

▼ Tomo-e Gozen transient surveyにおける  
深層学習によるReal/Bogus分類の改善

高橋一郎

東北大学

# 目次

- Tomo-e Gozen
- Real/Bogus分類と問題点
- 新モデルでの分類
- 実際のテストデータでの検証
- 新モデルの実装

+



# Tomo-e Gozen project

## Kiso Schmidt telescope

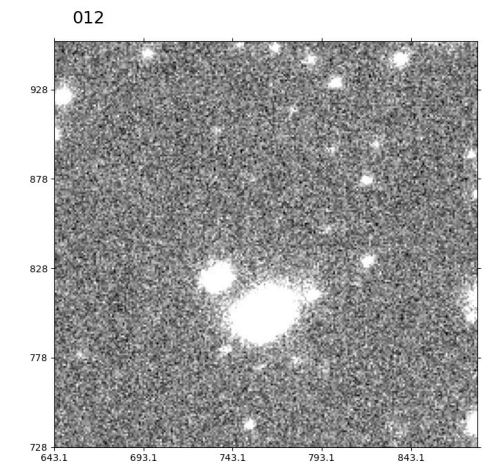
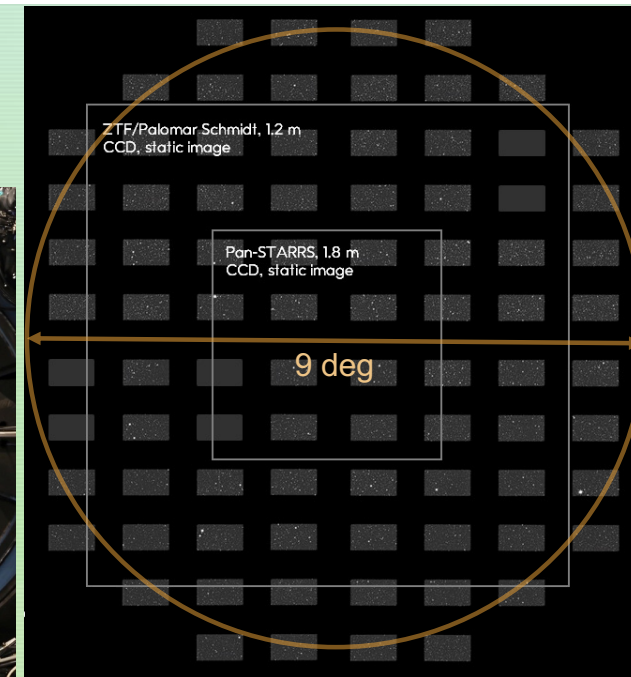
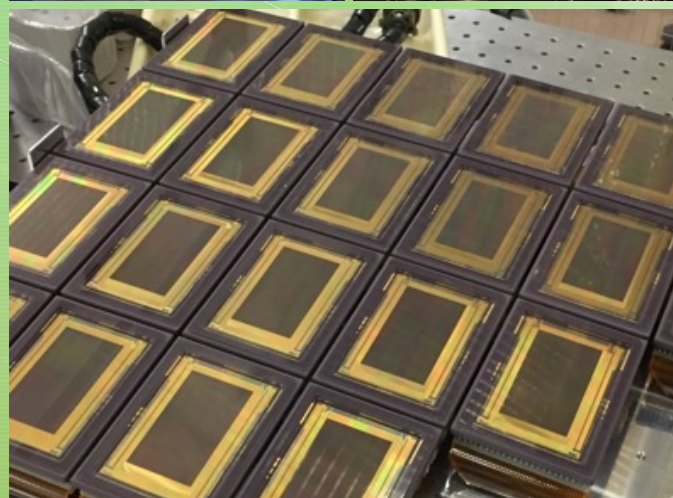
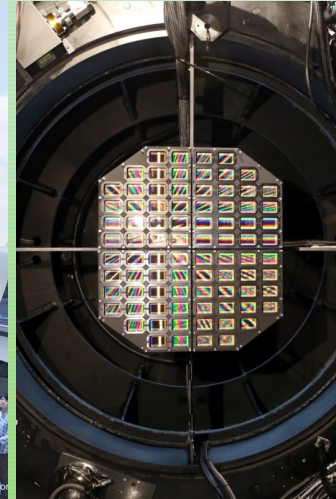
- 1.05 m
- FOV : 9 deg in diameter

## Tomo-e Gozen camera

- wide-field CMOS imager
- 84 CMOS sensors
- 20 deg<sup>2</sup> sky in 2 fps
- 30 TB /night

