

Detecting thousands of space debris by an AI-based video analysis

Motivation: It has been unknown whether there exist transient astronomical phenomena that emit visible light only for a second. Potentially a huge discovery space is open.

Method: We use Tomo-e Gozen, the world's most sensitive wide-field optical CMOS sensor array installed on Kiso Schmidt Telescope, and analyzed an unprecedentedly large volume (nearly 100TB) of the collected data. We developed a custom made deep neural network approach based on Single Shot MultiBox Detector (W. Liu et al. 2015).

Result: We identified approximately 1,500 flash events (glints) that appear for only half a second. We found that the majority of these events are caused by artificial satellites and space debris orbiting the Earth, which produce brief flashes by reflecting sunlight. The occurrence rate of such events is estimated to reach up to **ten million times per day over the entire sky**.

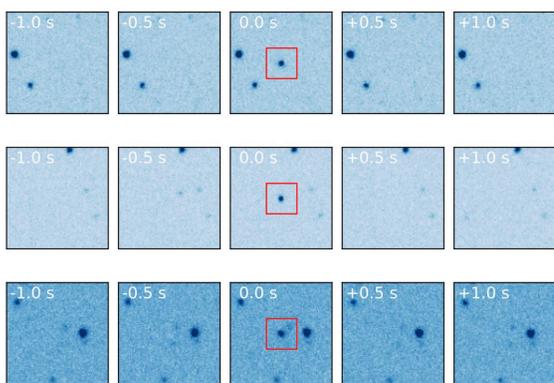


Fig. 1 Selected samples of short flashes identified by our detection software. The red bounding box is automatically set by our detection AI model.

Other samples with movie can be found here https://www.astr.tohoku.ac.jp/~masaomi.tanaka/press_glint/

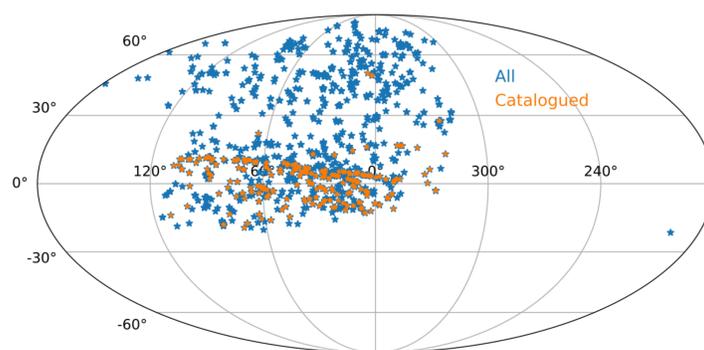


Fig. 2 Distribution of the detected glints in the sky. The blue points show all the glints while the orange points show those associated with catalogued objects.

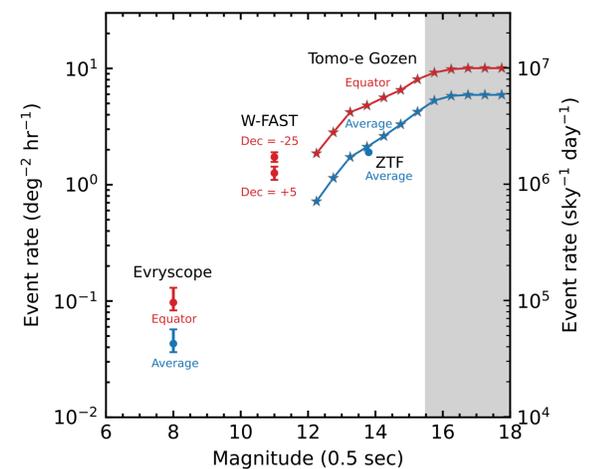


Fig. 3 The derived event rate of second-timescale glints as a function of magnitude (brightness).

Publication: M. Tanaka, I. Takahashi, N. Yoshida et al., "Second-timescale Glints from Satellites and Space Debris Detected with Tomo-e Gozen" *The Astrophysical Journal* 994 (2025) 175

CosmoGLINT: A transformer model that speaks "galaxies"

Motivation: Modelling star-forming galaxies is crucial for analysis of data from upcoming large galaxy surveys and line intensity mapping. LLMs can be used to generate quickly a number of galaxy catalogues, which we can utilize to derive statistical quantities and to develop data analysis pipelines.

