

木曾シュミットシンポジウム2022  
2022年7月5日

# 機械学習によるTomo-e Gozen動画 データからの短時間突発現象の検出

東北大学  
高橋一郎

共同研究者：田中雅臣(東北大学),上田修功(NTT コミュニケーション科学 基礎研究所),  
富永望(国立天文台), 酒向重行, 大澤亮, 吉田直紀(東京大学)

# Tomo-e Gozen Real/Bogus分類

## 前回の木曾シンポ発表

<https://arxiv.org/abs/2206.12478>

ソースコードとデータセットも公開

arXiv > astro-ph > arXiv:2206.12478

Astrophysics > Instrumentation and Methods for Astrophysics

[Submitted on 24 Jun 2022]

### Deep-learning Real/Bogus classification for the Tomo-e Gozen transient survey

Ichiro Takahashi, Ryo Hamasaki, Naonori Ueda, Masaomi Tanaka, Nozomu Tominaga, Shigeyuki Sako, Ryou Ohsawa, Naoki Yoshida

We present a deep neural network Real/Bogus classifier that improves classification performance in the Tomo-e Gozen transient survey by handling label errors in the training data. In the wide-field, high-frequency transient survey with Tomo-e Gozen, the performance of conventional convolutional neural network classifier is not sufficient as about  $10^6$  bogus detections appear every night. In need of a better classifier, we have developed a new two-stage training method. In this training method, label errors in the training data are first detected by normal supervised learning classification, and then they are unlabeled and used for training of semi-supervised learning. For actual observed data, the classifier with this method achieves an area under the curve (AUC) of 0.9998 and a false positive rate (FPR) of 0.0002 at true positive rate (TPR) of 0.9. This training method saves relabeling effort by humans and works better on training data with a high fraction of label errors. By implementing the developed classifier in the Tomo-e Gozen pipeline, the number of transient candidates was reduced to  $\sim 40$  objects per night, which is  $\sim 1/130$  of the previous version, while maintaining the recovery rate of real transients. This enables more efficient selection of targets for follow-up observations.

Comments: 14 pages, 17 figures, 2 tables. Published in PASJ. The source code is available at this <https> URL.  
Subjects: Instrumentation and Methods for Astrophysics (astro-ph.IM); High Energy Astrophysical Phenomena (astro-ph.HE)  
Cite as: arXiv:2206.12478 [astro-ph.IM]  
(or arXiv:2206.12478v1 [astro-ph.IM] for this version)  
<https://doi.org/10.48550/arXiv.2206.12478>  
Related DOI: <https://doi.org/10.1093/pasj/psac047>

Search or jump to... Pull requests Issues Marketplace Explore

ichiro-takahashi / tomoe-realbogus Public

<> Code Issues Pull requests Actions Projects Wiki Security Insights Settings

main 1 branch 0 tags Go to file Add file Code

ichiro-takahashi Initial Commit 8eb3ee 6 days ago 1 commit

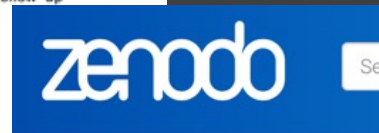
- data/raw/real\_bogus1 Initial Commit 6 days ago
- src Initial Commit 6 days ago
- README.md Initial Commit 6 days ago

README.md

### Deep-learning Real/Bogus classification for the Tomo-e Gozen transient survey

This repository contains the Python scripts for the experiments in the paper:

Ichiro Takahashi, Ryo Hamasaki, Naonori Ueda, Masaomi Tanaka, Nozomu Tominaga, Shigeyuki Sako, Ryou Ohsawa, Naoki Yoshida, Deep-learning real/bogus classification for the Tomo-e Gozen transient survey, Publications of the Astronomical Society of Japan, 2022, psac047, <https://doi.org/10.1093/pasj/psac047>



June 22, 2022

## Dataset for the Real/Bogus classifier in the Tomo-e Gozen transient survey

Takahashi, Ichiro; Hamasaki, Ryo; Ueda, Naonori; Tanaka, Masaomi; Tominaga, Nozomu; Sako, Shigeyuki; Ohsawa, Ryou; Yoshida, Naoki

This is a subset of the dataset used in the following paper:

