Machine learning for transient science abstracts

Monday 11th December 2023

Tom Killestein (University of Turku)
Sifting supernovae with deep-learned source classification
Invited

Large scale time-domain surveys (GOTO, ZTF, ATLAS) are generating vast quantities of transient discoveries, underpinned by deep learned source classifiers to identify astrophysical transients, reject false-positives, and provide source dispositions to aid in human interpretation. This talk will discuss developments from the GOTO survey in source classification, provide a wider context for the use of machine learning in source classification, and look ahead to the upcoming challenges of upcoming facilities (Vera Rubin Observatory, Argus Array), where a factor 100 increase in alert volume promises to force the development of new and novel architectures to meet the demands of this next generation of transient survey.

Joshua Weston (Queen's University Belfast)
Improvements to the ATLAS Real-Bogus Classifier

We present a Convolutional Neural Network (CNN) for use in the Real-Bogus classification of transient detections made by the Asteroid Terrestrial Impact Last Alert System (ATLAS) and subsequent efforts to improve performance since initial development. As the latest transient surveys increase their scope and depth a greater number of bogus detections will be generated, necessitating improved classification techniques. In such surveys the number of alerts made greatly outstrips the capacity for human scanning, necessitating the use of machine learning aids to reduce the number of false positives presented to annotators. ATLAS has three functioning telescopes each with a different detector and optical performance and here we investigate if training a CNN on data from the individual units improves results compared to a combined training data set. We take a sample of recently annotated data from each of the three operating ATLAS telescopes with ~340,000 real with ~340,000 real (mostly known asteroids) and ~1,030,000 bogus detections per model. Retraining our current CNN architecture with this data we see an improvement to the median False Positive Rate (FPR) from 2.97% to 1.11% with a fixed missed detection rate of 1.00%. Further investigations on the effect of reducing input image size show a negative impact to the false positive rate. Finally architecture adjustments and comparisons to contemporary CNNs indicate our retrained classifier provides an optimal FPR. We conclude that the periodic retraining and readjustment of classification models on survey data can yield significant performance improvements as changes to the data collection process lead to new features in the model input over time.

Anwesha Sahu (University of Warwick)

Machine Learning Applications for Images in Astronomy

In this talk, I will share some applications of convolutional neural networks and computer vision for classification, object identification and big data problems in radio astronomy. Such methods can be applied to automate object identification and classification as upcoming large scale astronomy projects currently face the key challenge of processing copious amounts of data. This project used images from the MeerKAT Galactic Plane Survey and VLA FIRST Survey for training and generating synthetic data, with the goal of identifying and classifying radio objects of interest in the MeerKAT Galactic Plane Survey.

Alex Andersson (University of Oxford)

Discovering radio transients with machine learning and citizen science

Current and upcoming interferometers can now sample wide swathes of the radio sky with unprecedented sensitivity and cadence. As a result, we can now discover novel radio transients across an immense range of astrophysical regimes - from flare stars to FRBs. I will discuss recent, serendipitous discoveries being made with the MeerKAT radio telescope and how we can make the most of new facilities coming online. This includes how citizen scientists have scoured through commensal data and uncovered 100s of new variable and transient sources. This is the first crowdsourcing project dedicated to radio transients in this manner and has uncovered variable sources as different as nearby flare stars, pulsars and AGN. I will detail unsupervised machine learning techniques being applied to speed up the search for interesting and anomalous sources. These anomaly detection models can, with the use of active learning strategies, be customised to find not only anomalies, but those that are verified as interesting systems for a particular science case. The tested models show great success in recovering transients in our large dataset and these techniques hold great promise for searching large datasets for rare and unique systems..

