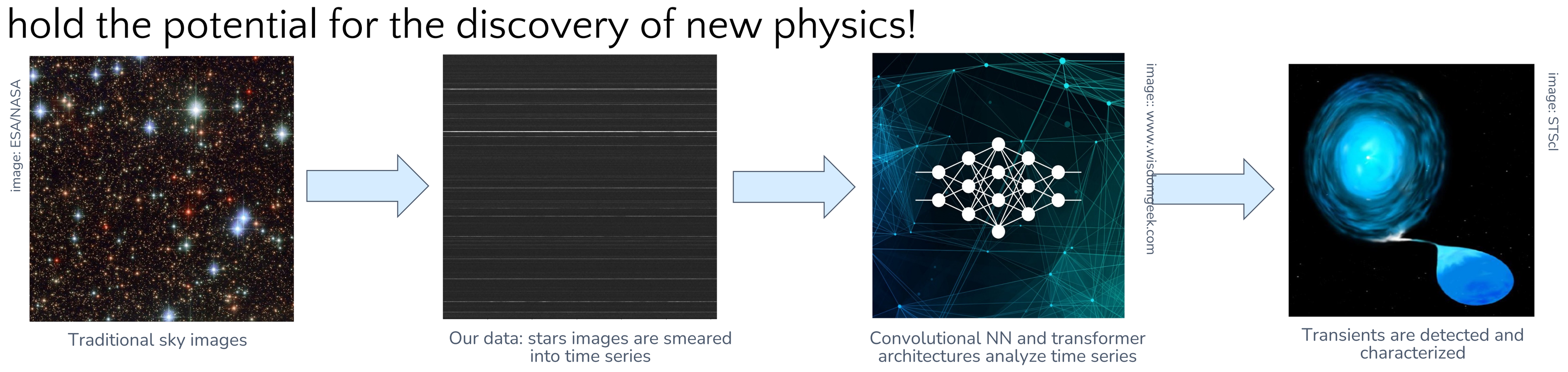
TLDR (Summary)

I am developing novel data collection and AI-aided analysis methodologies that will allow us to systematically and thoroughly investigate the dynamic universe at sub-second timescales by integrating stars along the image over time. The sub-second timescales of astrophysical variability are largely unexplored, but can reveal the nature of known phenomena by characterizing their short-term variability and even hold the potential for the discovery of new physics!



Motivation: Why care about sub-second transients?

Transient astronomy, the study of phenomena in the universe that change on human timescales, is an opportunity to discover everything from fundamental physics to the most energetic phenomena ever observed. Synoptic surveys provide opportunities find transients never before seen nor even hypothesized.

Transients are characterized by the time scale over which they vary, their "characteristic time scale", and by their energy output. **Figure 1** shows the zoo of known optical transients in a phase-space of peak magnitude - characteristic time scale. The hours-to-years portion of the graph is well explored, with many known phenomena. However, the minutes-to-fractions-of-seconds region is nearly empty. Several phenomena that are expected to vary on sub-second timescales (e.g. [2, 3] see **Fig. 2**) but observations in this space are particularly challenging .

