Coming to terms with AI

BY ROD MCNEIL

A machine might be called intelligent if its response to questions could convince a person that it was human, a test proposed by Alan Turing in 1950 [1]. The author considers potential applications of artificial intelligence (AI) using machine learning and deep learning techniques in clinical ophthalmology.

"Computer science is about automating stuff, and artificial intelligence is about automating everything"

Deep learning pioneer Alex Krizhevsky, Principal Machine Learning Architect at Toronto-based AI company DESSA (Source: https://qz.com/1307091/)

Deep learning models have demonstrated robust classification performance for major eye diseases and interpretation and synthesis of electronic patient record data based on deep learning techniques shows promise [2,3]. Automated image interpretation for screening, referral decision-making and patient monitoring is likely to play a routine role in frontline eye care, in part to address staggering outpatient appointment demand and reduce false positive referral rates.

New research on AI and biomarkers for eye disease was highlighted in presentations during a news conference at the 2019 annual meeting of the Association for Research in Vision and Ophthalmology, including AI-based approaches for detection of diabetic eye disease from fundus photography, automated characterisation of choroidal neovascularisation (CNV) activity in optical coherence tomography angiography (OCTA) images and estimation of haemoglobin A1c (HbA1c) from retinal photographs.

Deep learning can automatically detect severity and presence of diabetic eye disease from colour fundus photographs only

A proof-of-concept study demonstrated for the first time that a deep learning model could automatically detect the severity or presence of diabetic macular oedema (DMO) from colour fundus photographs only [4]. The deep learning algorithm was based on highly curated datasets from the phase III RISE and RIDE clinical trials (sample of more than 700 patients and ≥17,000 images). The deep learning model successfully identified colour fundus photographs with presence of central subfield thickness ≥250μm and ≥400μm. Results suggest that deep learning models could potentially be used to augment diabetic retinopathy (DR) screening programmes by detecting the presence and quantifying the severity of DMO, thereby enhancing referral patterns.

Machine learning for automated detection and characterisation of CNV activity over time in SS-OCTA images

Volumetric characteristics of CNV are difficult to quantify in swept source-OCTA images due to difficulties in making volumetric manual annotations. To address this challenge, Dr Sisternes and colleagues developed a method to automatically detect the presence of CNV and to segment and quantify CNVs in a volumetric manner in SS-OCTA images [5]. Automated characterisation of CNV activity may provide a promising approach to monitoring CNV activity in patients over time, noted study authors.

Detection of vascular and metabolic diseases from retinal images

Retinal images may provide information on systemic vascular and metabolic diseases. Tham and colleagues undertook a study to evaluate the performance of a newly developed deep learning system in estimating HbA1c from retinal photographs [6]. Results showed, in the validation dataset, the overall mean error between deep learning system-predicted and actual serum HbA1c measurements was 0.87%, the Bland Altman plot mean difference was 0.18% (95% Limits of Agreement, -2.39 to 2.76%), indicating fair agreement with slight overestimation by the deep learning system. Implications include the potential development of a deep learning system matched with smartphone technology that allows home monitoring of diabetes.

Wong et al. developed and validated a deep learning system with a convolutional neural network using retinal images to detect chronic kidney disease [7]. Results support the feasibility of an AI deep learning system to automatically detect chronic kidney disease from retinal photographs.
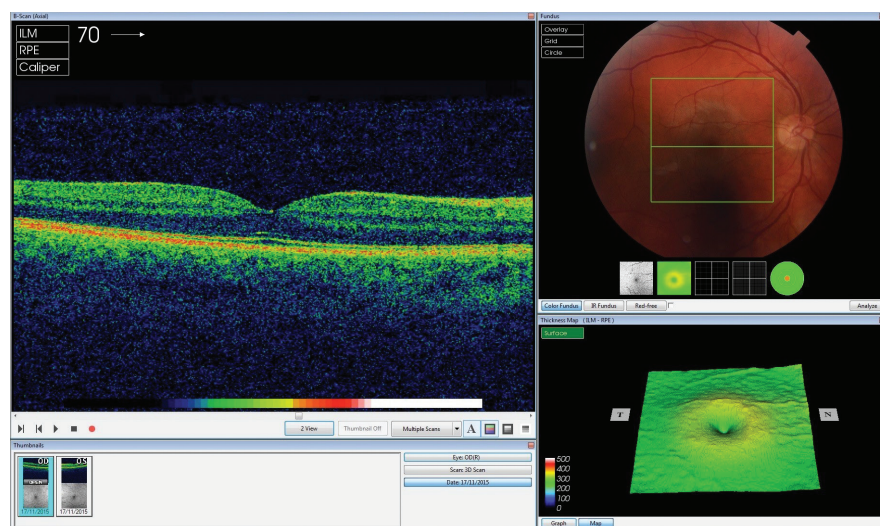