## CMOS image sensor

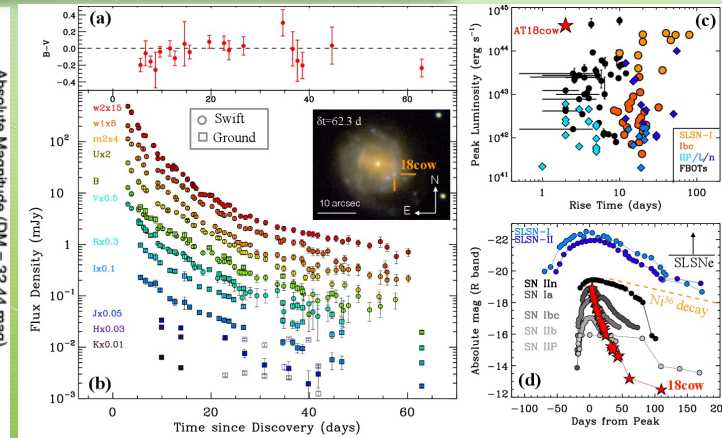
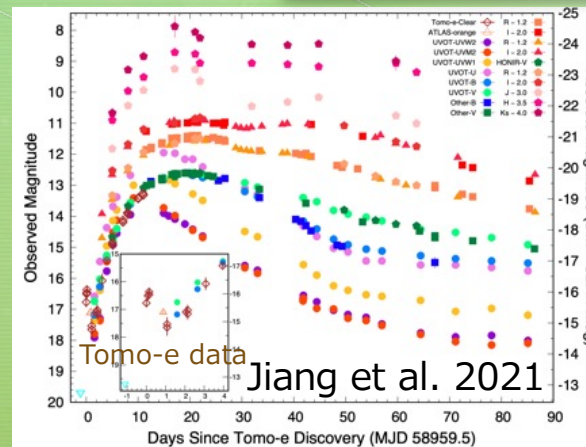
- 2000 x 1128 pixels
- Non-cooling
- low read noise of 2 e<sup>-</sup>



# Tomo-e Gozen transient survey

- 2種類のサーベイ
- All-sky survey
  - 露出時間 : 9秒 (18 frame)
  - $\sim 12000 \text{ deg}^2/\text{day}$
- High-cadence survey
  - 露出時間 : 6秒 (12 frame)
  - $\sim 3000 \text{ deg}^2/ 30 \text{ min}$
  - 1日に3回観測
- 主なターゲット : 最初期超新星  
Rapid transient

Tomo-e Gozen Sky Atlas (前回WS瀧田発表)

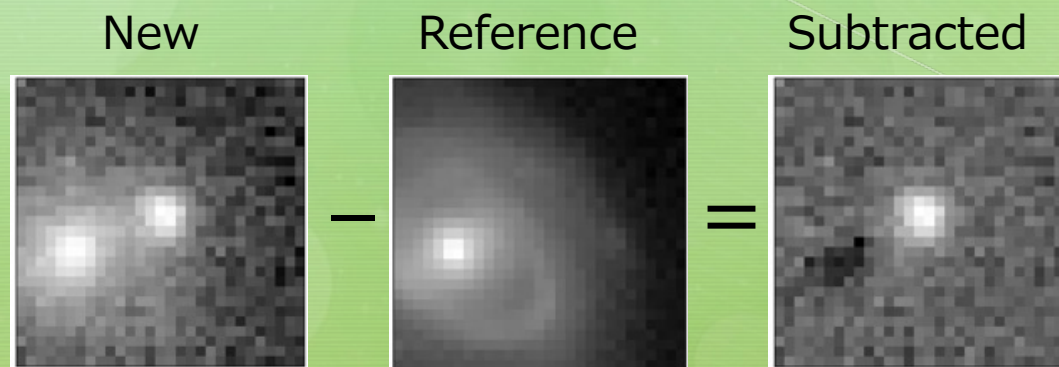


昨日のKavli IPMUのプレスリリース

Margutti et al. 2019

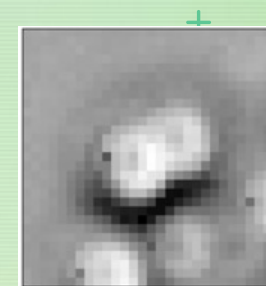
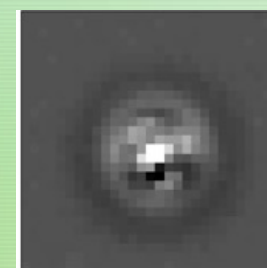
# Transient detection

- 差分を取ることでtransientを検出
- ReferenceにはPS1の画像を使用
  - 現在Tomo-e referenceへ切り替え中
- 誤検出 (Bogus) が多い ( $10^6/\text{night}$ )
  - 機械学習で分類
- Real or Bogus

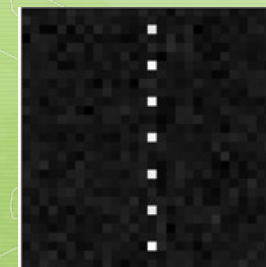
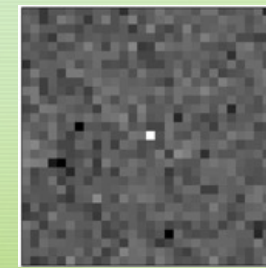


## Bogus examples

差分の失敗



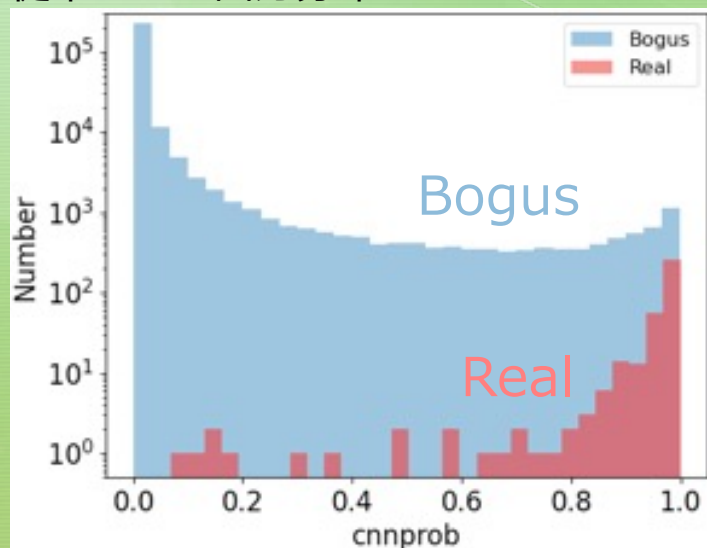
観測装置由来



# Previous Real/Bogus Classification

- モデルはシンプルなCNN
  - 入力: 29x29 image x 3 (ref., new and sub.)
- 学習データ
  - Real : シミュレートした星を埋め込み
  - Bogus : 実際のもの
  - 数百万サンプル