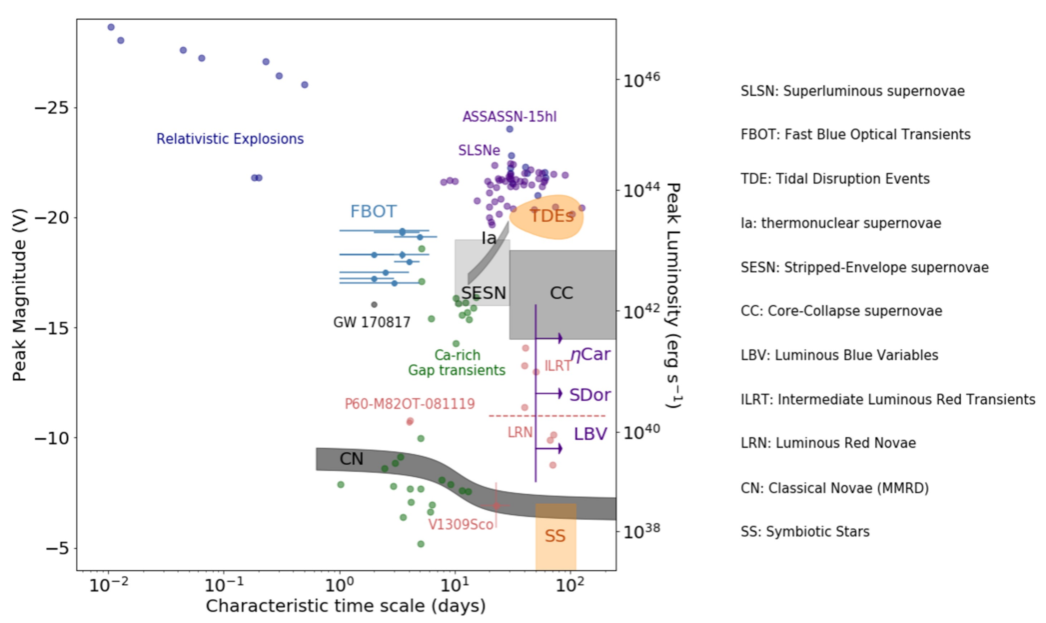


Figure 1 (modified from [1]): Known transient phenomena, categorized by characteristic timescale and peak magnitude

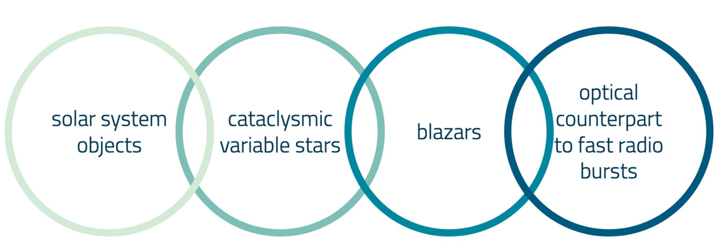


Figure 2: Expected sub-second optical transients
Solar system objects: occultations of Kuiper belt objects
Cataclysmic variable stars: episodic violent stellar bursts
Blazars: active galactic nuclei with relativistic jets
Fast radio bursts: highly energetic mysterious radio sources

Traditional observational methods cannot detect transients at this timescale.

Data: Continuous-readout images

In traditional astronomical observations, the camera shutter opens, takes in photons, closes, and then the positions and intensities of the captured light is read out. These exposure and readout cycles require seconds or minutes to complete.

However, two nontraditional observing modalities, **trailing and continuous-readout [4],** enable resolution at sub-second timescales by leaving the shutter open during slew or readout. This has the effect of integrating each astrophysical image along one spatial dimension, so that the resulting image has one time dimension and one spatial dimension (see **Fig. 3).** This methods allow us to resolve the data at second or sub-second sampling rates.

Our images were taken in continuous-readout mode at the Zwicky Transient Facility [ZTF, 5] in a special program lead by PIs Andreoni and Mahabal. The images are sampled approximately 300x per second. 450GB of data are currently available, and the dataset will triple in the next few years.