Figure 1: Analysing an OCT scan. *Source: Courtesy of DeepMind Technologies.*
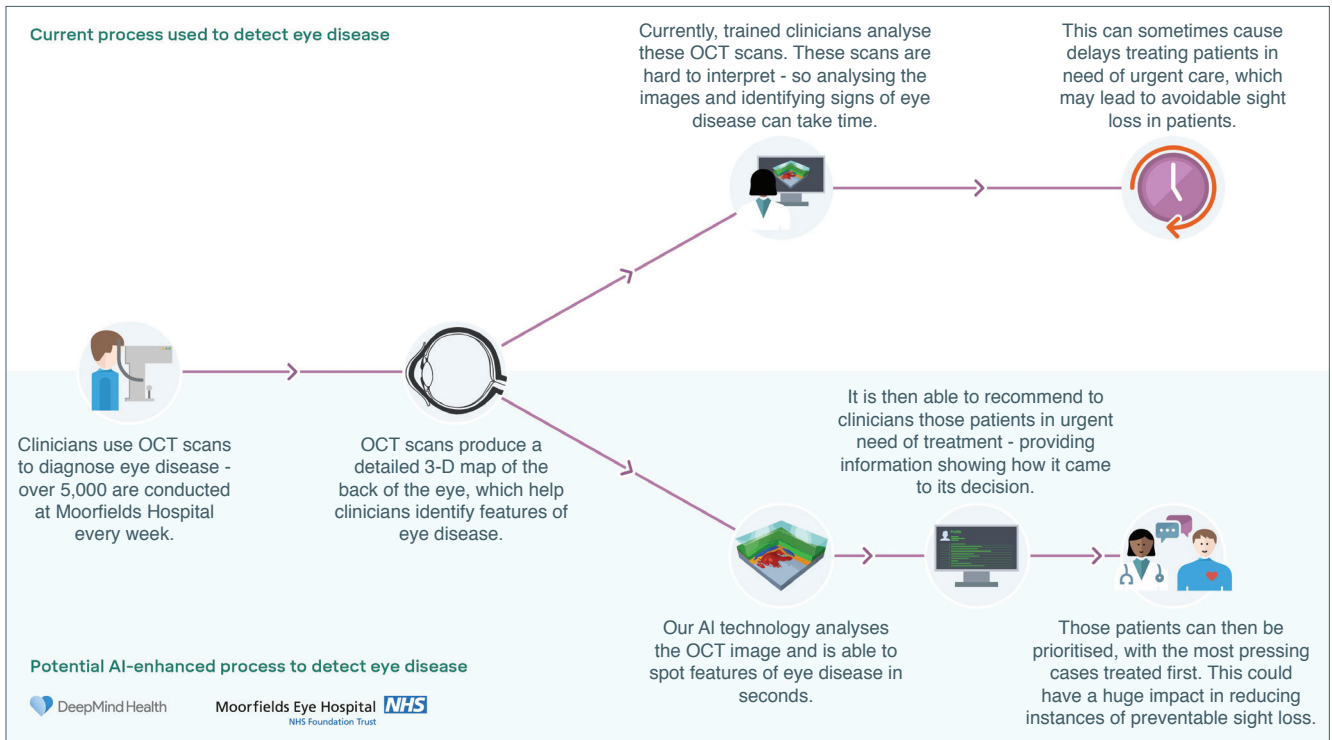
Figure 2: Illustration of how an AI-enhanced process could help clinicians detect eye disease. *Source: Courtesy of DeepMind Technologies.*

AI-based diagnostic system for detection of referable DR in primary care settings

IDx-DR (IDx Technologies Inc.) is the first AI-based diagnostic system with US Food and Drug Administration (FDA) clearance and a Class IIa CE mark for the autonomous detection of referable DR. Assessment for referable DR uses a standard two-image per eye protocol (one optic disc centred image and one macula centred image from a fundus camera with at least 1000 by 1000 pixels per image), and the IDx-DR delivers test results within one minute. In a pivotal clinical study of 900 participants, the device performance for accurately detecting the presence of more than mild DR was equal to or better than human graders, demonstrating observed sensitivity for more than mild DR of 87.4%, with observed specificity of 89.5% [8]. The FDA evaluation noted that the high accuracy of the IDx-DR makes the potential risk of false negatives low.