従来CNNの出力分布



従来CNNモデルの構造

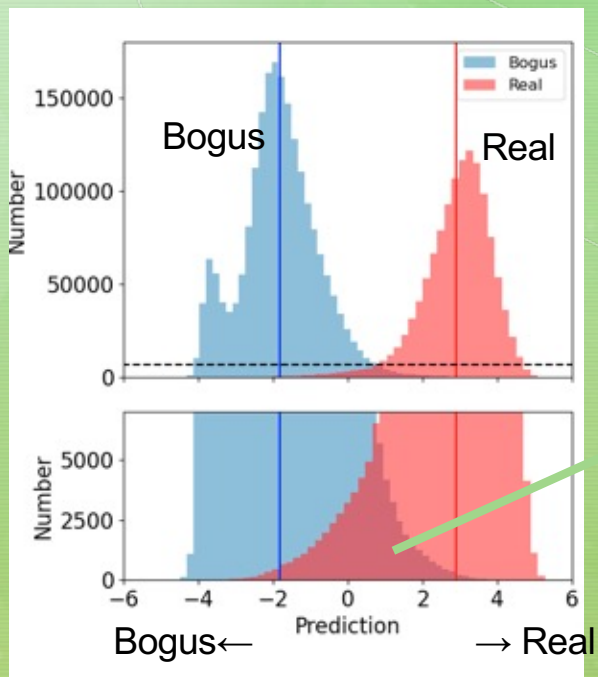
Layer (type)	Output Shape	Param #
conv2d_1 (Conv2D)	(None, 25, 25, 32)	2432
max_pooling2d_1 (MaxPooling2D)	(None, 12, 12, 32)	0
conv2d_2 (Conv2D)	(None, 10, 10, 64)	18496
max_pooling2d_2 (MaxPooling2D)	(None, 5, 5, 64)	0
dropout_1 (Dropout)	(None, 5, 5, 64)	0
flatten_1 (Flatten)	(None, 1600)	0
dense_1 (Dense)	(None, 128)	204928
dense_2 (Dense)	(None, 64)	8256
dense_3 (Dense)	(None, 1)	65
Total params: 234,177		
Trainable params: 234,177		
Non-trainable params: 0		

