Equivariant Neural Network for Estimating Black Hole Spin and Inclination Angle

Takahashi et al., in prep.

<u>Mikiya Takahashi</u>
<u>(National Institute of Technology, Tokyo College)</u>
Yuta Asahina (U. of Tsukuba), Ken Ohsuga (U. of Tsukuba),
Rohta Takahashi (NIT, Tomakomai C.), Takumi Ogawa (NIT, Ibaraki C.),
Masayuki Umemura (U. of Tsukuba)

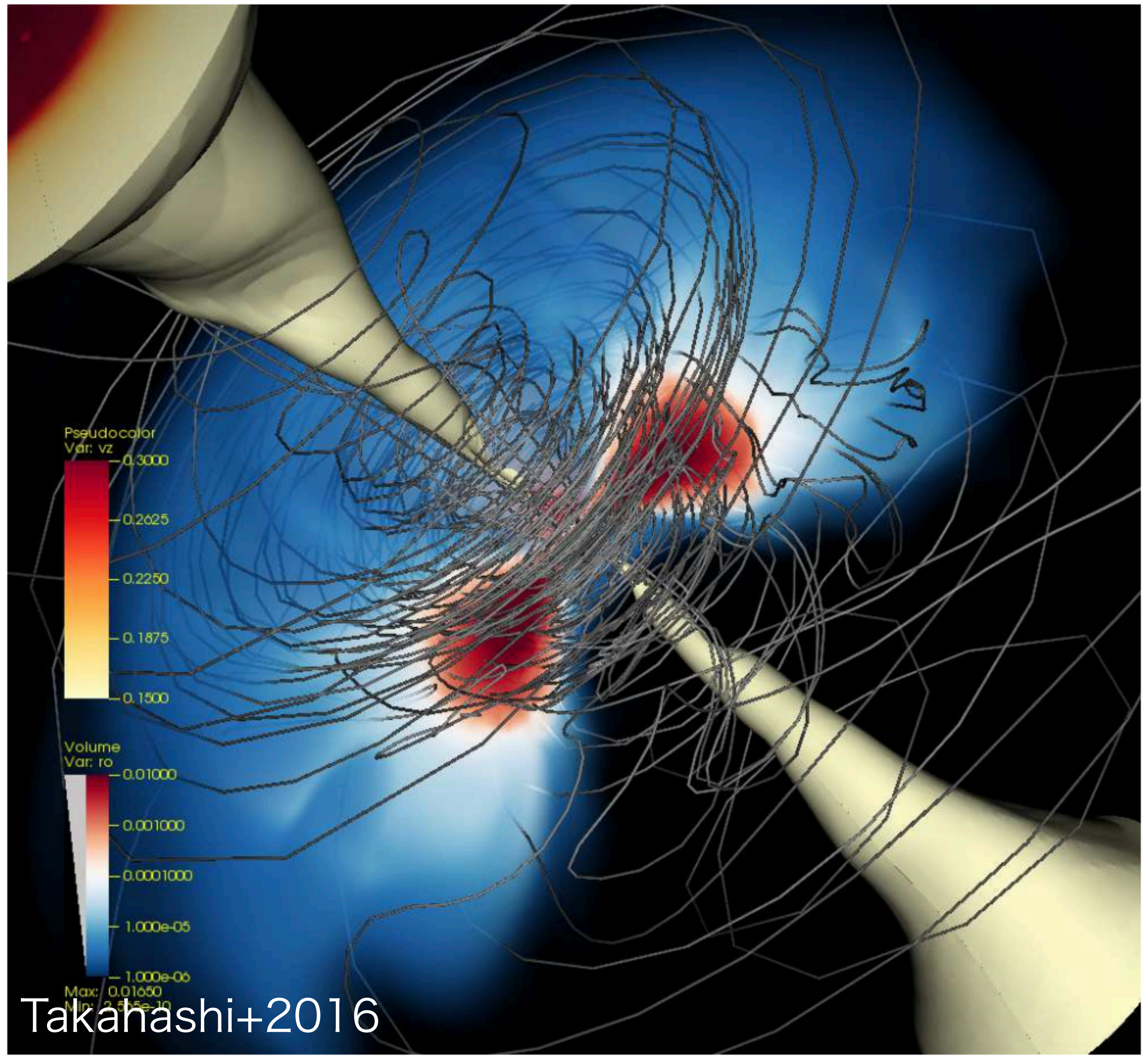
2025/9/15-9/19   East Asian Numerical Astrophysics Meeting 10

website

Why is Black Hole Spin Important?