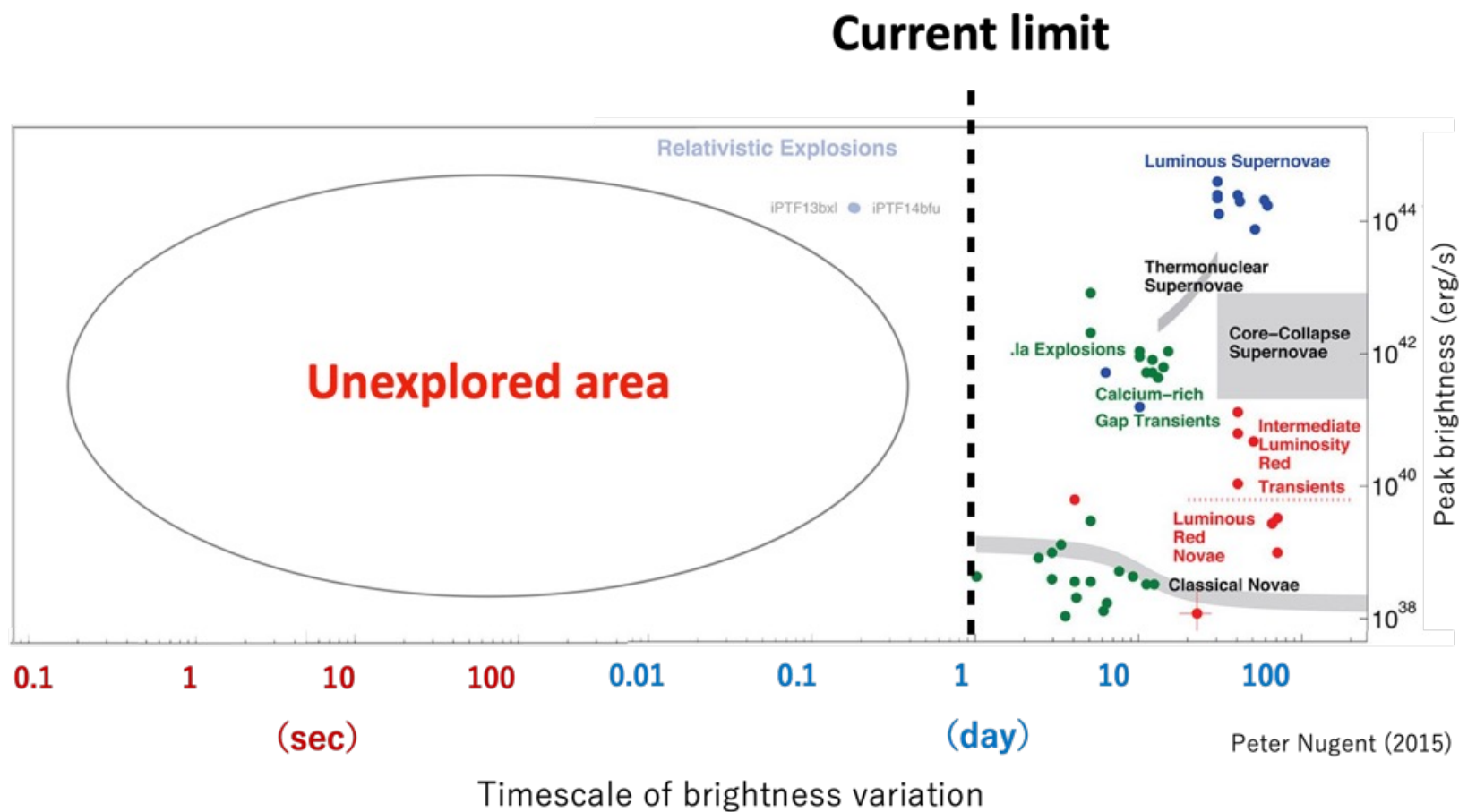
Ichiro Takahashi, Ryo Hamasaki, Naonori Ueda, Masaomi Tanaka, Nozomu Tominaga, Shigeyuki Sako, Ryou Ohsawa, Naoki Yoshida, Deep-learning real/bogus classification for the Tomo-e Gozen transient survey, Publications of the Astronomical Society of Japan, 2022, psac047, <https://doi.org/10.1093/pasj/psac047>

- Tomo-e Gozen transient survey
- 機械学習による新しいReal/Bogus分類器
- 誤検出の数を従来の1/100以下に減らした