IDx is also developing algorithms that can detect disease across a number of imaging modalities, focusing on fundus photography and OCT, including signs of age-related macular degeneration (AMD) from standard retina images, as well as prototype algorithms to detect and track glaucoma indicators, Alzheimer's disease, cardiovascular disease and stroke risk.

Automated medical image interpretation and triage

As part of a joint research partnership between Moorfields Eye Hospital and Google DeepMind, researchers applied a novel deep learning architecture to a clinically heterogeneous set of OCT scans from patients referred to the eye hospital with symptoms suggestive of macular pathology [9]. The system's performance in making a correct referral recommendation reached or exceeded that of clinical experts on a wide range of sight-threatening retinal diseases, with an accuracy of 94% (overall error rate, 5.5%). Referral accuracy was maintained when using tissue segmentation from a different scanning device type.

The deep learning framework features two separate neural networks: a segmentation network that creates a detailed device-independent tissue segmentation map and a classification network that analyses this map and provides diagnoses and referral suggestions (Figures 1-3). The training set for the classification
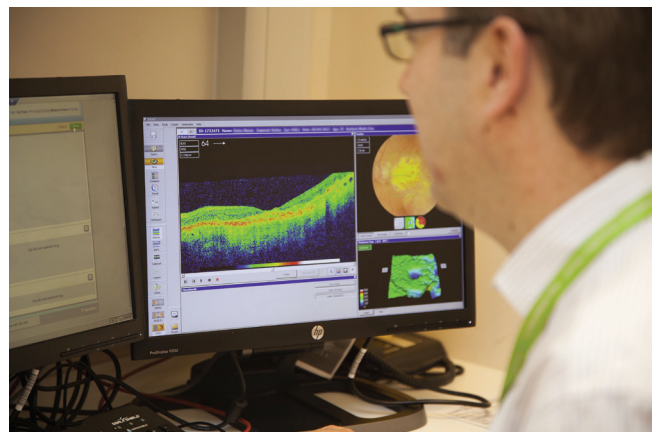

Figure 3: Consultant Ophthalmologist Pearse Keane, Moorfields Eye Hospital, London, analysing an OCT scan. *Source: Courtesy of DeepMind Technologies.*

network was 14,884 OCT scan volumes of 7621 patients. The system also explains its decisions: first, with an interpretable segmentation or visual 'map' of pathology features identified on the OCT images and second, generating predicted probabilities for the level of confidence it has in its diagnoses and referral suggestions.

Another joint AI technology project involves analysis of OCT scans of up to 7000 patients previously treated for unilateral nAMD, to try to predict CNV conversion in the fellow non-affected eye. This may ultimately lead to the development of an AI-based system that can help identify imminent CNV converters and provide reassurance to those considered at low risk of converting at two years.

Toward image-based clinical management of patients with early / intermediate AMD

Grechenig and colleagues from the Medical University of Vienna proposed a fully-automated deep learning method capable of segmenting drusen in OCT in a precise and reproducible manner [10]. The authors noted the need for advanced medical image computing methods that can objectively measure distinct changes in drusen morphology to assess conversion risk to later AMD stages

"The single greatest threat to reliance on data-driven technology is the actual or possible presence of bias"

and prognosis determination.

'Deciphering AMD by deep phenotyping and machine learning' (PINNACLE) is a 5-year European AI study of patients with AMD, funded by the Wellcome Trust. The study aims to teach computers to analyse high resolution retinal images, to identify what eye changes appear in patients with AMD, and to identify the structural changes that lead to and are associated with cell degeneration in the retina in patients with early AMD.