分類成績が頭打ちに  
RealがBogusに埋もれる  
追観測のターゲット選定の障害

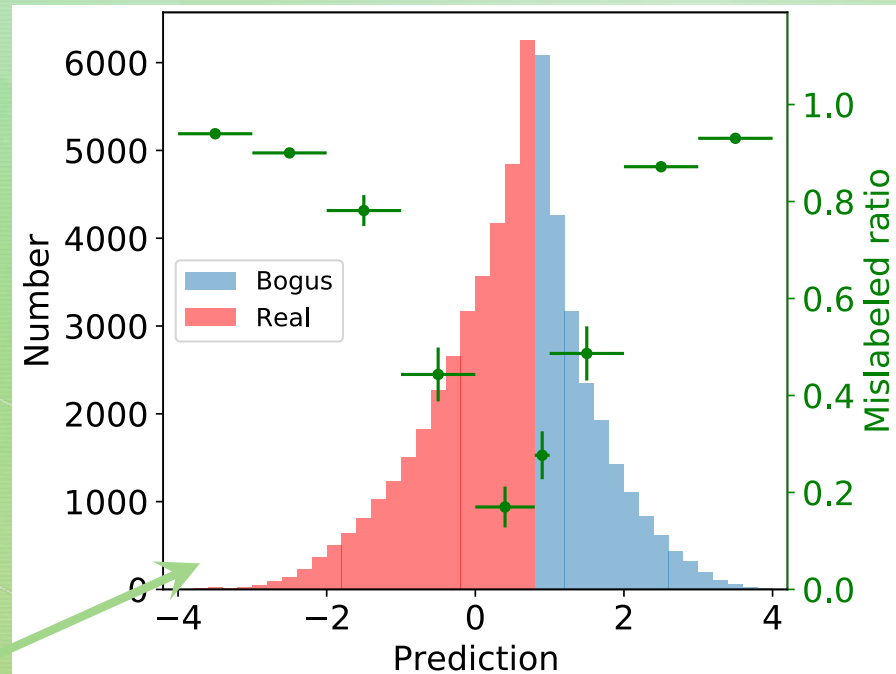
# Cleaning mislabel in training data

- 学習データにラベル間違いが存在
- 学習データ自体をCNNで分類
  - 誤分類しているものを標本調査
  - ラベル間違いの比率を計算

学習データの出力分布



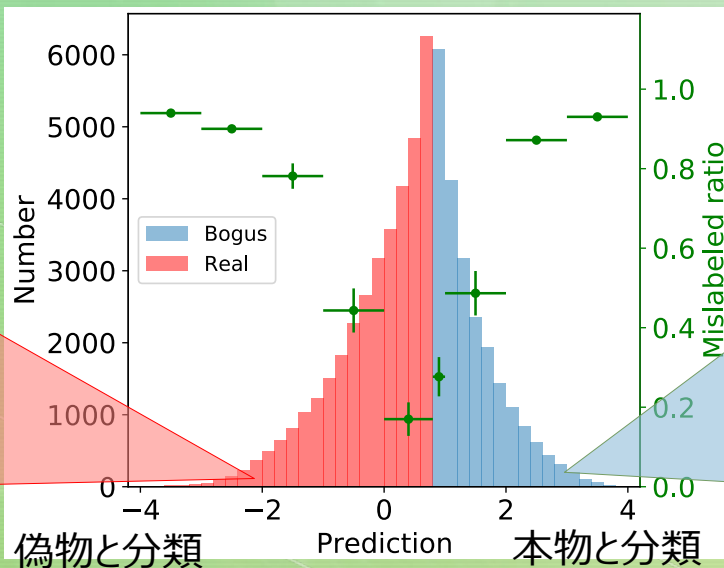
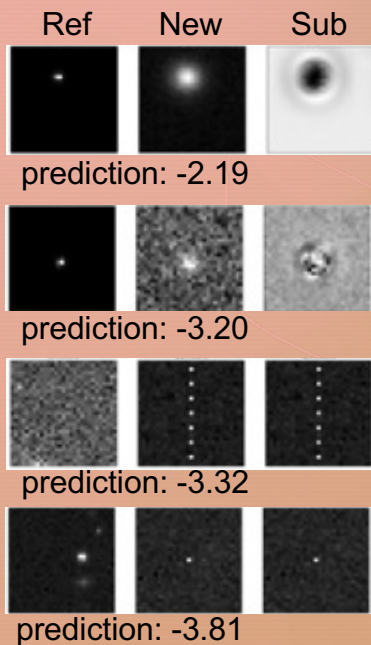
後分類の分布と誤ラベルの比率



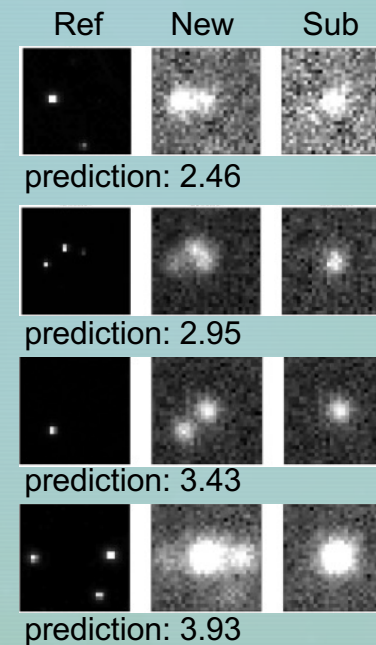
裾野に行けば行くほど誤ラベルが増える

# Examples of mislabel

Label : Real



Label : Bogus

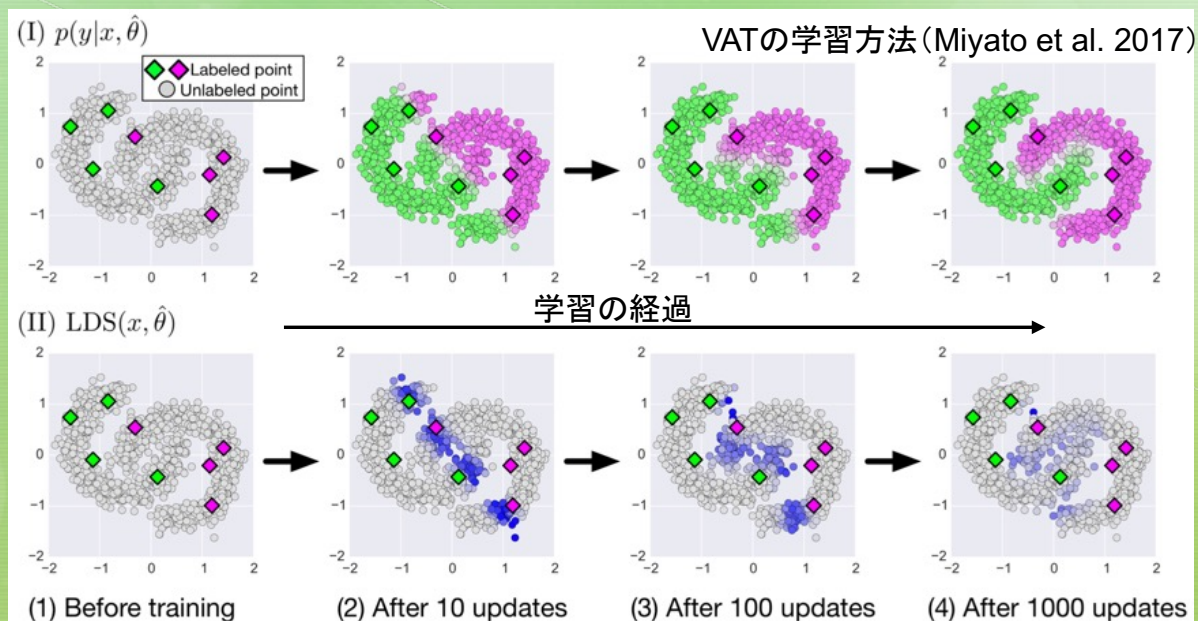


- 機械は正しく分類している
- ラベル間違いの混入率 : 0.6% (Bogus), 1.0% (Real)
- 新手法 : 分類を2段階に分ける
  - 1段階目で誤分類している学習データを「ラベルなし」にして半教師あり学習を行う
  - Virtual Adversarial Training (VAT, Miyato et al. 2016)