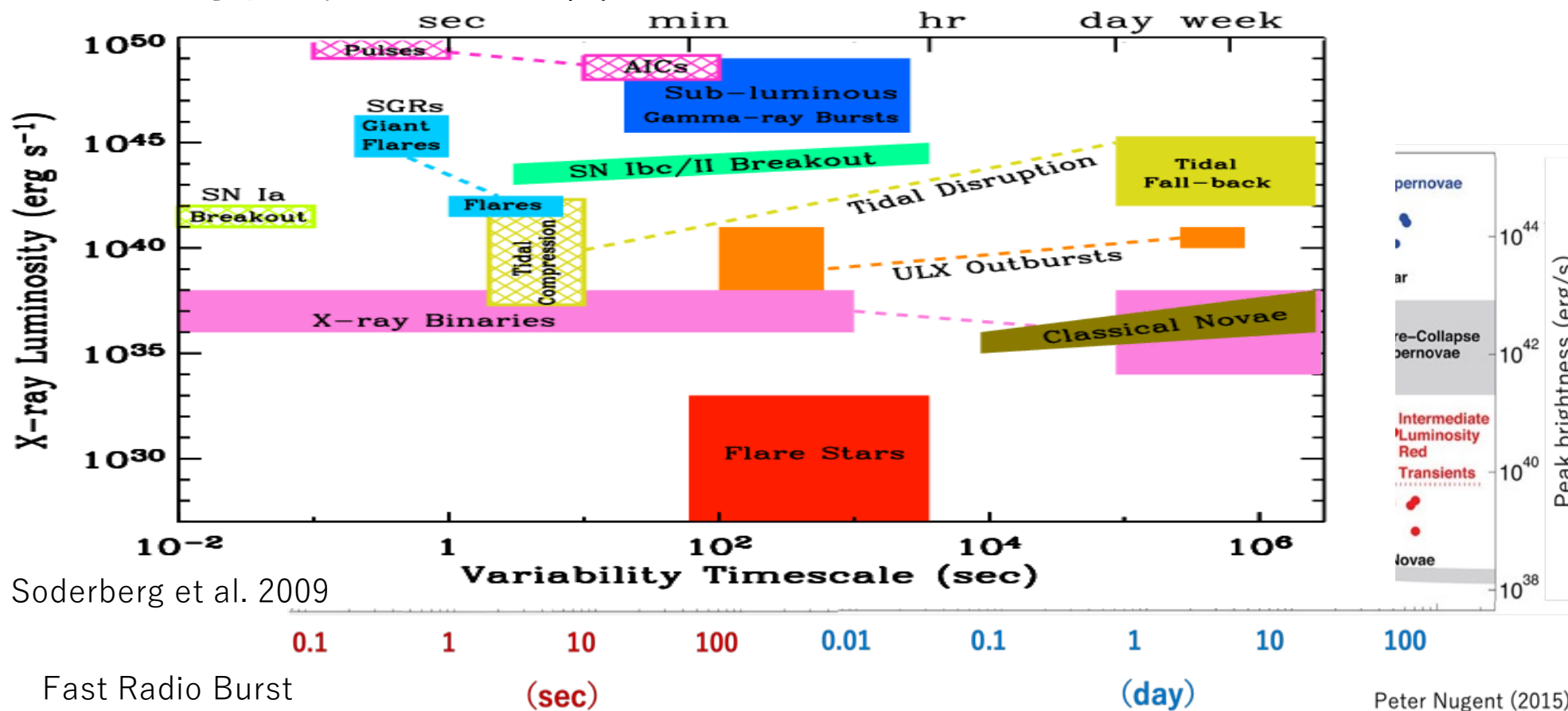
# 目次

- 短時間変動天体
- Tomo-e Gozenn 動画データ
- 機械学習による動画からの天体検出
- 結果
- 議論
- まとめ

# 短時間変動天体



# 短時間変動天体

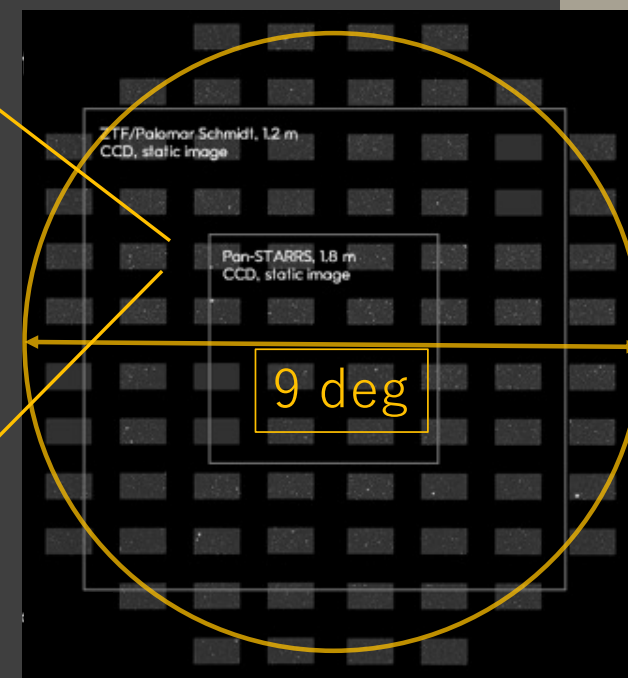
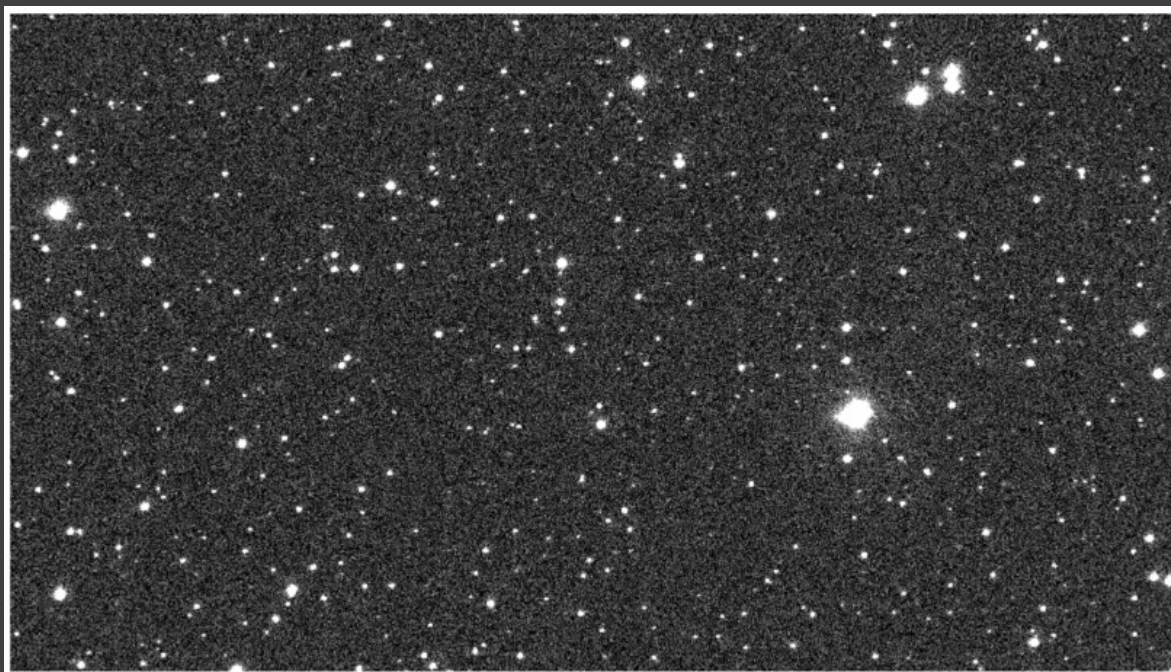
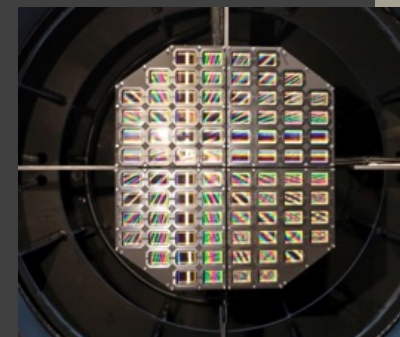


Fast Radio Burst  
(新納さん発表)

M dwarf超高速フレア Timescale of brightness variation  
(逢澤さん発表)