- **For verification of gravitational theory**
  - black hole (BH) spacetime is (almost) determined by the mass and spin

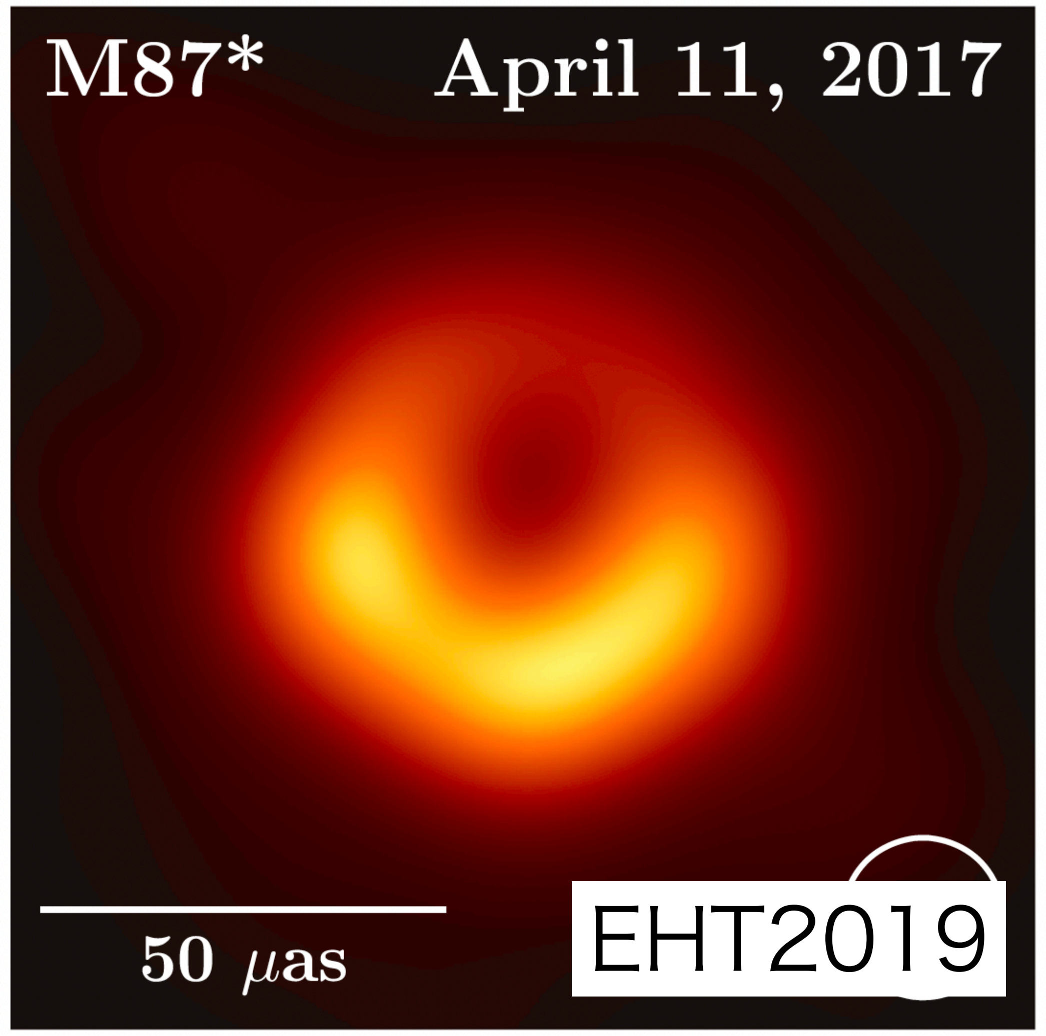
- **For clarification of mechanism of high energy phenomena**
  - black hole spin drives powerful outflows (jet or wind)
  - these outflows affect the structure not only in the vicinity of the black hole, but also on a galactic scale