Alexander Gagliano (Harvard/MIT)

Multi-Modal Insights for Early Photometric Classification of Supernovae in the Era of LSST

In this talk, I will present a "First Impressions" network designed and validated for early photometric classification of supernovae. I will discuss the design decisions that have motivated the recurrent neural network, including transfer learning to ensure performance when shifting from synthetic to observed samples; a custom weighted loss to prioritize accurate early predictions over late ones; and the inclusion of host galaxy photometry to inform the model's final classification. At both early and late phases, this method achieves comparable or superior results to the leading classification algorithms with a simpler network architecture. I will end by discussing the value of spatial information and Rubin's intermediate data products for improved early classification. In the impending era of petabyte-scale sky surveys, early photometric classification will be our first line of defense for translating photons to scientific insight.

Justyn Maund (University of Sheffield)
Light curve and spectrum classification of transients

TBD

Tuesday 12th December 2023

Matt Grayling (University of Cambridge)
Augmenting supernova training sets using Generative Adversarial Networks
Invited

The vast data sets that will be provided by upcoming surveys such as LSST means that spectroscopic follow-up will only be possible for a small fraction of observed transients, making reliable photometric classification of great importance moving forward. However, one of the main issues faced by such classifiers is a lack of training data, particularly of rarer transient classes. One approach to mitigate for this is to generate synthetic data to train these models, often in the form of simulations. However, as gaps remain in our understanding of the physical processes involved in many types of transient simulated samples are often not representative of the observed population of objects.

I will present my work focusing on using data-driven methods for data augmentation, specifically the use of Generative Adversarial Networks (GANs). These networks have been successfully applied in other fields, such as medicine, in order to generate realistic training data to augment sparse data sets and boost classifier performance. GANs have the potential to generate realistic multi-band light curves which span the full variability of the observed population without making any assumptions about the underlying physics. By comparing GAN-generated supernova light curves with the sample on which they were trained, I will demonstrate that this technique can indeed be used to generate physically realistic training data and has the potential to improve the performance of photometric classification techniques.

Erin Hayes (University of Cambridge)

GausSN: Bayesian Time Delay Estimation for Strongly Lensed Supernovae

Time delay cosmography with strongly lensed supernovae (SNe) is an exciting local probe of H0 that is independent of the local distance ladder. One of the most important ingredients in H0 estimates from strongly lensed SNe is the time delay between the appearance of the multiple images of the SN. In this talk, I describe GausSN – a new method for extracting time delays from multi-band photometric observations of resolved lensed SNe images using Gaussian Processes. Our methodology improves upon existing time delay estimation methods by including a fully Bayesian exploration of the parameter space and a novel treatment of microlensing, all with minimal assumptions about the underlying shape of the light curve. We demonstrate the ability of GausSN to recover accurate and precise time delays using simulations of lensed SNe data as expected from the Roman Space Telescope. Compared to existing methods, GausSN recovers time delays which are on average as close or closer to the truth with better calibrated uncertainties. With the upcoming Rubin Observatory's Legacy Survey of Space and Time and Roman Space Telescope, we expect to discover tens to hundreds of lensed SNe, which, with tools like GausSN, will provide important independent evidence for investigating the Hubble tension in the coming decade.

Shan Raza (University of Warwick)

Building Computational Tools for Deployment of AI algorithms in Pathology

Development of AI algorithms for giga pixel pathology images is a challenging tasks due to the sheer size of the images and requirement for processing the images at multiple resolutions. At TIA centre, we have build an open-source computational pathology toolbox for researchers which is freely available to download at this link https://github.com/TissueImageAnalytics/tiatoolbox. The toolbox has been downloaded more than 150K times during the past three years and has been used to benchmark various algorithms in computational pathology.

Heloise Stevance (University of Oxford)

Just because you can doesn't mean you should

Invited

The number of published articles mentioning machine learning tools has grown dramatically in the last few years, but many very promising proof-of-concepts have failed to deliver when applied in situ. This is not a problem that is limited to astronomy and the transient field - all over science and data science, researchers rush to apply a new method without spending enough time defining the scope of the question that needs to be answered, what bench marks must be met for a real use-case to be achieved, and what validation metrics are the most appropriate. In this talk I want to address what fundamental solution design methodologies need not be forgotten if we want to ensure our published models are not more artificial than they are intelligent.