Combined image-based diagnosis and interpretation of EHR data for rapid, accurate and total diagnoses

An image-based deep learning framework using transfer learning techniques demonstrated diagnostic performance comparable to that of human experts in classifying retinal OCT images for the presence of CNV and DMO [12]. A convolutional neural network trained on an ImageNet dataset of common categories was adapted to significantly increase the accuracy and shorten the training duration of a neural network trained on an input dataset of retinal OCT images. Occlusion testing confirmed that the network made its decisions using accurate distinguishing features from the input images.

"Using a transfer learning technique reduced the amount of high-quality OCT images required to train the deep learning system to provide accurate diagnosis of common retinal diseases requiring urgent specialist referral," explained study investigator and lead contact Professor Kang Zhang, Institute for Genomic Medicine, Institute of Engineering in Medicine, and Shiley Eye Institute, University of California, San Diego, La Jolla, CA, USA, speaking with the author in a telephone interview.

The same transfer learning technology was applied for paediatric pneumonia detection using chest x-ray images, accurately distinguishing bacterial and viral pneumonia, achieving high accuracy of 92.8%.

Prof Zhang added that the application of AI-based machine learning classifiers beyond image-based diagnoses also shows promising potential, for example to evaluate high-volume electronic patient record data to detect trends in clinical features and diagnoses.

Investigators applied an automated natural language processing system using deep learning techniques to extract clinically relevant information from electronic health records (EHRs) [3]. Data points derived from over 1.3 million paediatric outpatient visits presenting to a major referral centre in China were analysed to train and validate the AI framework. The neural network demonstrated high diagnostic accuracy across multiple organ systems and was comparable to experienced paediatricians in diagnosing common childhood diseases.

"The benefits of such an AI system in analysing massive amounts of EHR data to augment diagnostic evaluations and provide clinical decision support could likely be applied across medical specialties, including ophthalmology," observed Prof Zhang. "A combination of automated image-based diagnosis and interpretation of EHR data based on deep learning techniques might potentially be utilised to synthesise the entire clinical feature classifications and provide rapid, accurate and total diagnoses mimicking a human physician."

The potential role of AI and deep learning in glaucoma

Research suggests that automated optic disc assessment may help discriminate glaucomatous disease from no disease. MacCormick and colleagues developed a novel glaucoma detection algorithm capable of accurate, fast and interpretable glaucoma diagnosis based on automated image segmentation and analyses of colour photographs [11]. The algorithm was tested on two separate image datasets, ORIGA and RIM-ONE. The spatial probabilistic algorithm, using a dataset 100-times smaller than that required for deep learning algorithms, accurately distinguished glaucomatous and healthy discs on internal and external validation (area under the receiver operating characteristic curve 99.6% and 91.0% respectively).

Investigators plan to evaluate the algorithm using data from the Ocular Hypertension Treatment Study and the UK Glaucoma Treatment Study to assess diagnostic ability and progression detection. The ultimate hope is that this spatial model could be utilised for smartphone disc photography assessment and used in conjunction with clinician assessment to increase diagnostic ability.

Guidance and standards on AI in health and secondary use of health data

A position statement on AI in health from the Royal College of Physicians recommends a transparent approach to explaining the evidence base for new technology, ensuring testing takes place using real-life data that is diverse enough to represent the intended population [13]. Where possible, findings should be subject to peer-reviewed publication. Healthcare managers and industry should also ensure that electronic record systems adhere to existing information standards.

With respect to clinical uses, the Royal College of Physicians states: "Context and conversation, particularly around difficult decisions, is a crucial part of determining treatment plans that can only be supported, not replaced, by technology." New technology should support physicians in taking a more person-centred approach to care. High-quality data is needed to ensure the conclusions drawn from AI are valid and safe. This applies in particular to the use of clinical data recorded in electronic records by health professionals. Clinical headings in the record need to be standardised so that data from different organisations can be integrated without changing the meaning.

A code of conduct for data-driven health and care technology from the UK government provides guidance on behaviours expected from those developing, deploying and using such technologies, including AI techniques [14]. Developers should show what type of algorithm is being developed or deployed, the ethical examination of how the data is used, how its performance will be validated and how it will be integrated into health and care provision. The single greatest threat to reliance on data-driven technology is the actual or possible presence of bias, notes the guidance, recommending that any commercial arrangement needs to identify where this manifests and how it is managed, by whom, and at what cost. Also, outputs should be transparent and explainable.