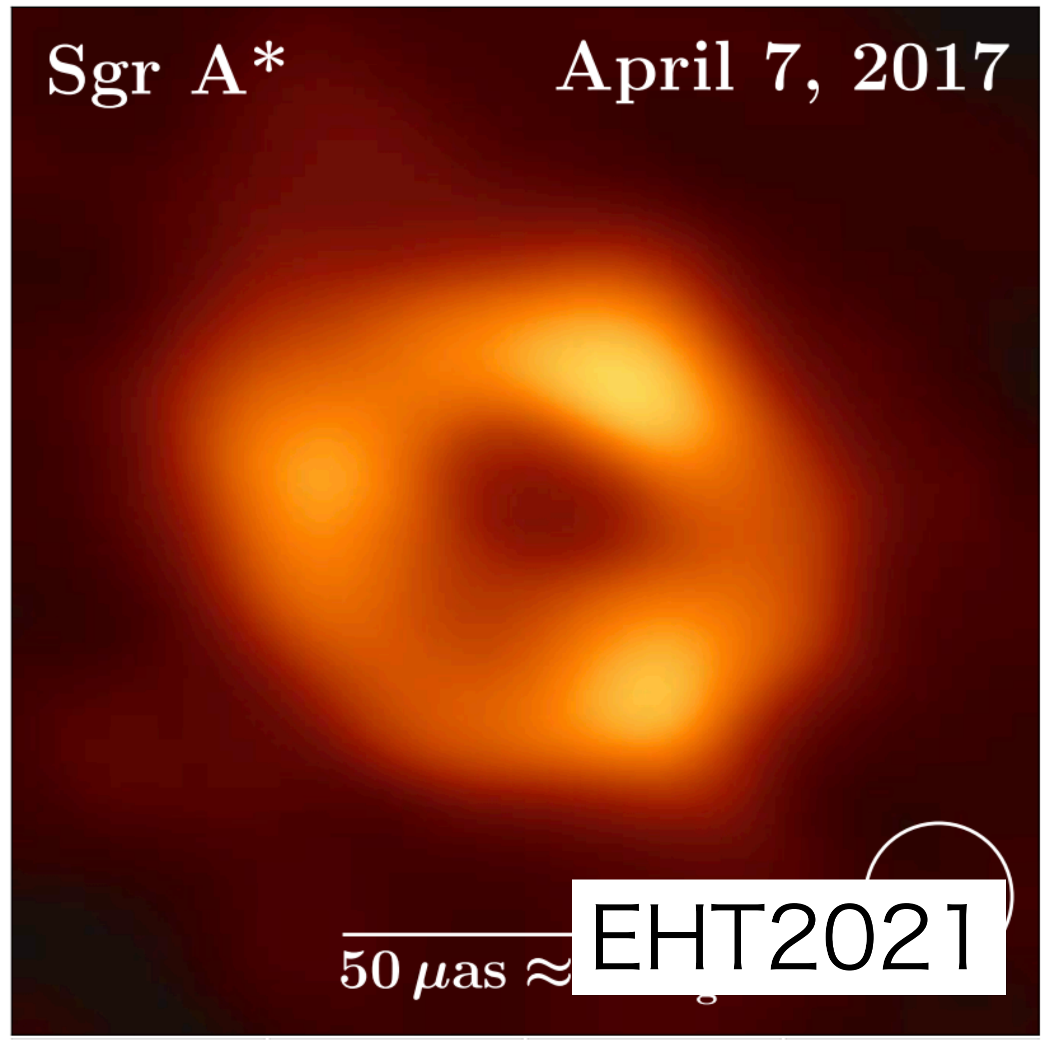


It is essential to determine the black hole spin for considering the structure from the vicinity of the BH to galactic scale.

