An assertion

• Astronomy, like biology, geology and exo-planetology is a phenomenological subject
  • In contrast, physics aims to distil phenomena to laws and ultimately a few numbers

• Phenomenological subjects have three distinct steps
  • Discovery
    • The Universe has more imagination than any one of us (Zwicky)
  • Systematic study of phenomena: Search for patterns
  • Understanding via a "physical model"
    • Physics is a subject that aims to build models to explain patterns by the use of laws and rules
Mathematics & Technology

- Mathematics is an integral part of natural sciences because it allows signal levels to be calculated
  - The origins of mathematics lie in poetry and practical needs
- Discovery is the blood of science. It is largely a function of technology
  - The last several decades has seen spectacular progress in astronomy
    - My view is that it is entirely due to technological advances (and not due to insight nor brilliance of astronomers)
  - Biology had a later start but has an acceleration larger than that of astronomy
Mathematics & Technology: Example

• Example: Planetary motions
  • Planetary Motion noticed by ancient scientists
  • With the first telescope Galileo discovered moons of Jupiter
  • Newton noted the falling of an apple
  • Brahe made careful observations of planetary positions
  • Kepler distilled the data into three laws of Kepler
  • Newton develop the theory of Gravity (inverse square law force)
  • Newton & Leibniz developed calculus
  • Birth of modern Mechanics
  • Gauss computed trajectories and made predictions.
This is how it all began
Chemistry in the young Universe was simple
Time Domain Astronomy (Supernovae)
Time Domain Astronomy: Fritz Zwicky @ the Palomar 18-inch telescope
A star dies: A supernova is born & new elements made
We are star dust!

From Universetoday.com
A factory to discover supernovae
Palomar Transient Factory
Google Mastermind Turns to the Stars

Wayne Rosing on Revolutionizing Astronomy from Google
Palomar Transient Factory: Hardware, Software & Grayware!
Software, software & more software

Time from exposure to candidates in the subtraction database: 20 - 40 minutes

Supernova, good!

Subtraction artifacts, bad!
Spectroscopy on demand: Robotic Spectral Energy Distribution Machine
Rainbow camera: acquisition camera
Center aperture: hyper-spectrograph

\[ g \sim 40\% \]
\[ u \sim 8\% \]
\[ i \sim 30\% \]
\[ r \sim 45\% \]
Comparison with other Spectrographs

Type Ia Supernovae
- ZTF18achaqm | $z = 0.017$ | $\phi = -14$ d
  - LT $\Delta \phi = -1$ d
  - NTT $\Delta \phi = +1$ d

- ZTF18acdwohd | $z = 0.036$ | $\phi = -3$ d
  - NTT $\Delta \phi = -1$ d

- ZTF18acbxsgz | $z = 0.025$ | $\phi = -11$ d
  - DUP $\Delta \phi = +4$ d

- ZTF18acbwonm | $z = 0.060$ | $\phi = +4$ d
  - NTT $\Delta \phi = +2$ d

Type II Supernovae
- ZTF18abojmr | $z = 0.038$ | $\phi = +0$ d
  - P200 $\Delta \phi = +3$ d

- ZTF18abwkrbr | $z = 0.010$ | $\phi = +30$ d
  - NOT $\Delta \phi = +4$ d

- ZTF18aczpxr | $z = 0.036$ | $\phi = -10$ d
  - Keck $\Delta \phi = -2$ d

- ZTF18aciqqaw | $z = 0.009$ | $\phi > +15$ d
  - P200 $\Delta \phi = -3$ d

Type Ib/c Supernovae
- ZTF18abivorn | $z = 0.020$ | $\phi = +5$ d
  - P200 $\Delta \phi = +2$ d

- ZTF18abktmbr | $z = 0.021$ | $\phi = -13$ d
  - P200 $\Delta \phi = +1$ d

- ZTF18abwknv | $z = 0.003$ | $\phi = +14$ d
  - LT $\Delta \phi = -1$ d
  - P200 $\Delta \phi = +5$ d

- ZTF18acsirbc | $z = 0.037$ | $\phi = +0$ d
  - Keck $\Delta \phi = -8$ d

Normalized Flux vs. Wavelength [Å]
Methodology & Technological Developments

- **Phase I (Palomar Transient Factory):** 2009-2012
  - Machine Learning (ML) for classification
  - Same night classification & follow up

- **Phase II (intermediate Transient Factory):** 2013-2016
  - Mixed cadence observing
  - Multi-band observing
  - Robotic IFU spectroscopy (SEDM on P60)
  - Demonstration of a Needle in a haystack Search (robust & rapid ML)

- **Phase III (Zwicky Transient Factory):** 2018-2020
  - Pre-cursor to LSST for Time Domain Astronomy
  - Industrialization of alert distribution
Innovations