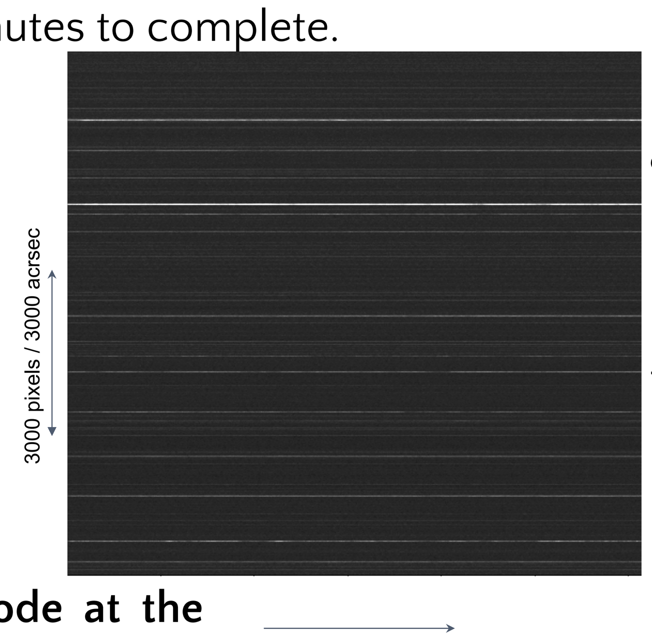
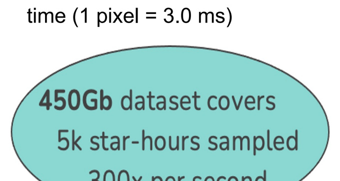


Figure 3: an example of our ZTF data



Trailing images while slewing has been proposed for the upcoming Vera C. Rubin LSST [6], leading to a 10ms sampling rate.

Neural Networks for Sub-Second Astronomical Transient Discovery

Shar Daniels, University of Delaware | FASTLab
Federica Bianco, UD
Igor Andreoni, UMD
Ashish Mahabal, CalTech

Acknowledgements

This work is supported by an Amazon Web Services grant (PI Andreoni) and by the University of Delaware DARWIN computing system: DARWIN – A Resource for Computational and Data-intensive Research at the University of Delaware and in the Delaware Region, which is supported by NSF under Grant Number: 1919839, Rudolf Eigenmann, Benjamin E. Bagozzi, Arthi Jayaraman, William Totten, and Cathy H. Wu, University of Delaware, 2021, URL:https://udspace.udel.edu/handle/19716/29071

References

[1] Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. (2019), The Astrophysical Journal, 873, 111.
[2] McLaughlin, Maura A., A. G. Lyne, D. R. Lorimer, M. Kramer, A. J. Faulkner, R. N. Manchester, J. M. Cordes et al. "Transient radio bursts from rotating neutron stars." Nature 439, no. 7078 (2006): 817-820.
[3] T. C. Nihei et al 2007 AJ 134 1596
[4] Bianco, F. B., Protopapas, P., McLeod, B. A., Alcock, C. R., Holman, M. J., & Lehner, M. J. (2009), The Astronomical Journal, 138, 568.
[5] Bellm, E. C., Kulkarni, S. R., Graham, et al. (2019), Publications of the Astronomical Society of the Pacific, 131, 018002.
[6] Thomas, D., & Kahn, S. M. (2018), The Astrophyscial Journal, 868, 38.
[7] Thomas, D. et al 2019, https://docushare.lsstcorp.org/docushare/dsweb/Get/Document-30601/thomas_startrails_mini.pdf.
[8] LeCun, Yann; Bengio, Yoshua (1995). "Convolutional networks for images, speech, and time series". In Arbib, Michael A. (ed.). The handbook of brain theory and neural networks (Second ed.). The MIT press. pp. 276–278.
[9] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł. and Polosukhin, I., 2017. Attention is all you need. Advances in neural information processing systems, 30.

Methodology: Sliding-window CNN

Continuous-readout images require custom-made analysis tools. Our planned methodology includes the application of convolutional neural networks (CNNs) [8] and transformers [9]. Here we present preliminary results for our CNN-based analysis.

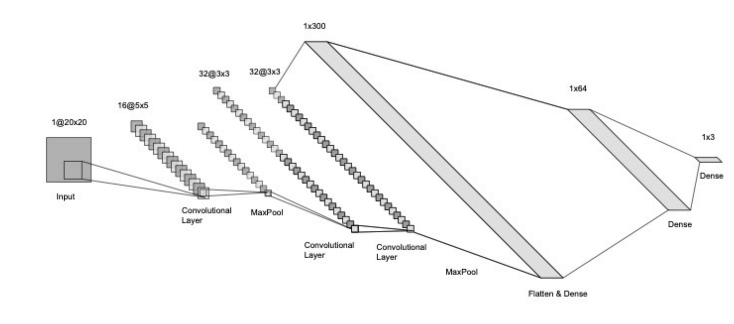


Figure 4: CNN architecture

Sliding-window CNN model:
Our sliding-window CNN scans the images, classifying each pixel in the context of a larger square window of pixels around it and looking for pre-determined patterns of transient behavior (brightening and dimming). **For each pixel, the CNN outputs the probability that it belongs to each of three classes: background, star-streak, or star-streak with transient.**