# Tomo-e Gozen 動画データ

- 2fpsの動画データを撮れる
- 全ての動画をチェックしきれていない



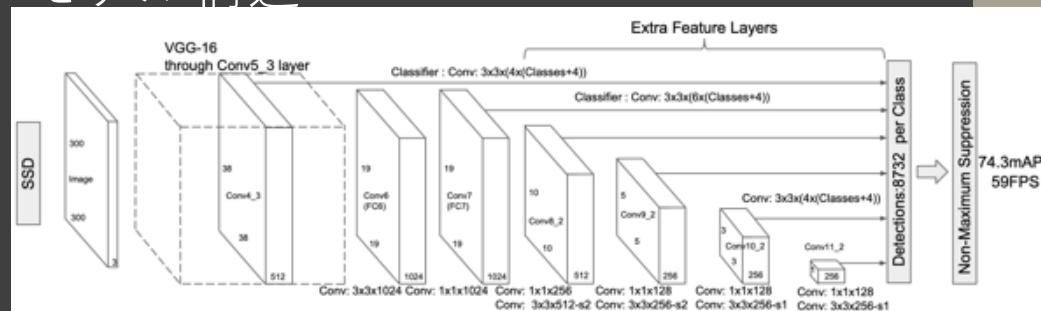
機械学習技術を応用して短時間変動天体を高速に検出する



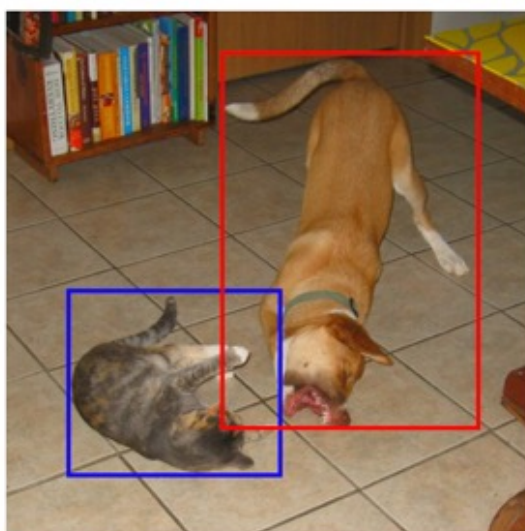


# SSDの検出方法

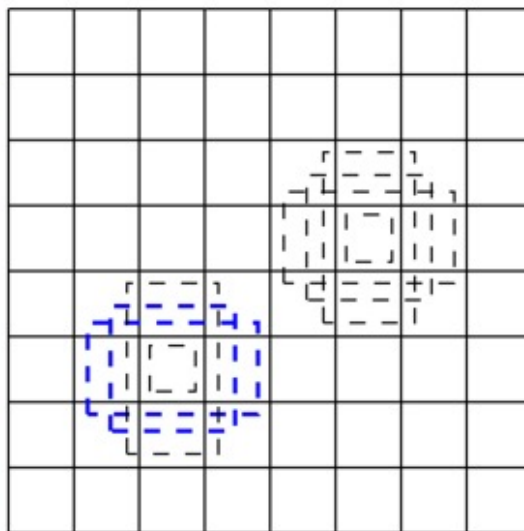
## モデル構造



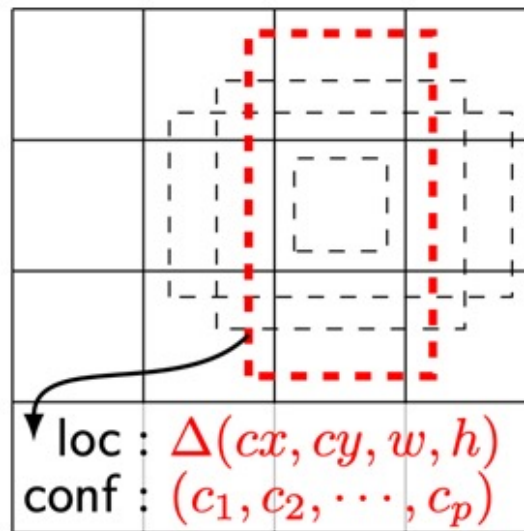
- 様々な大きさ、形のbounding boxで画像を走査してその都度種類を分類



(a) Image with GT boxes



(b)  $8 \times 8$  feature map

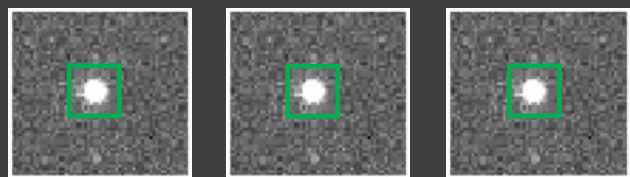


(c)  $4 \times 4$  feature map



# 検出テスト

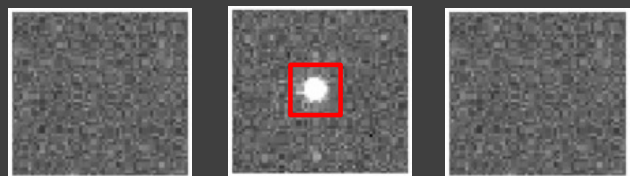
- 動画データの各フレームでSSDによる物体検出
  - NTTチームでモデルを開発
- 連続する3フレームを入力
- 1フレームだけ光る天体(Flash)を検出



Class0:  
Normal Star



Class1:  
Background



Class2:  
Flash

- データセット
  - 本物のTomo-e Gozenの動画データ
- 学習データ
  - 2019年11月
  - 2fps
  - 3.5TB
- テストデータ
  - 新納さん提供のFRB探査データ
  - 2020年12月
  - 1fps
  - 0.65 TB