• Real-time data transmission (microwave link)
• Robust (fault-tolerant) & real-time pipelines
• Telescope sequencing not done daily but weekly via a sophisticated algorithm
• Application of Machine Learning (ML) for quantitative assessment of candidates
• Devised tools for “follow up”
  • “Marshal” (aggregator of information)
  • “Broker” (help take decision on potential value of candidates using contextual data, past data and other catalogs)
Astronomy is, despite astronomers, useful
Towards Automated Discovery of the Universe
S. Kulkarni
Principal Investigator

M. Graham
Project Scientist

E. Bellm
Survey Scientist
ZTF already motivated a second company

http://www.soroco.com/
Field of view & aperture comparison (ZTF: wide & shallow)
Zwicky Transient Facility: First Light
ZTF: The big picture

- SEDM: Future Plans
- Instrument IO Panel
- New Optical Bench
- Removable Enclosure
- PSU Management
- 2X NPS (Remote Shutoff)
- QDC Manifold Glycol Plumbing
- Dry Air Supply
- IO Bulkhead Connectors

...
Alert structure: AVRO format

63 x 63 pixel 32-bit images

Unique spatially matched alert name

Rolling 30-day window light curve

• ZOGY parameters
• Real-bogus score
• Star/galaxy score
• 3 nearest PS1 sources
• Nearest SS object
• Alert history

https://github.com/ZwickyTransientFacility/ztf-avro-alert
## Where can I get alerts?

<table>
<thead>
<tr>
<th>Service</th>
<th>Basic web search</th>
<th>User-defined filters</th>
<th>Notifications</th>
<th>Kafka streams</th>
<th>API</th>
<th>Bulk access</th>
</tr>
</thead>
<tbody>
<tr>
<td>LCO MARS</td>
<td>Yes</td>
<td>No</td>
<td>-</td>
<td>No</td>
<td>JSON</td>
<td>No</td>
</tr>
<tr>
<td>ANTARES</td>
<td>No</td>
<td>Python</td>
<td>Slack</td>
<td>Yes</td>
<td>Python</td>
<td>(Yes)</td>
</tr>
<tr>
<td>LASAIR</td>
<td>Yes</td>
<td>SQL</td>
<td>-</td>
<td>No</td>
<td></td>
<td>No</td>
</tr>
<tr>
<td>UW</td>
<td>No</td>
<td>No</td>
<td>-</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>ALERCE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ALERCE is coming soon.
Projects in the works

• Replacing image subtraction algorithms (optimized for Gaussian/Poisson noise) with optimized for better ROC
• Detection of streaks (fast moving objects: asteroids & spacecraft debris) against a variety of artifacts
• Adaption of ML for changes in pipeline processing
• Replace spectral identification with reliable ML/DL scoring
• Making “Needle in Haystack” searches routine (rapid decision making with inputs from vast data bases)
• Higher level of engagement with Zoonverse (Citizen Science/Oxford)
# ZTF = 0.1 LSST

<table>
<thead>
<tr>
<th></th>
<th>ZTF</th>
<th>LSST</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of sources</td>
<td>1 billion</td>
<td>37 billion</td>
</tr>
<tr>
<td>No. of detections</td>
<td>1 trillion</td>
<td>37 trillion</td>
</tr>
<tr>
<td>Annual visits per source</td>
<td>1000 (2+1 filters)</td>
<td>100 (6 filters)</td>
</tr>
<tr>
<td>No. of pixels</td>
<td>600 million (1320 cm² CCDs)</td>
<td>3.2 billion (3200 cm² CCDs)</td>
</tr>
<tr>
<td>Field of view</td>
<td>47 deg²</td>
<td>9 deg²</td>
</tr>
<tr>
<td>Hourly survey rate</td>
<td>3750 deg²</td>
<td>1000 deg²</td>
</tr>
<tr>
<td>Nightly alert rate</td>
<td>1 million</td>
<td>10 million</td>
</tr>
<tr>
<td>Nightly data rate</td>
<td>1.4 TB</td>
<td>15 TB</td>
</tr>
</tbody>
</table>
Next Show in 2023: Ultimate Celestial Cinematography
Detection of Streaks

Dmitry Duev & Ashish Mahabal
Zwicky Transient Facility
Caltech
Fast moving objects appear as streaks

Images are taken and subtracted with a reference image to eliminate static objects and reveal the streaks
1. Near-Earth Asteroids
2. Satellite debris
3. Artifacts
Artifacts

We chose not to use transfer learning to initialize or freeze layer weights for the deep models and trained all our models from scratch. The reason is that the available pre-trained networks are trained on drastically different image data sets and thus do not necessarily capture the features relevant to this work.