What next?

Moves to substitute clinicians with machines for selected first-line hospital eye services may not be far off. In June 2019, NHS Chief Executive Simon Stevens announced a global call for evidence from technologists for how the NHS can best incentivise the use of carefully targeted AI and machine learning technologies across the NHS from April 2020 and beyond.

Ophthalmologists emphasise that robust clinical validation as well as multi-class classification capabilities will be required to support general adoption of AI solutions within ophthalmology practice.

Professor Tunde Peto, Queen's University Belfast, discussing potential clinical applications of AI with deep learning programs in a debate at the 2019 Annual Congress on Controversies in Ophthalmology in Dublin, commented: "We are moving towards a better understanding of what AI and deep learning can achieve in ophthalmology. But AI is not quite there yet in terms of validated automated clinical applications for detecting and / or grading DR, AMD, glaucoma and other ophthalmic disorders."

Prof Peto said that automated applications using AI offer significant potential benefits in clinic settings, particularly for high-volume retinal screening and triage and for monitoring of ocular pathology, adding: "Human input is absolutely essential in determining clinical relevance, as pragmatic common sense is needed when deciding whether to treat or observe. Also, treatment modalities change over time and management practice needs to adapt or adjust accordingly."

Remote evaluation of retinal images using deep learning image assessment systems may prove effective for automated detection of a range of referable eye disorders. Expected clinical applications include deployment in rapid access macular clinics, digital surveillance pathways for stable diabetic eye disease patients, maculopathy referral refinement clinics, glaucoma monitoring clinics, as well as community-based tele-ophthalmology screening programmes.

TAKE HOME MESSAGE

- Deep learning models have demonstrated robust classification performance for major eye diseases.
- Application of AI-based machine learning classifiers beyond image-based diagnoses shows promising potential, for example to evaluate high-volume electronic patient record data to detect trends in clinical features and diagnoses.
- Automated applications using AI offer significant potential benefits in clinic settings, particularly for high-volume retinal screening and triage and for monitoring of ocular pathology.
- Outputs should be transparent and explainable, permitting real-time interaction.
- Robust clinical validation as well as multi-class classification capabilities are required to support general adoption of AI solutions within ophthalmology practice.

JARGON BUSTER

Artificial intelligence: The science and technology of making machines smart (artificially intelligent) and intuitive. AI describes advanced technologies that allow machines to carry out complex tasks effectively – tasks that would require intelligence if a person were to perform them, according to the Academic Health Science Network.

Machine learning: A form of artificial intelligence that allows computer systems to learn from examples, data and experience. Through enabling computers to perform specific tasks intelligently, machine learning systems can carry out complex processes by learning from data, rather than following pre-programmed rules. Machine learning algorithms are programs that tell a computer how to learn to solve a problem.

Bayes' theorem: A theory that specifies how to handle uncertainty by updating the probability for a particular event, phenomenon or hypothesis in response to data.

Bias (sampling): Selection of data or samples in a way that does not represent the true parameters (or distribution) of the population. Bias in training data leads to bias in algorithms: machine learning is a data driven technology and the characteristics of the data are reflected in the properties of the algorithms.

Big data: Large and heterogeneous forms of data that have been collected without strict experimental design. Big data is becoming more common due to the proliferation of digital storage, greater ease of data acquisition and the higher degree of interconnection between devices.

Convolutional neural network (CNN): A model designed to incorporate prior knowledge to compensate for all the data not included in the use of a smaller training dataset. Compared with standard feedforward neural networks with similarly-sized layers, CNNs have fewer connections and parameters and so are easier to train (using transfer learning techniques).

Deep learning: A machine learning method which composes details together to obtain more abstract, higher level, features of the data through composition of mathematical functions. Powerful deep learning algorithms often involve a large number of these levels.

Neural network: A complex mathematical system for identifying patterns in images or data.

Reinforcement learning: An approach to machine learning in which an agent learns to interact with its environment, receiving inputs, and making sequential decisions so as to maximise future rewards. An important feature in this context is that it is often only after the agent makes a number of decisions that it learns of the payoff resulting from the set of choices. One challenge in reinforcement is thus to work out which of the decisions were 'good' and which less so.