Method: We develop CosmoGLINT, a Transformer-based generative framework designed to create realistic galaxy populations from outputs of simulations of gravitational dynamics. CosmoGLINT autoregressively generates sequences of galaxy properties – including star formation rate, distance to the main galaxy, and radial and tangential velocities – conditioned on the mass of local density peaks.

Result: Trained on outputs of a cosmological hydrodynamic simulation, the model reproduces key statistical properties of the original data, including the voxel intensity distribution and the power spectrum both in real and redshift space. The original hydrodynamics simulation takes millions of CPU-hours on supercomputers, whereas our model can generate a galaxy catalogue in a minute.

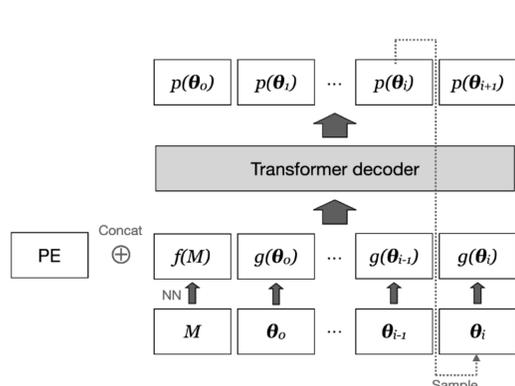


Fig. 4 The architecture uses halo mass M and the sequence of galaxy properties θ as inputs. They are embedded into a latent space using neural networks f and g . The sequence is passed through a Transformer decoder, which outputs a sequence of predicted probability density functions $p(\theta_0), \dots, p(\theta_i)$.

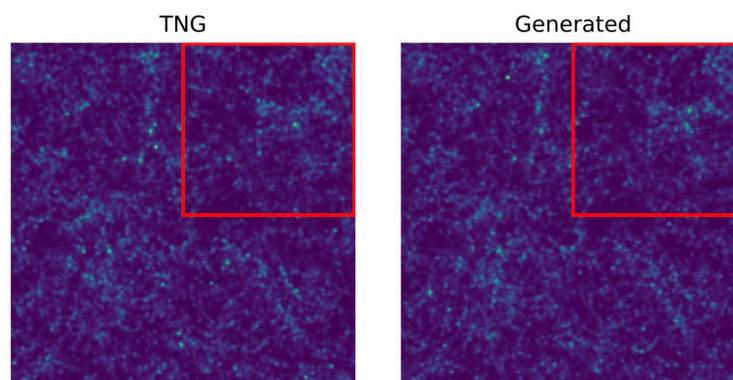


Fig. 5 We compare the spatial distributions of the galaxies in the numerical simulation (left) and of the generated galaxies (right). The overall brightness as well as large-scale distribution are well reproduced by our model.

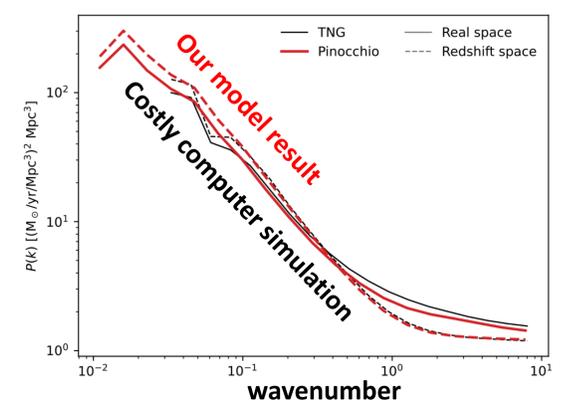


Fig. 6 Power spectra of the spatial distribution of galaxies weighted by star formation rate. The dashed lines are for those measured in velocity (redshift) space. Our model reproduces well the velocity distribution of the galaxies.

Publication: K. Moriwaki, R. Jun, K. Osato, N. Yoshida, "CosmoGLINT: cosmological generative model for line intensity mapping with transformer" *Monthly Notices of the Royal Astronomical Society* 545 (2026) 2124