# 結果

## Selection

- 総検出数：43845
- Bounding Boxの面積とスコアで抽出
- 抽出後：404

## 統計（目で見て分類）

星の端： 206

移動天体： 87

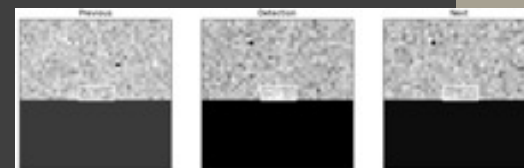
流星： 64

宇宙線： 25

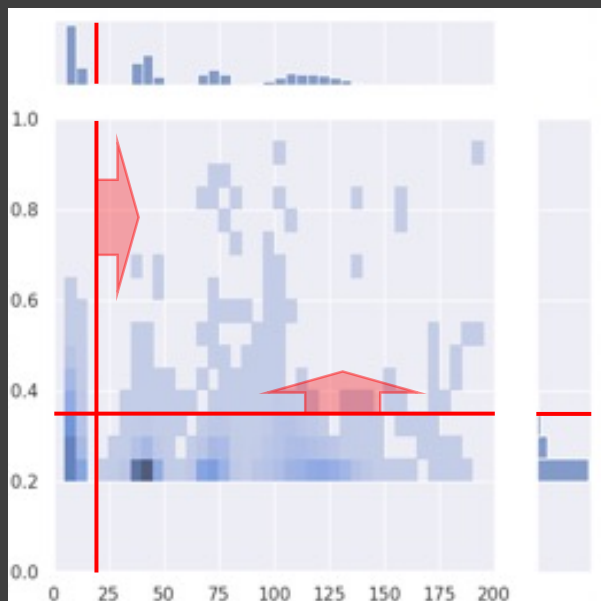
画像の端： 16

Flash候補： 5

曇り： 1



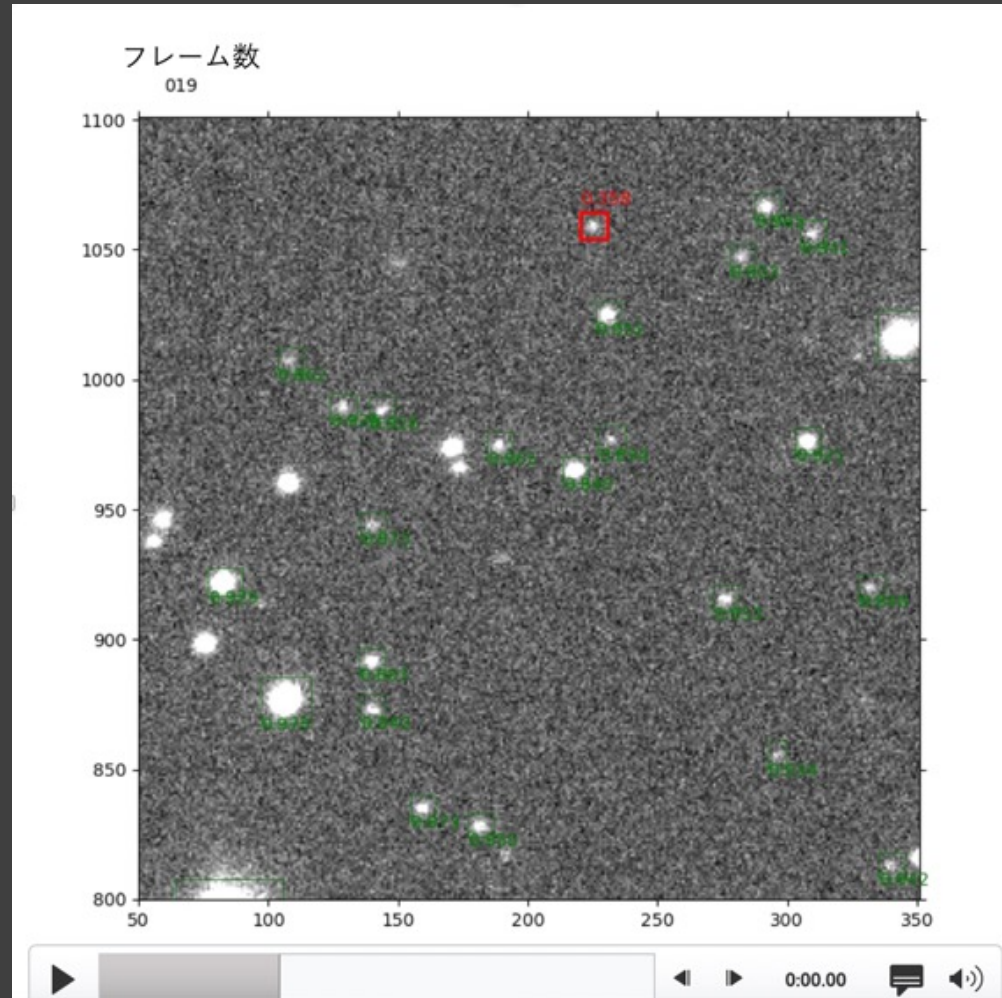