# Virtual Adversarial Training (VAT)

- 敵対的学習の一種
  - 分類器が誤りやすいノイズを付加したデータで学習 -> ノイズ耐性
- 予測分布が最も大きく変わる方向へノイズを加え、予測分布の差が小さくなるように学習
  - Local Distributional Smoothness (LDS)
  - LDS 小 -> 境界がデータ集合から遠くなる-> よりロバスト
  - LDSの計算には正解ラベル情報がいらぬ -> 半教師あり学習が可能



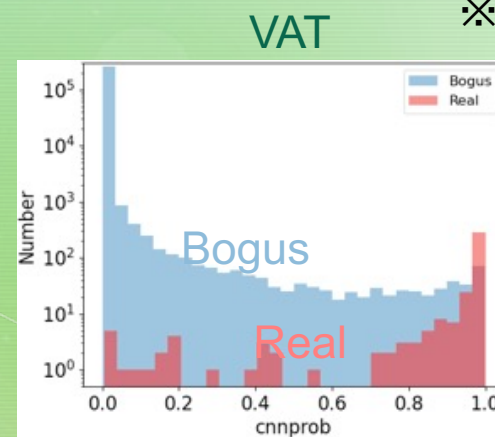
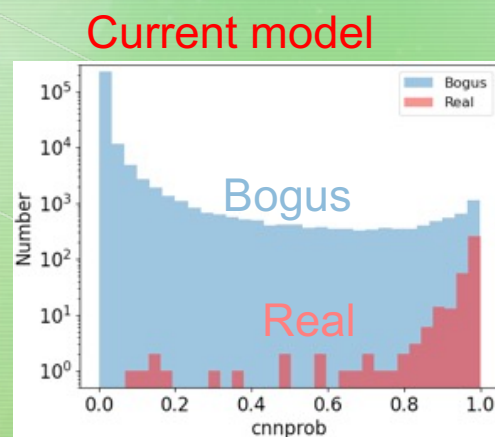
# Evaluation with actual test dataset

	Accuracy	Precision	Recall	F1 score
Current model	0.974	0.050	0.981	0.095
VAT	0.998	0.436	0.942	0.596

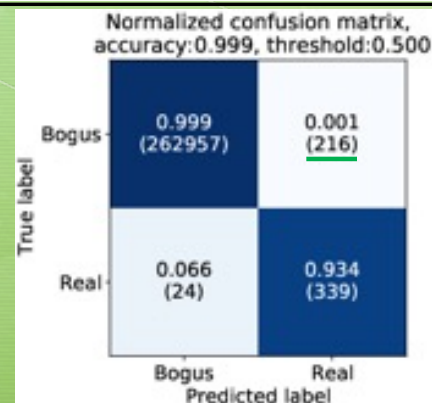
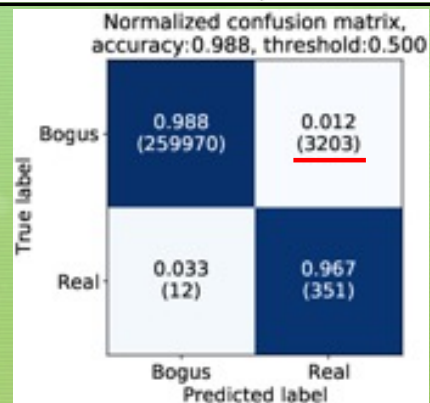
+