Training data preparation:
We generated a synthetic transient dataset by implanting brightening events in >18,500 star-streaks. Each implantation is generated by multiplying the star-streak flux by as a gaussian profile increase:

$$F * ((A-1) * (e^{-(x-M)^2/(2\sigma^2)}) + 1)$$

F is the initial streak flux, M is the location of the transient along the streak, A is the brightening amplitude, and σ is the transient's duration.

Parameter	Value(s)
A	10, 8, 6, 4, 3, 2, 1.5, 1.3, 1.1, 1.05, 1.01
M	10
σ	3

Table 1: Parameter values.

See **Fig. 5** for an example and **Table 1** for the distribution of parameters in our implantations. **Future work will include increasing the diversity of transient morphologies.**

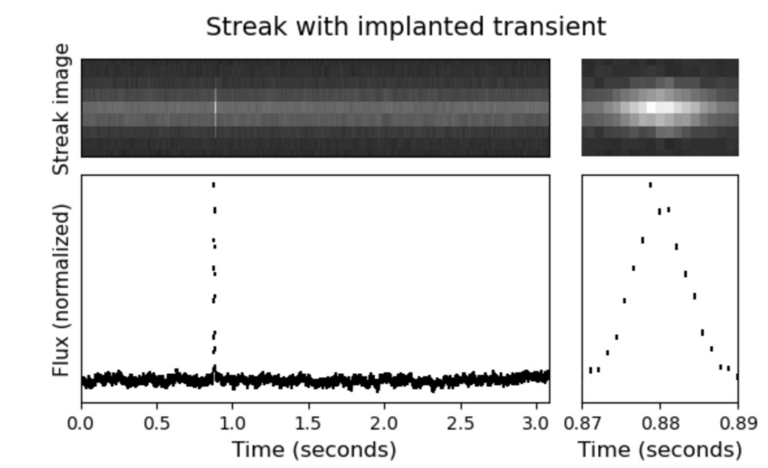


Figure 5: Example of a transient implantation. Top: the implanted streak (L), zoomed in (R) Bottom: resulting time series, aperture photometry

Preliminary results

The efficiencies of the CNN's transient retrieval for each transient brightness are shown in **Fig. 6.** Detected transients are shown in **Fig. 7** (true detection) and **Fig. 8** (false positive).

