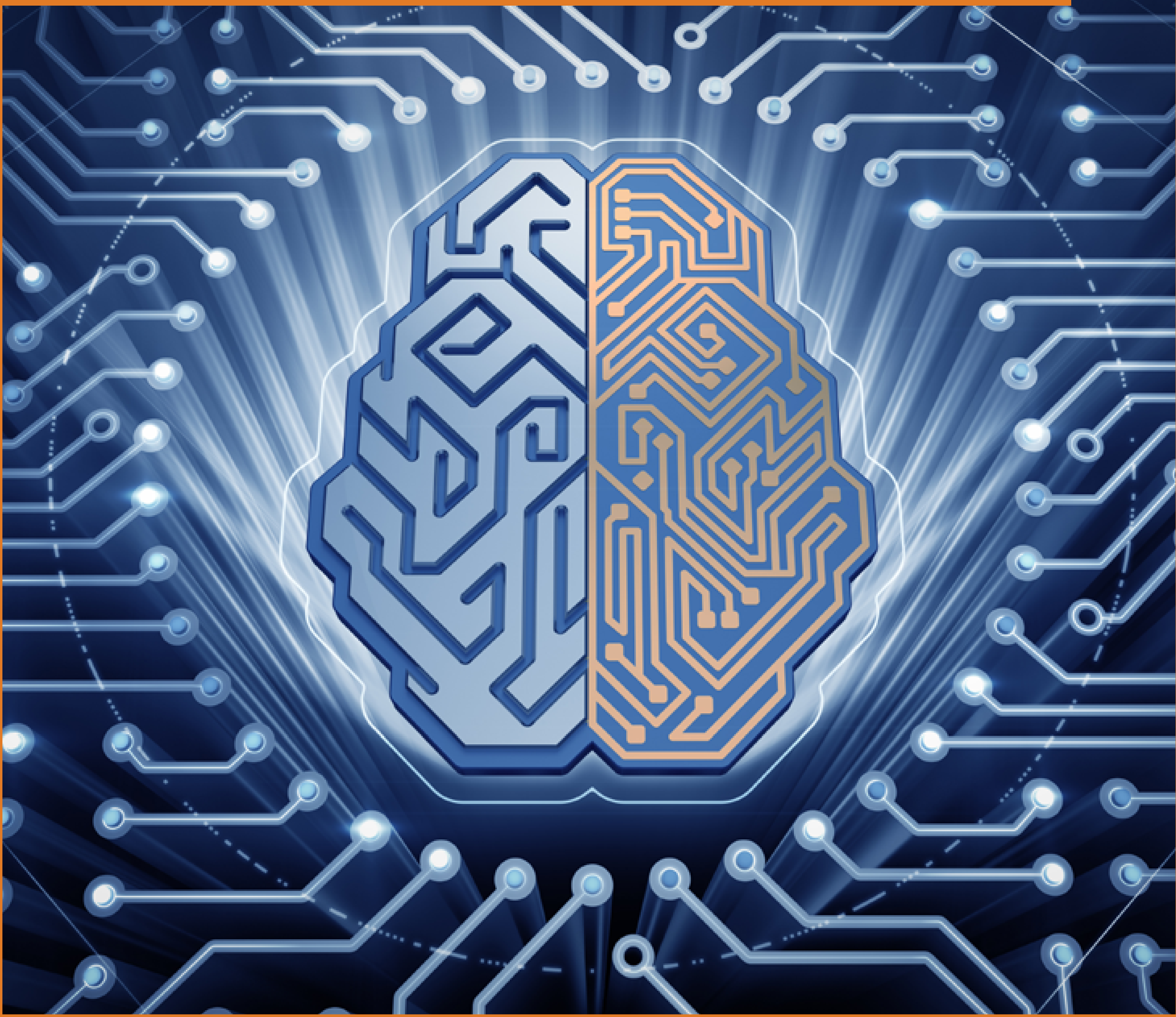
November/December 2021

INDUSTRIALNEWS

The premier source of industry news

EVOLVING AI APPLICATIONS



HOW AI IS ACCELERATING RESEARCH AND DEVELOPMENT

Features:

- The next phase in the evolution of insight
- Missing data links in the IoT revolution
- Factors for successful smart manufacturing

THE AI-BACKED PHYSICS REVOLUTION

95% of our Universe is a mystery to us. Particle physics experiments, like those done at the Large Hadron Collider, explore particle collisions using photon detectors to better understand the basic building blocks of nature. For reliable results, huge volumes of data are needed, however, this brings unique challenges; noise accumulation and scalability to list a few. Extremely scalable and reliable algorithms are becoming increasing vital. AI presents a promising solution to this, especially if it can be demonstrated that greater results can be achieved from machine learning models without the demands for initial large training datasets, as this will minimise the computational expense.