※しきい値: probability=0.5

出力分布

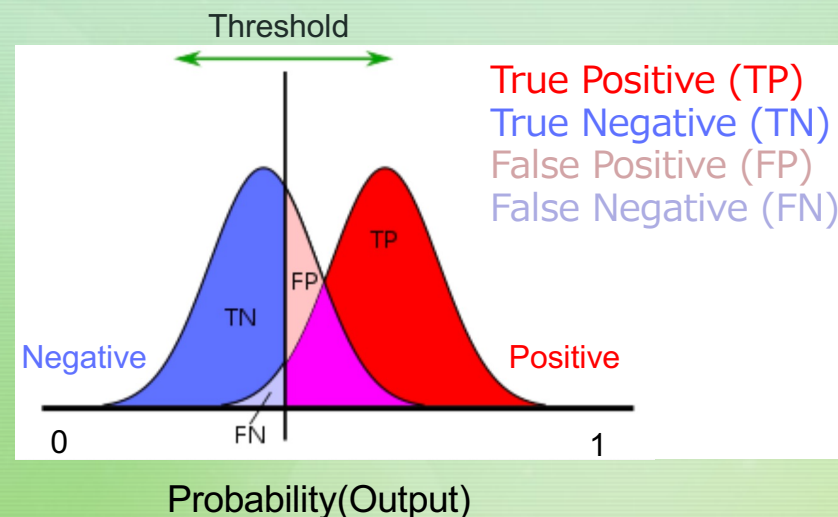
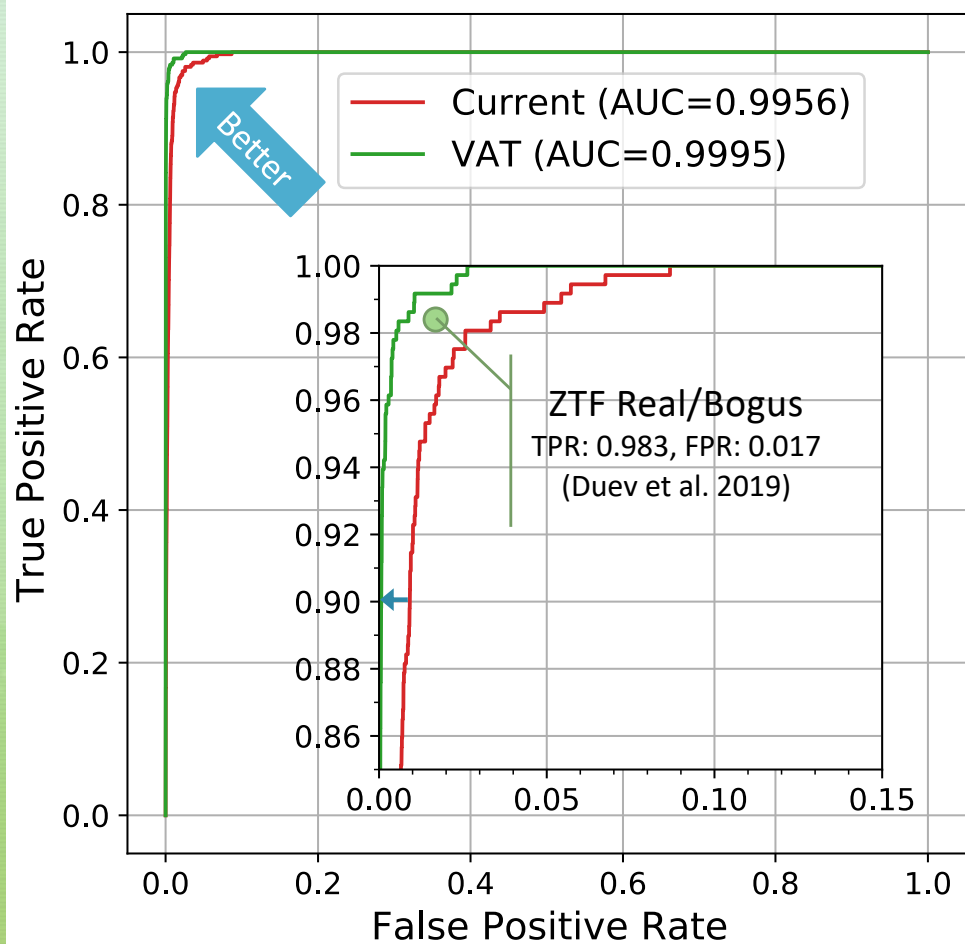


Confusion matrix



# Evaluation with actual test dataset

## ROC curve



$TPR = TP / (TP + FN)$   
: 正しく拾えているか

$FPR = FP / (TN + FP)$   
: 誤検出の少なさ

## Confusion matrix

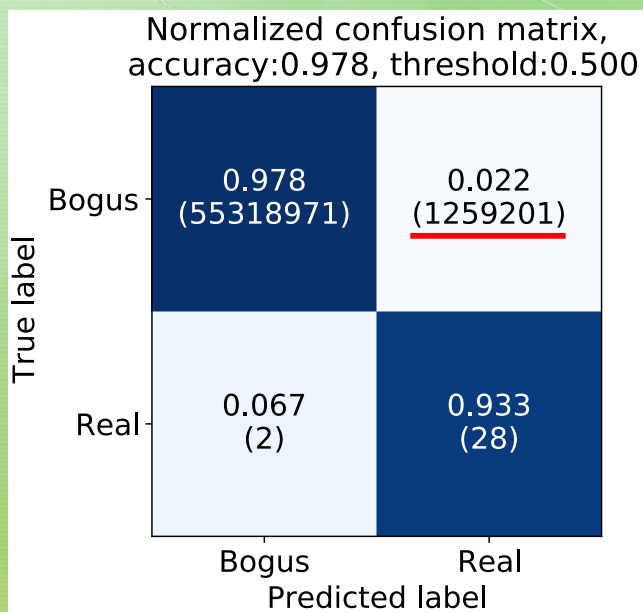
Actual	Bogus	TN	FP
	Real	FN	TP
		Bogus	Real
		Predicted	