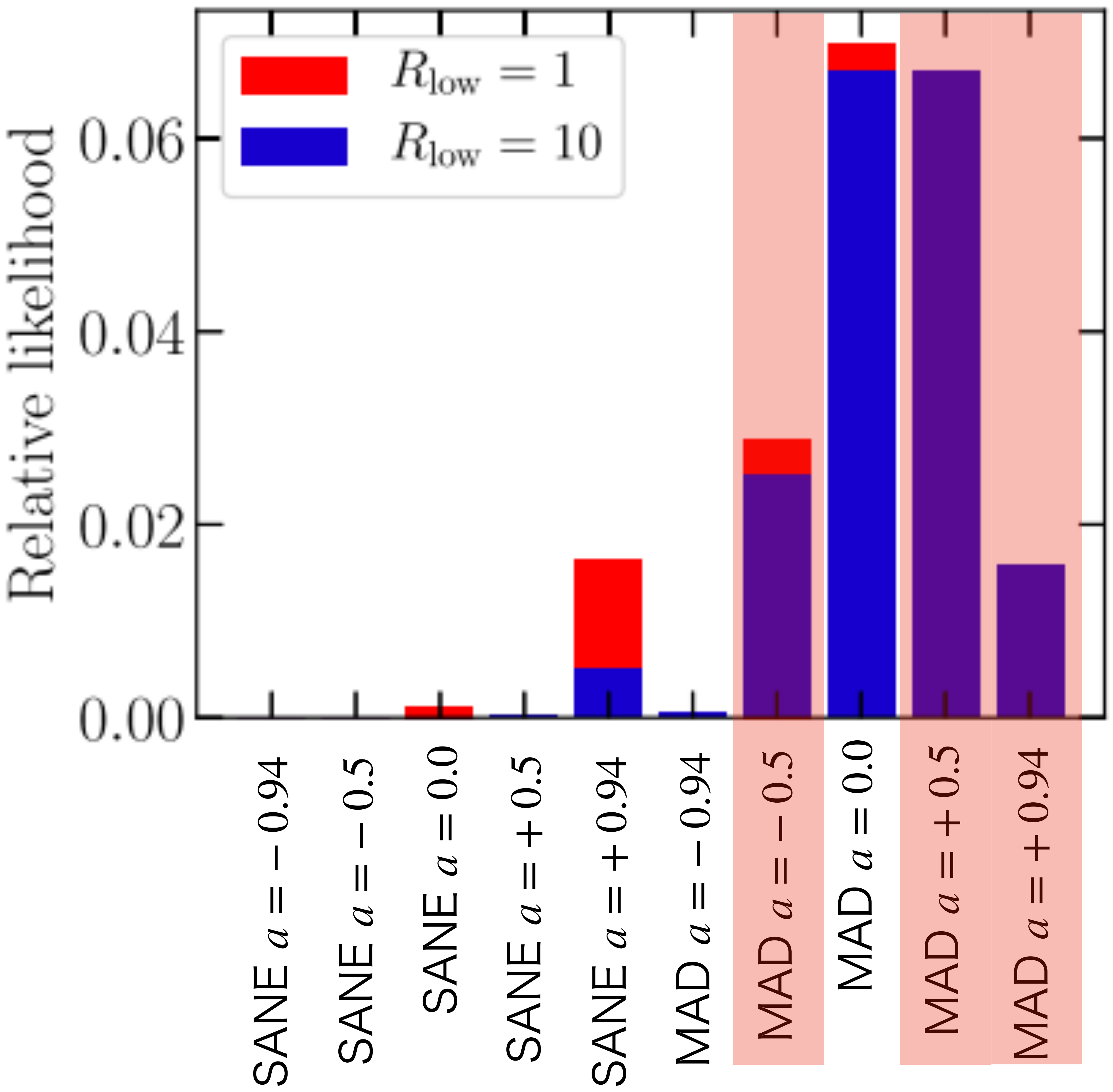
Current Understanding of Black Hole Spin