Advancement in Detector Physics relies on enhancing detector technology to improve the spatial resolution, temporal resolution, and data analysis software routines.

This unique collaboration with Photek and the University of Leicester, funded by the Royal Commission for the Exhibition of 1851, aims to develop a new-generation single photon imaging detector. This novel detector with a high image resolution and rapid image refresh rate will have extensive impacts on the world of microscopy. Such a device will allow visualisation of organic and inorganic samples through quantum imaging at a far higher quality than previously possible.

Accomplishment of this innovative solution will create an entirely new class of device with picosecond timing resolution, megapixel spatial resolution and pairing this with robust, AI-enabled, powerful data analytical software, will enable new scientific techniques for investigation of many microscopic processes.

The microchannel plate-based detector has 256 channels in a 16 ×16-pixel layout. Each pixel has a maximum photon rate of 480,000 Hz due to the electronics. As there are 256 channels the timing electronics of the detector are capable of measuring 120 Mega counts per second (Mcps). Simply put, these detectors produce huge volumes of data, in the region of 10 gigabytes per second per detector. With applications such as particle physics requiring hundreds of detectors this huge volume of data paired with current algorithmic software routines is creating a bottleneck in the path to real time data processing software.

Thus, bringing us to the crux of this project; developing software analysis which can detect and visualise the actions of a single photon every microsecond or ideally every nanosecond to bring us closer to real time data processing. How are we hoping to achieve this? AI.

An objective of this research is to explore how greater learning can be achieved in areas where the input data is scarce or difficult to attain. Machine learning (ML) has become a ubiquitous solution for data handling within a vast range of scientific industries, and it is widely accepted that the volume of training data is one of the determining factors in the success of the model. This research queries the data's parameters as well as the volume of training samples. Experimental evidence shows that ML models with small numbers of training samples, but where each sample has multiple features, have been successful at drawing accurate and reliable conclusions. The hypothesis is that this controversy can be explained because there is often structure to the data. By controlling how the data is represented in our models by optimising the architectures we can learn from much smaller training sets with strong guarantees.

This has immense capabilities to

by Amelia Markfort, Industrial Fellow 2021,
Royal Commission for the Exhibition of 1851

revolutionise scientific methods, particularly quantum computing, biological imaging, and particle physics. A prime example is the vast potential for exploitation of ML in space science instrumentation. Ever more powerful space-based detectors, producing enormous datasets at planetary distances stretch satellites' telemetry capabilities. This makes the prospect of on-board data reduction and high-level analysis via powerful ML techniques very attractive, if not essential for learning more about our universe.

Another use of ML is particle identification in particle physics. In this framework the high volumes of data complicate the data analysis to an unmanageable scale. Research aims to explore how a system can be built so that regardless of the parameters of the data, you can take advantage of high volumes of data and control its representation to calculate the limits of its application, consequently allowing for robust and reliable data analysis frameworks to be built.

Other processing tasks such as calibrations and other imaging analytics are essential in other detector applications, understanding how to model generalisable ML systems for these tasks, is necessary for efficient and reliable experimental data.

Beyond the context of detector physics, this research has broader impacts, currently the use of AI is becoming ubiquitous but the approach to it in many scenarios is like alchemy. Which ML model should be used? Convolutional, LSTM, recurrent? How many layers will make the model optimal? How many neurons in each layer? The answers are often unknown.

This collaboration is well placed as it creates a distinct problem in an industry application with known bounds and limited scope, allowing us to explore neural network models and their hyperparameters; the number of layers, the epochs (number of iterations of training) and learning rate to list a few. These can be used to optimise the model and bring greater understanding to the conservative knowledge we have about AI networks. Creating a framework to build robust and reliable ML models by manipulating the parameters of the data will remove the computation expense and requirement of high volumes of training data and is applicable and advantageous to every sector within worldwide industry.