スコア



Boxの面積

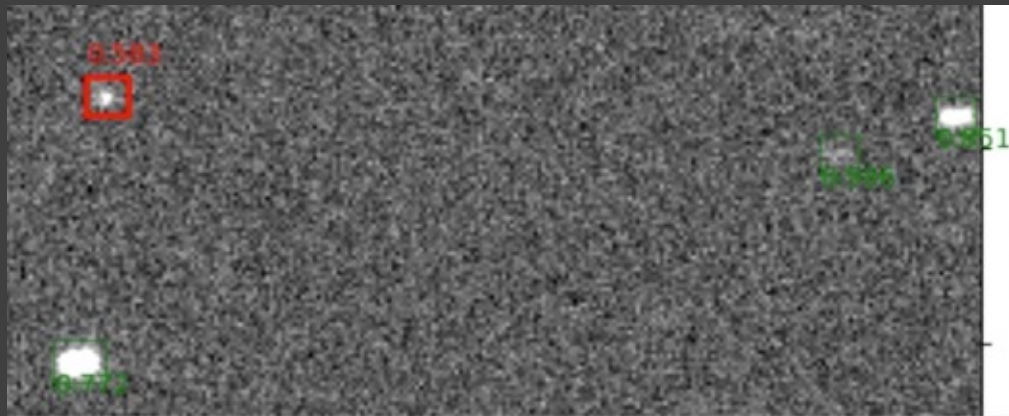
# Flash候補 1

19フレーム目

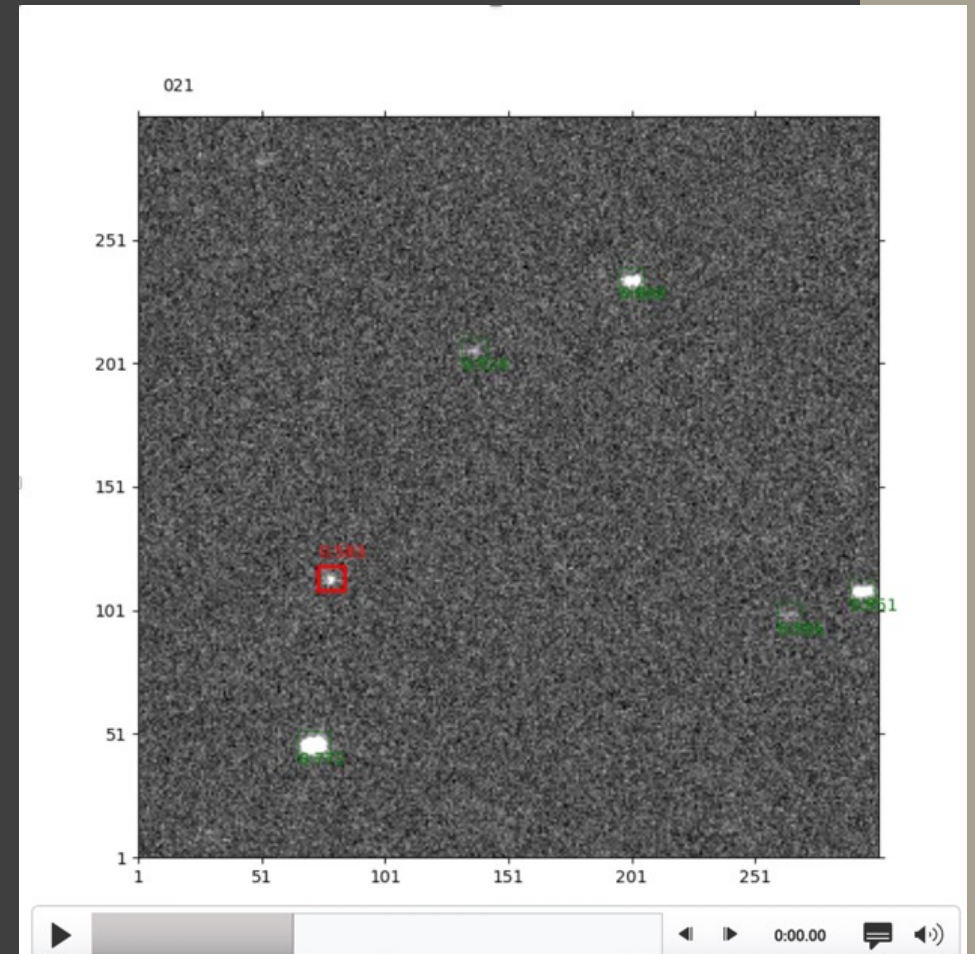


# Flash候補 2

21フレーム目

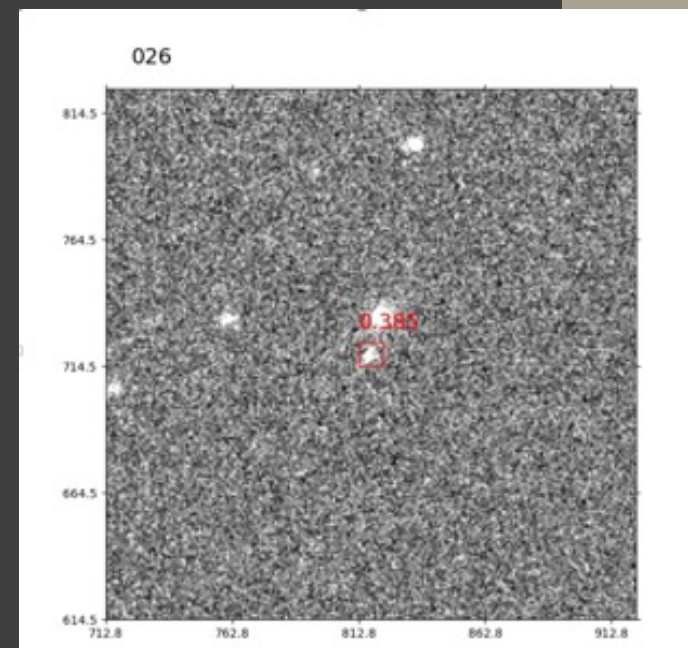
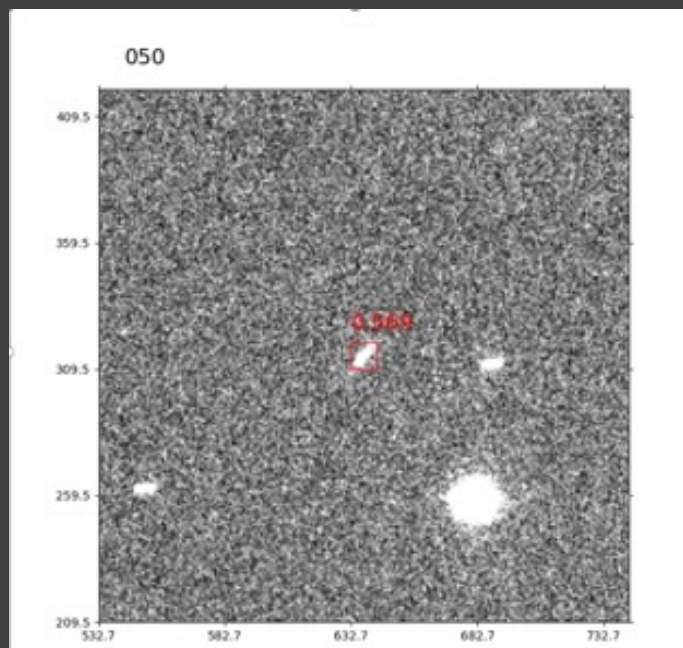
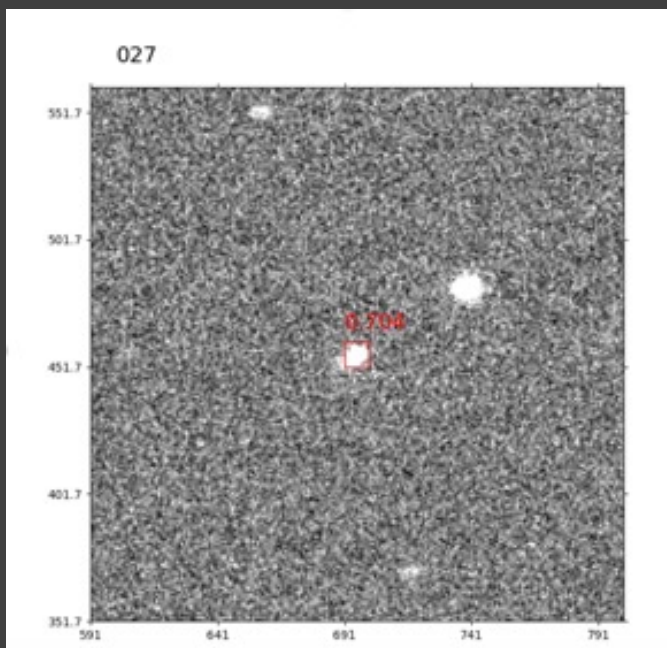