Supervised learning: An approach to machine learning which relies on training data that has been labelled, often by a human. A label could be a categorisation into one or more groups: this is known as classification.

Test data: Data that is used to test the functioning of a machine learning system or verify its outputs.

Training data: Data that can be used to train machine learning systems, having already been labelled or categorised into one or more groups.

Unsupervised learning: An approach to machine learning that uses data which has not been labelled. Commonly it will seek to determine characteristics that make the data points more or less similar to each other and will attempt to represent the data in a summary form, such as through clusters or common features.

According to DeepMind, unsupervised learning is a paradigm designed to create autonomous intelligence by rewarding computer programs for learning about the data they observe without a particular task in mind.

Source (unless otherwise stated and excluding descriptions of CNN and neural network): The Royal Society. Machine learning: the power and promise of computers that learn by example. April 2017. Report available at: www.royalsociety.org/machine-learning

References

1. Turing A. Can machines think? *Mind* 1950;**59**:433-60.

2. Ting DSW, Pasquale LR, Peng L, et al. Artificial intelligence and deep learning in ophthalmology. *Br J Ophthalmol* 2019;**103(2)**:167-75.

3. Liang H, Tsui BY, Ni H, et al. Evaluation and accurate diagnoses of pediatric diseases using artificial intelligence. *Nat Med* 2019;**25(3)**:433-8.

4. Willis JR, Arcadu F, Benmansour F, et al. Deep learning predicts OCT measures of diabetic macular thickening from color fundus photographs. Poster presentation at the 2019 Annual Meeting of the Association for Research in Vision and Ophthalmology, April 27-May 02, 2019, Vancouver, Canada. Abstract No. A0222.

5. Sisternes LD, Bagherinia H, Makedonsky K, et al. Automated volumetric choroidal neovascularization segmentation and quantification in swept-source OCT angiography using machine learning. Poster presentation at the 2019 Annual Meeting of the Association for Research in Vision and Ophthalmology, April 27-May 02, 2019, Vancouver, Canada. Abstract No. A0330.

6. Tham YC, Liu Y. Ting D, et al. Estimation of Haemoglobin A1c from Retinal photographs via Deep Learning. Poster presentation at the 2019 Annual Meeting of the Association for Research in Vision and Ophthalmology, April 27-May 02, 2019, Vancouver, Canada. Abstract No. A0140.

7. Wong TY, Xu D, Ting D, et al. Artificial intelligence deep learning system for predicting chronic kidney disease from retinal images. Poster presentation at the 2019 Annual Meeting of the Association for Research in Vision and Ophthalmology, April 27-May 02, 2019, Vancouver, Canada. Abstract No. A0152.

8. Abràmoff MD, Lavin PT, Birch M, et al. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. *npj Digital Medicine* 2018;**1**:39.

9. De Fauw J, Ledsam JR, Romera-Paredes B, et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nat Med* 2018;**24(9)**:1342-50.

10. Grechenig C, Asgari F, Gerendas BS, et al. Fully-automated drusen segmentation in OCT using deep learning with Pyramid U-net. Poster presentation at the 2019 Annual Meeting of the Association for Research in Vision and Ophthalmology, April 27-May 02, 2019, Vancouver, Canada. Abstract No. A0212.

11. MacCormick IJC, Williams BM, Zheng Y, et al. Accurate, fast, data efficient and interpretable glaucoma diagnosis with automated spatial analysis of the whole cup to disc profile. *PLoS One* 2019;**14(1)**:e0209409.

12. Kermany DS, Goldbaum M, Cai W, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell* 2018;**172(5)**:1122-31.

13. Royal College of Physicians. Artificial intelligence (AI) in health. Position Statement. 3 September 2018. Available at: https://www.rcplondon.ac.uk/projects/outputs/artificial-intelligence-ai-health

14. Department of Health and Social Care. Guidance: Code of conduct for data-driven health and care technology. Updated 19 February 2019. Available at:https://www.gov.uk/government/publications/code-of-conduct-for-data-driven-health-and-care-technology/initial-code-of-conduct-for-data-driven-health-and-care-technology

AUTHOR

Rod McNeil,

Independent Journalist and Consultant.

Watch our interview with Pearse Keane at ESCRS 2018

https://www.eyenews.uk.com/features/video-interviews/post/in-conversation-with-pearse-keane-2018