computer learning could reduce uncertainty and improve confidence in weather forecasting.

Lewis, M.J., Palmer, T., Hashemi, R. et al. Wave-tide interaction modulates nearshore wave height. doi.org/10.1007/s10236-018-01245-z

Hashemi, M.R., et al.. and Lewis, M., 2016. An efficient artificial intelligence model for prediction of tropical storm surge. Natural Hazards, 82(1), pp.471-491.

Ensembles of Deep Neural Networks for Semi-supervised Learning.

1st supervisor: Prof Ludmila Kuncheva, School of Computer Science and Electronic Engineering
2nd supervisor: School / Franck Vidal, School of Computer Science and Electronic Engineering

In semi-supervised learning, some of the data have labels but most of the data is unlabelled (Engelen and Hoos, 2020). This type of data is widespread because labelling is often infeasible, destructive, or too expensive. Consider as an example a sample from a population affected by a pandemic. A small

proportion of the people in the sample would have been tested for the disease, returning a positive or a negative output. This will be the labelled part of the data. The rest of the sample will be unlabelled but could still be of great help in extracting an accurate proxy for the test, which can then be deployed to the whole population. Deep learning has hijacked the research in machine learning and pattern recognition for a good reason - the undisputed success of large-scale models for complex data. Ensemble models are known to be more accurate than single models, which explains the interest in ensembles of Deep Learning Neural Networks (DLNN). Ensembles of DLNN for semi-supervised data are just taking off, with a few, heuristically crafted models. Considerable amount of knowledge on standard classifier ensembles has been accumulated in the past twenty years. Grounding our work in that, we will attempt to improve on the state-of-the-art in semi-supervised learning. Our analysis will gauge the need for a DLNN ensemble based on data size and characteristics. We will examine the contribution of diversity within the DLNN ensemble to the ensemble performance as well as the extent of usability of unlabelled data. New ensemble methods will be devised based on robust training and testing protocols, a suite of classifier combination methods, and a novel idea about feature transformation. In the first instance, we will illustrate the proposed methods on videos related to animal monitoring.

1. Van Engelen, J.E. and H, H. Hoos, A survey on semi-supervised learning. Machine Learning 109, 373–440 (2020). https://doi.org/10.1007/s10994-019-05855-6
2. Laine, S and T Aila, Temporal ensembling for semi-supervised learning- arXiv preprint arXiv:1610.02242, 2016 - arxiv.org https://arxiv.org/abs/1610.02242
3. Kuncheva L.I. Combining Pattern Classifiers. Methods and Algorithms, Wiley, 2nd edition, 2014

Bringing big-data to social science

1st supervisor: Simon Willcock / School of Natural Sciences
2nd supervisors William Teahan / School of Computer Science and Electronic Engineering;) & **Prof Jonathan Roberts** / School of Computer Science and Electronic Engineering

Project description: We live in a period of unprecedented data availability, but not all data are equal – with quantitative data sometimes viewed as more reliable, robust and/or useful than qualitative data. This is particularly problematic when conducting the interdisciplinary research necessary to address the most important global challenges, such as climate change and sustainable development. For example, the benefits human's derive from nature are categorized into three types: provisioning services (products obtained from ecosystems; e.g. food), regulating services (benefits obtained from the regulation of ecosystem processes; e.g. regulation of air quality), and cultural services (non-material benefits people obtain from ecosystems; e.g. cultural heritage or spiritual enrichment). Whilst both provisioning and regulating services can be quantified at local and global scales, cultural services are often viewed as 'unquantifiable', being spatially and temporally distinct, intangible, subtle, mutable and intuitive in nature, based on ethical and philosophical perception – thus largely unique to the individual. As such, most nature-based research is dominated by the relatively easily quantified provisioning and regulating services, which are readily monetized to enable comparisons across services. The same is not true of cultural services and how to combine these data to holistically value nature's contributions to people is unknown. We seek to address this here.
Using existing data from seven national surveys across Wales (~1000 respondents per survey; 3 surveys complete [Jan-Jun '20], 1 ongoing, 3 planned in Jan-Jun '21). Using Supercomputing Wales, we will use bespoke Natural Language Processing (NLP) to analyse the quantitative data within these surveys,

understanding how people's reasons for spending time in greenspace change from before, during and after the ongoing coronavirus crisis. These qualitative data contain free text responses in both English and Welsh, and our cross-language analysis would compare responses to see if there are any specific language differences, as well as differences between genders and socioeconomic groups. Finally, advanced visualization techniques will be developed to enable the comparison of the qualitative free text responses and quantitative survey data, which includes distance travelled and length and regularity of visits. The ability to visualise both quantitative and qualitative data at national-scales may transform sustainable decision-making.

- Willcock, S., Camp, B. J., & Peh, K. S. H. (2017). A comparison of cultural ecosystem service survey methods within South England. Ecosystem Services, 26, 445-450.
- W.J. Teahan. 2018. "A Compression-Based Toolkit for Modelling and Processing Natural Language Text", *Information*, Vol. 9, No. 294. MDPI Publishers. doi:10.3390/infoxx010001.
- Rick Walker, Llyr ap Cenydd, Serban Pop, Helen C Miles, Chris J Hughes, William J Teahan, and Jonathan C Roberts. 2013. Storyboarding for visual analytics". *Journal of Information Visualization*.

Optimization of co-located offshore wind and wave energy arrays

1st supervisor: Prof. Simon Neill
2nd supervisor: Dr David Christie (SEEC research fellow) and TBA

Co-locating wave energy with offshore wind developments leads to synergies which can improve efficiency and reduce cost, including the advantages of a common consenting process, shared grid connection, and logistics. Power output can be smoothed by the phasing between wave and wind peaks within the same weather footprint, consideration of the swell component of waves, and further improved by strategic (micro-siting) within the array to modify the phase relationship between individual Wave Energy Converters (WECs). By absorbing wave energy, the WECs can reduce wave loading on the wind turbine structures, which in turn may modify the wave resource by reflection and diffraction. Energy yield, power quality, loading forces and cabling cost will depend on the spatial configuration of the devices within the array. This project will apply genetic algorithms to determine the optimal device configuration. The student will use a coupled wind/wave resolving model running on a Supercomputing Wales to determine the effect of WECs and wind turbines on the surrounding wave and wind fields. Combining with wind and wave data, boundary conditions from global models will allow an estimation of annual energy yield for a given configuration. A cost function will be synthesized from estimation of a variable contribution to the infrastructure cost to represent wave loading forces, as well as cabling costs (which will be affected by both the internal layout and phase relationship). The number of devices and spatial footprint will be constrained by optimization with respect to levelized cost.

- Neill, S.P. and Hashemi, M.R. 2018. Chapter 9 - Optimization, Fundamentals of Ocean Renewable Energy – Generating Electricity from the Sea, Academic Press, Pages 237-270.
- Neill, S. P., Vögler, A., Goward-Brown, A. J., Baston, S., Lewis, M. J., Gillibrand, P. A., ... & Woolf, D. K. (2017). The wave and tidal resource of Scotland. Renewable Energy 114, 3-17.
- Neill, S. P., Hashemi, M. R., & Lewis, M. J. 2014. Optimal phasing of the European tidal stream resource using the greedy algorithm with penalty function. Energy 73, 997-1006.