点源として写っている-> 1秒以下の変動



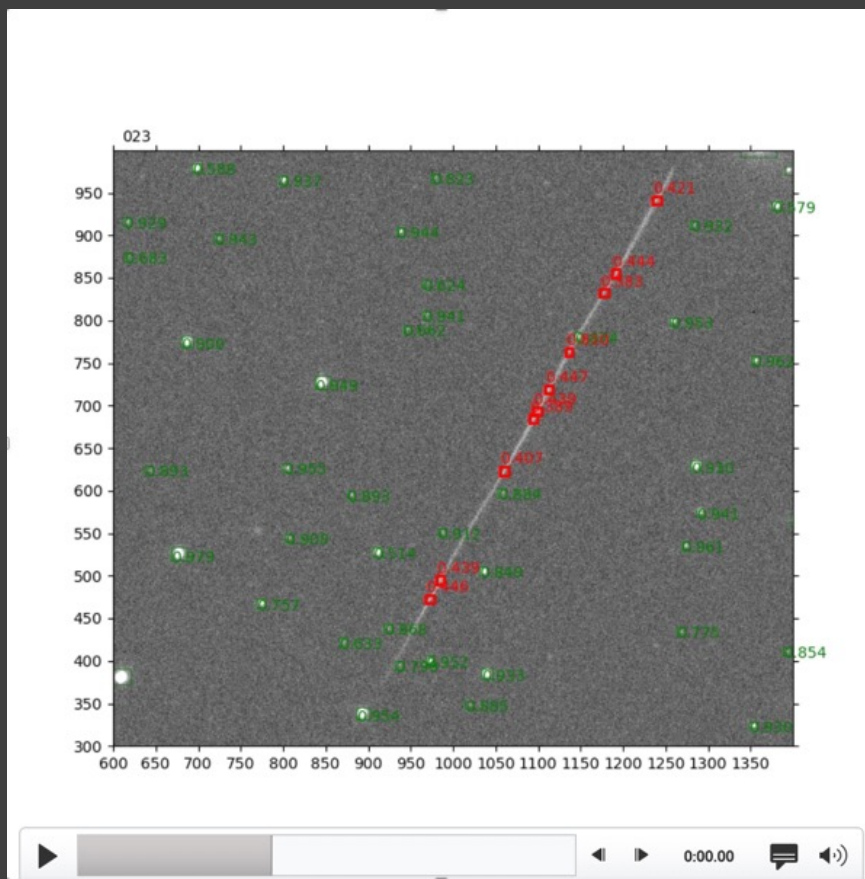


# 他の候補

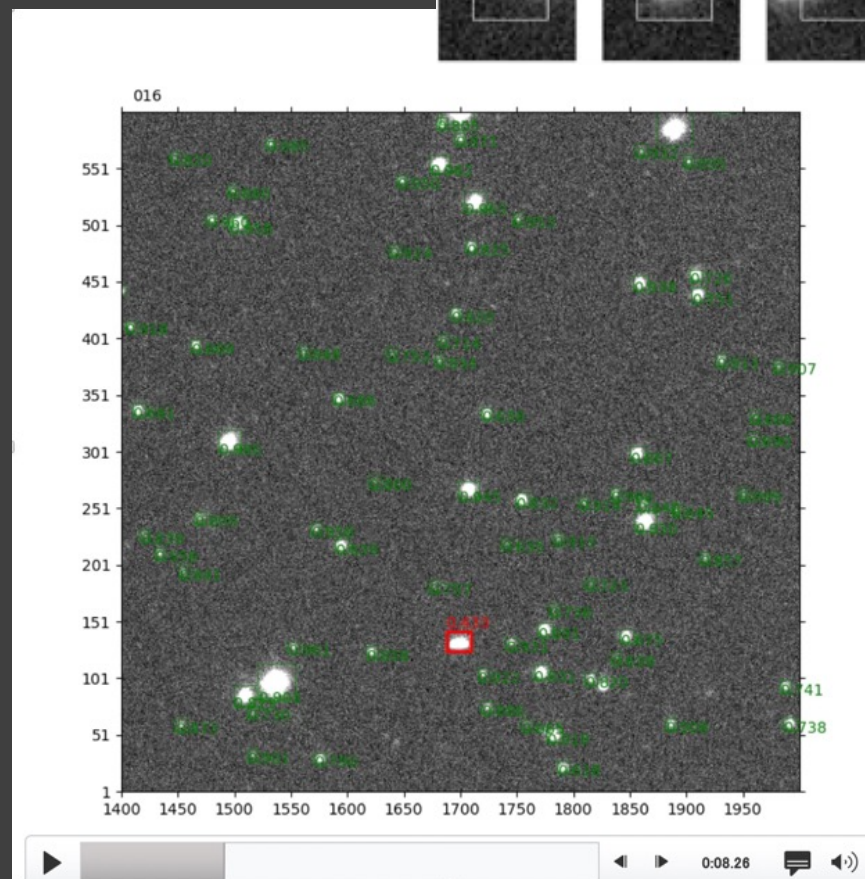


# 誤検出例

## 流星 (23フレーム目)



## 移動天体



# 議論

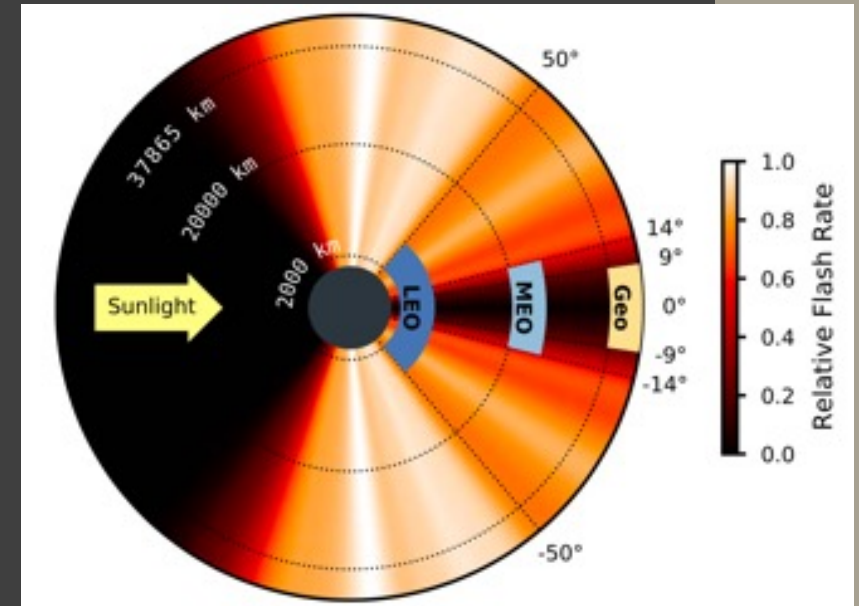
## • 処理速度

- 1 センサ 1 GPU (GeForce RTX 3090)ならリアルタイムで処理可能
- 60秒の動画データ (1fps) に対して40秒で処理

## • 候補天体の起源

- 太陽との離角：120~145 deg
- 人工衛星やデブリによる太陽光反射

Evryscope による地球影に対する flash rate 分布



Corbett et al. 2020

# 議論

## レート比較

• 2016 菊池修論 (FRB対応天体探査) :

• 地球影外  $\sim 4 \text{ event deg}^{-2} \text{ hour}^{-1}$  (10候補, 2fps)

• 地球影内  $\sim 1 \text{ event deg}^{-2} \text{ hour}^{-1}$  (4候補, 2fps)