Model	AUC(ROC)	FPR@TPR=0.9
Current	0.9956	0.00916
VAT	0.9995	0.00065

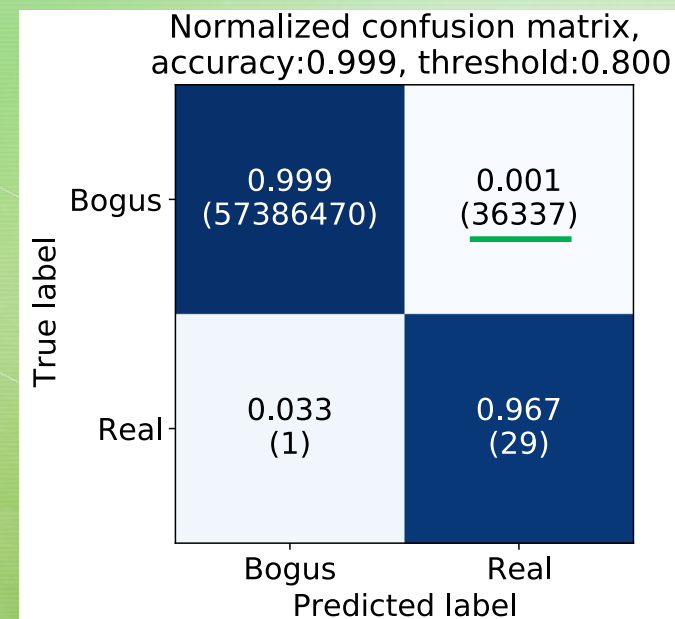
# Implementation of new model

- 今年の5月末にパイプラインに新モデルを実装
- 実装後の分類成績
  - 1週間の間に検出された候補(約6000万サンプル)

## Simple CNN



## VAT



# Misclassified samples

- 誤分類しているものをチェック
  - FNの1つとFPの~1%

Normalized confusion matrix,  
accuracy:0.999, threshold:0.800

True label	Bogus	0.999 (57386470)	0.001 (36337)
	Real	0.033 (1)	0.967 (29)
		Bogus	Real
		Predicted label	

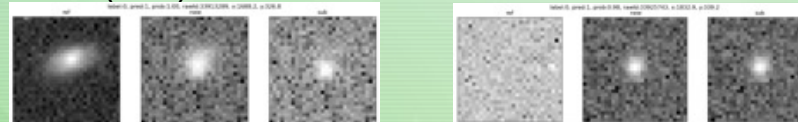
## False Negative (1)

Low S/N, Large galaxy (1)

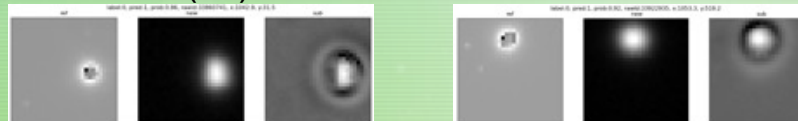


## False Positive (341/36337)

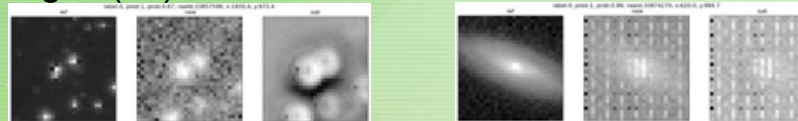
Real? (106)



Mask in ref (65)



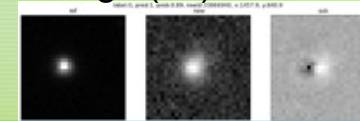
Bogus (64)



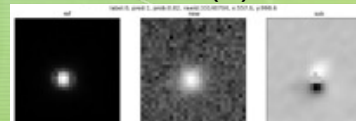
Low S/N (43)



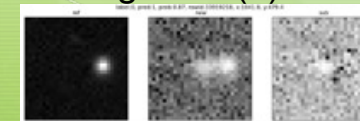
Moving? (29)



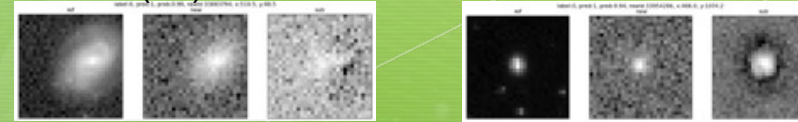
Mask in sub (6)



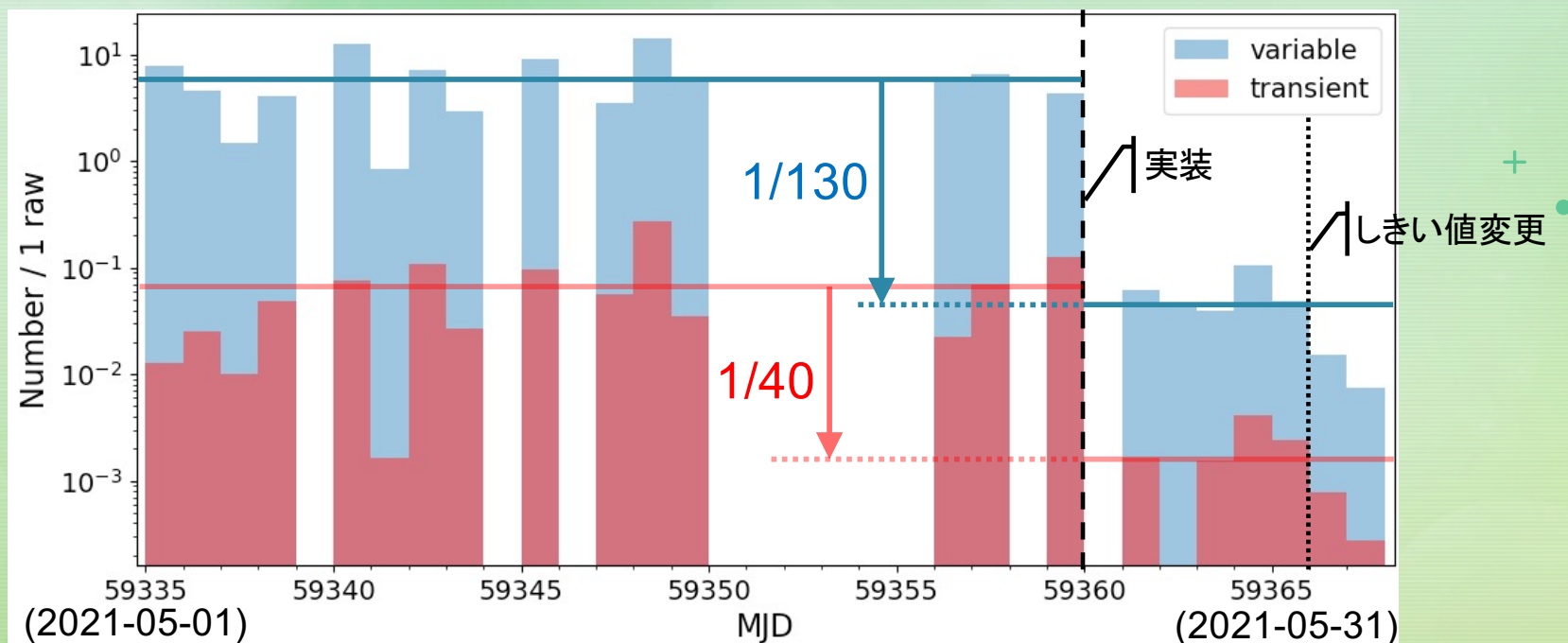
Tracking error (1)