While providing a similar sensitivity, DeepStreaks demonstrates a 50× better performance than the original MNRAS 000, 1–8 (2019)
DeepStreaks: results

• 96-98% true positive rate, depending on the night
  • Quantified by performance on test data sets and using known NEOs observed by ZTF

• Below 1% false positive rate, 50x-100x improvement over original RF classifier

• Near-real-time operations; below 10 min per day spent by human scanners vs ~hours with original RF classifier

• 16 confirmed new NEAs over 57 observable nights

• Another 20 “lost” due to insufficient follow-up
DL wish list for ZTF

- Convolutional Autoencoder: detect transients (based on Ashish/Nima’s TransiNet/ZwickyNet), rb
- Photometry time-series + image data (RNN/Attention + CNNs): variable science. For ZTF v2 + spectroscopy
- DeepStreaks v2: Convolutional Autoencoder-based detection + identification + PSF fit on sci+ref stacks
- DL-based deconvolution in densely populated regions

- Edge TPUs for inference: cheap and efficient power usage [~ $100 per device]
- More hands for (expert) data labeling!

<table>
<thead>
<tr>
<th>Model architecture</th>
<th>Desktop CPU*</th>
<th>Desktop CPU * + USB Accelerator (USB 3.0) with Edge TPU</th>
<th>Embedded CPU **</th>
<th>Dev Board † with Edge TPU</th>
</tr>
</thead>
<tbody>
<tr>
<td>MobileNet v1</td>
<td>47 ms</td>
<td>2.2 ms</td>
<td>179 ms</td>
<td>2.2 ms</td>
</tr>
<tr>
<td>MobileNet v2</td>
<td>45 ms</td>
<td>2.3 ms</td>
<td>150 ms</td>
<td>2.5 ms</td>
</tr>
<tr>
<td>Inception v1</td>
<td>92 ms</td>
<td>3.6 ms</td>
<td>406 ms</td>
<td>3.9 ms</td>
</tr>
<tr>
<td>Inception v4</td>
<td>792 ms</td>
<td>100 ms</td>
<td>3,463 ms</td>
<td>100 ms</td>
</tr>
</tbody>
</table>

* Desktop CPU: 64-bit Intel(R) Xeon(R) E5-1650 v4 @ 3.60GHz
** Embedded CPU: Quad-core Cortex-A53 @ 1.50GHz
† Dev Board: Quad-core Cortex-A53 @ 1.50GHz + Edge TPU
DeepStreaks: architecture

Convolutional-neural-network, deep-learning system designed to efficiently identify streaking FMOs

• "rb": bogus or real streak? Identify all streak-like objects, including actual streaks from FMOs, long streaks from satellites, and cosmic rays

• "sl": long or short streak?

• "kd": ditch or keep? Is this a real streak, or a cosmic ray/some other artifacts?
DeepStreaks: performance

Out of 210 streaks from real NEOs detected by the ZTF Streak pipeline at IPAC, 202 (96%) are correctly classified.

ROC curves of individual classifiers deployed as of February 2019

Completteness identifications using known NEOs observed by ZTF in October 2018 – January 2019.
Image subtraction for hunting transients without subtraction

Encoder-Decoder network encapsulates the essence of transientness

Sedaghat and Mahabal, 2017
"rb": streak or not?
"sl": long or short?
"kd": real or cosmic?

96% completeness
Human vetting down from several hours to few minutes

16 confirmed new NEAs over 57 observable nights
Another 20 “lost” due to insufficient follow-up

Featured in Minor Planet Center circulars as D. Streaks

Duev et al. 2019
Figure 11. Examples of $(dm, dt)$ Probability Distribution Functions. Smoothed 2D histograms are shown for SN Ia (top-left), SN IIp (top-right) and RR Lyrae (bottom-left), using bins of width $\delta t = 1$ day (x-axis), and $\delta m = 0.01$ (y-axis). The superimposed diamonds are from a single LC (of SN Ia). PDFs for the two SN types form a better fit than that of RR Lyrae (and SN Ia is a better fit than SN II P). Various metrics on probability distributions can be used to automatically quantify the degree of fitness. The decision tree used is shown at bottom-right.
(dmdt) Image representation to make the images CNN-ready

Mahabal et al., 2017

light curve with n points

23 x 24 output grid

n * (n-1)/2 points

Area equalized pixels
with Kshiteej Sheth
Classification Workflow

Light curves → Feature vectors → Dimensionality Reduction → Classification

Domain knowledge/subjectivity

Light curves → Density representation → Equi-area images → Convolutional Neural Network
Next steps: Domain adaptation by including archives

Using CRTS, PTF, Gaia, Pan-STARRS etc. data

CRTS transients
S Das, A Mahabal

Using light curves (with mag err)

To be applied to ZTF
Started:

- Transient hunting from images using deep learning

Planned:

- Deconvolution in crowded fields

Future:

- Real-real classifications by combining diverse data
  - light curves
  - meta data
- Period finding for all periodic sources
- LSTMs/RNNs based on light curves
- Combining diverse types of labels (e.g. marshal, Zooniverse etc.)
- Domain adaptation using archival data