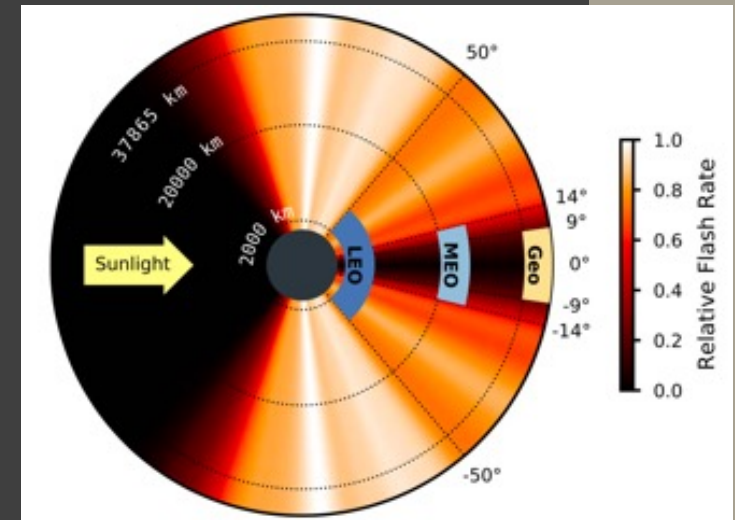
• 今回 (地球影外) :  $\sim 1 \text{ event deg}^{-2} \text{ hour}^{-1}$  (5候補, 1fps)

• Evrscope :  $\sim 0.04 \text{ event deg}^{-2} \text{ hour}^{-1}$

(限界等級  $m_g = 14.2$ , Corbett et al. 2020)

• 今後サンプル数を増やしてより正確な検証へ

Evrscope による地球影に対する flash rate



Corbett et al. 2020



# まとめ

- 短時間変動天体の検出のため、Tomo-e Gozenn 動画データへ機械学習を応用
- Single Shot Multibox detector (SSD)による物体検出
  - 1frameだけ光る天体
- 5つの候補天体を検出
  - 人工衛星やデブリによる太陽光反射
- 1 センサ1GPUならリアルタイム処理が可能
- レートは以前のものと同様に一致
  - 今後サンプル数を増やしてより正確な検証へ