Unclear (27)



# Transient candidates registered in DB

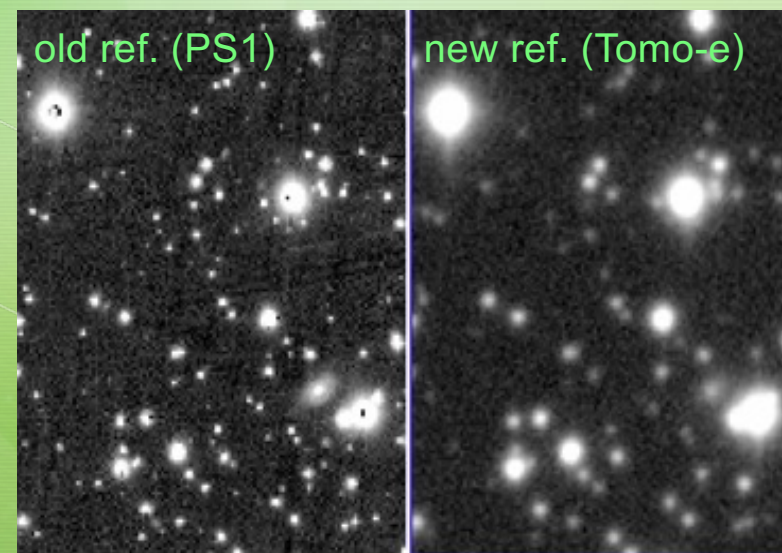
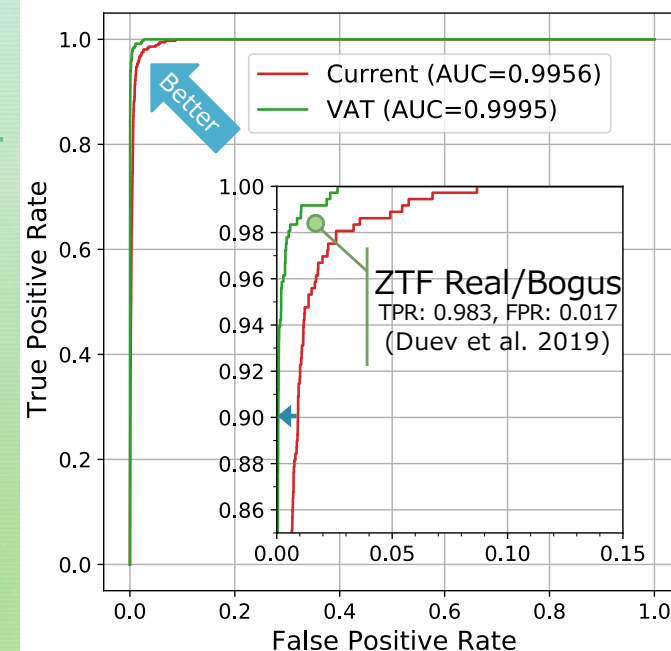


		実装前	実装後
本物 (TNS天体) の確率	variable	1 / 640494	1 / 1004
	transient	1 / 8738	1 / 90

transient候補が6000天体/dayから150天体/day (人がチェックできるレベル) へ

# まとめと今後の予定

- Tomo-e Gozen transient survey
- Real/Bogus分類
- ラベル間違いの存在
- 仮想敵対的学習 (VAT) を適用
- 1%のラベル間違いを機械が修正することで  
性能が1桁以上向上
- 今後 : reference更新
  - モデルも更新
  - FPRがさらに減る可能性



# 最後に

- 観測装置の出力は宝が眠ったデータの山
  - できるだけ出力やパラメータを残してほしい
  - 機械への入力が多いほど良い
  - データが巨大：Tomo-e Gozen, TriCCS
  - 機械学習などによるスマートな処理が必須
- 教訓：機械は間違った学習データも学習してしまう
  - 学習データをきれいに作る
    - 人の目や手が入る以上限界はある（人によるブレ、時間によるブレ）
  - 学習データのクリーニングも機械に任す
- 装置開発への機械学習応用のご相談お待ちしております