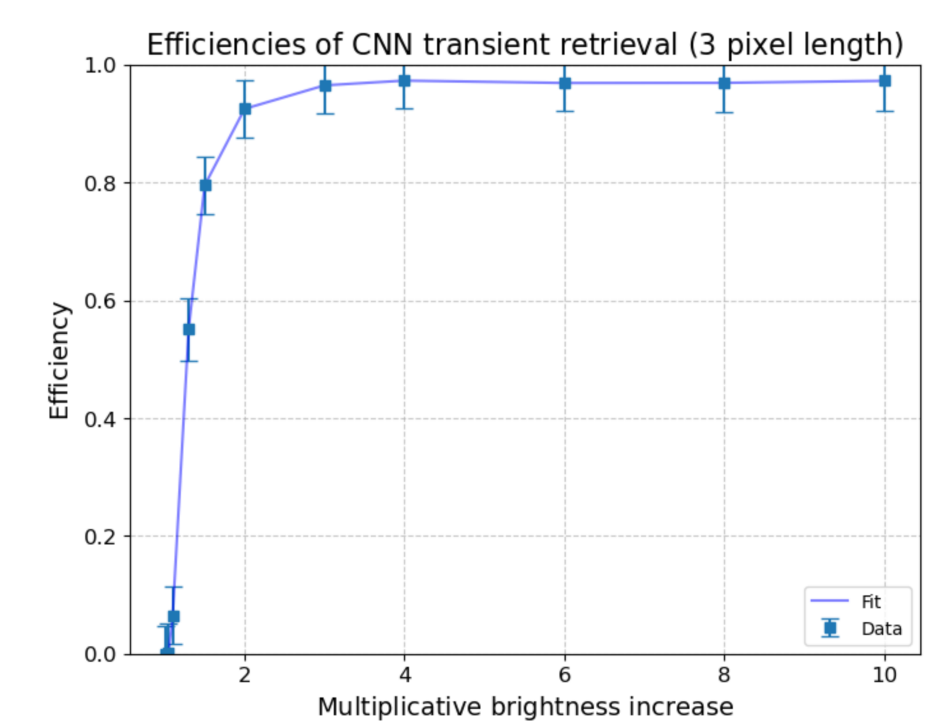


Figure 6: Transient retrieval efficiencies (correctly identified transients/total transients), with Poisson error.

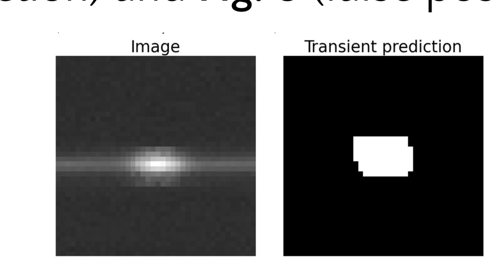


Figure 7: Example of our CNN's successful detection of an implanted transient. Left is the input image with implanted transient; Right is the binary pixel classifications of the CNN as transient (white) and not transient (black).

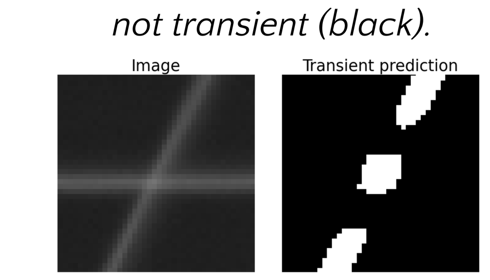


Figure 8: Example of our CNN's detection of a satellite crossing in front of our star-streak, leading to a false positive.

The CNN can detect over 90% of transients with 2x brightness increases and greater.

Future work: Transformers

Transformers [9] are a class of neural networks for language prediction based on the attention mechanism, which enables variable correlations to be established between elements of a sentence. The encoder element of the transformer can enable classification of multivariate time series.

Continuous read-out sky images are especially suited to transformer based analysis as every star therein is a multivariate time series. **Using transformers would allow us to extend our detection of transients to a characterization of their transient behavior.**

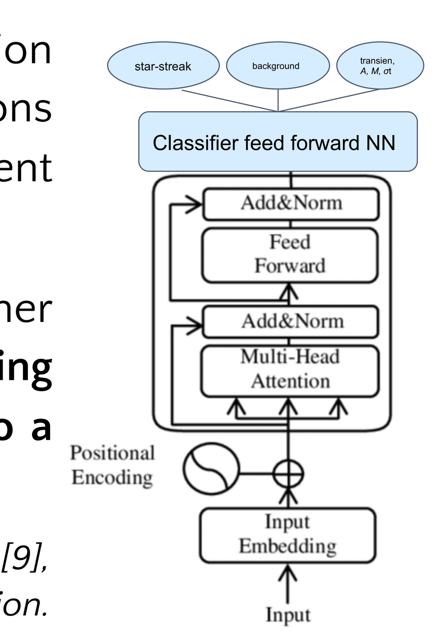


Figure 8: Modified transformer architecture: the encoder is identical to [9], connected to a feedforward NN with a 3-neuron output layer for classification.