- The BH spin is still unclear from the event horizon scale observation (EHT)



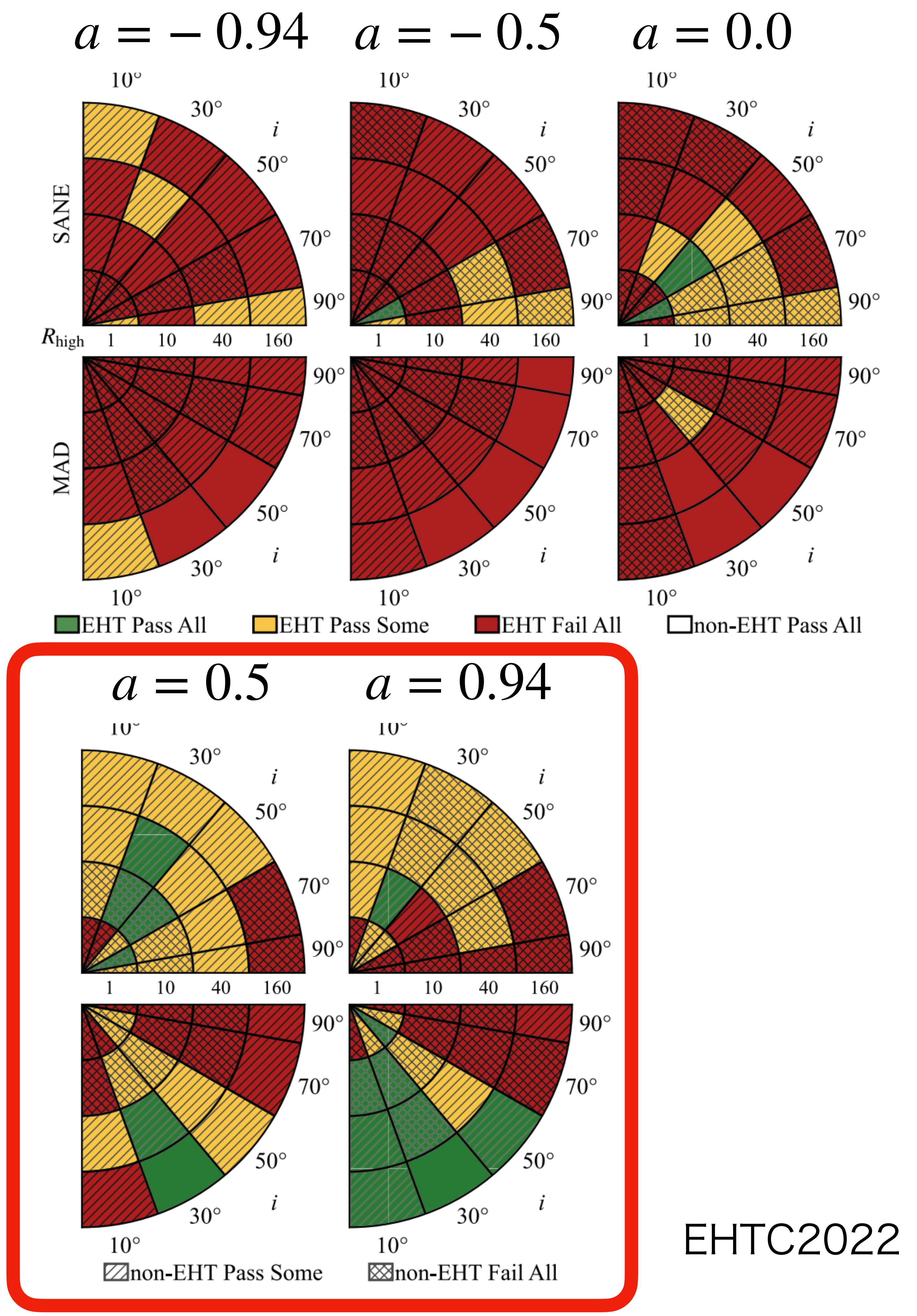


M87*



EHTC2021

SgrA*



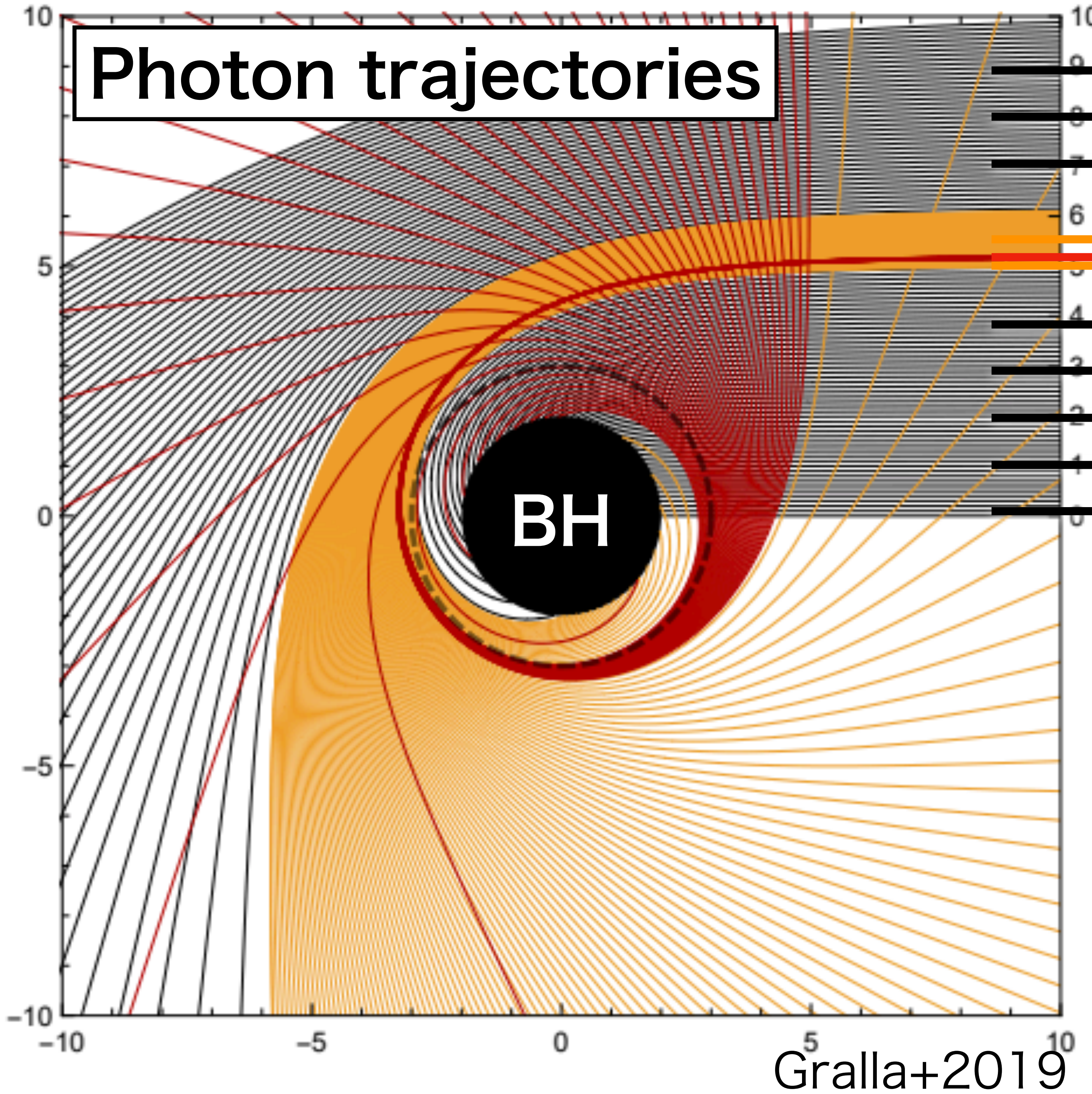
EHTC2022

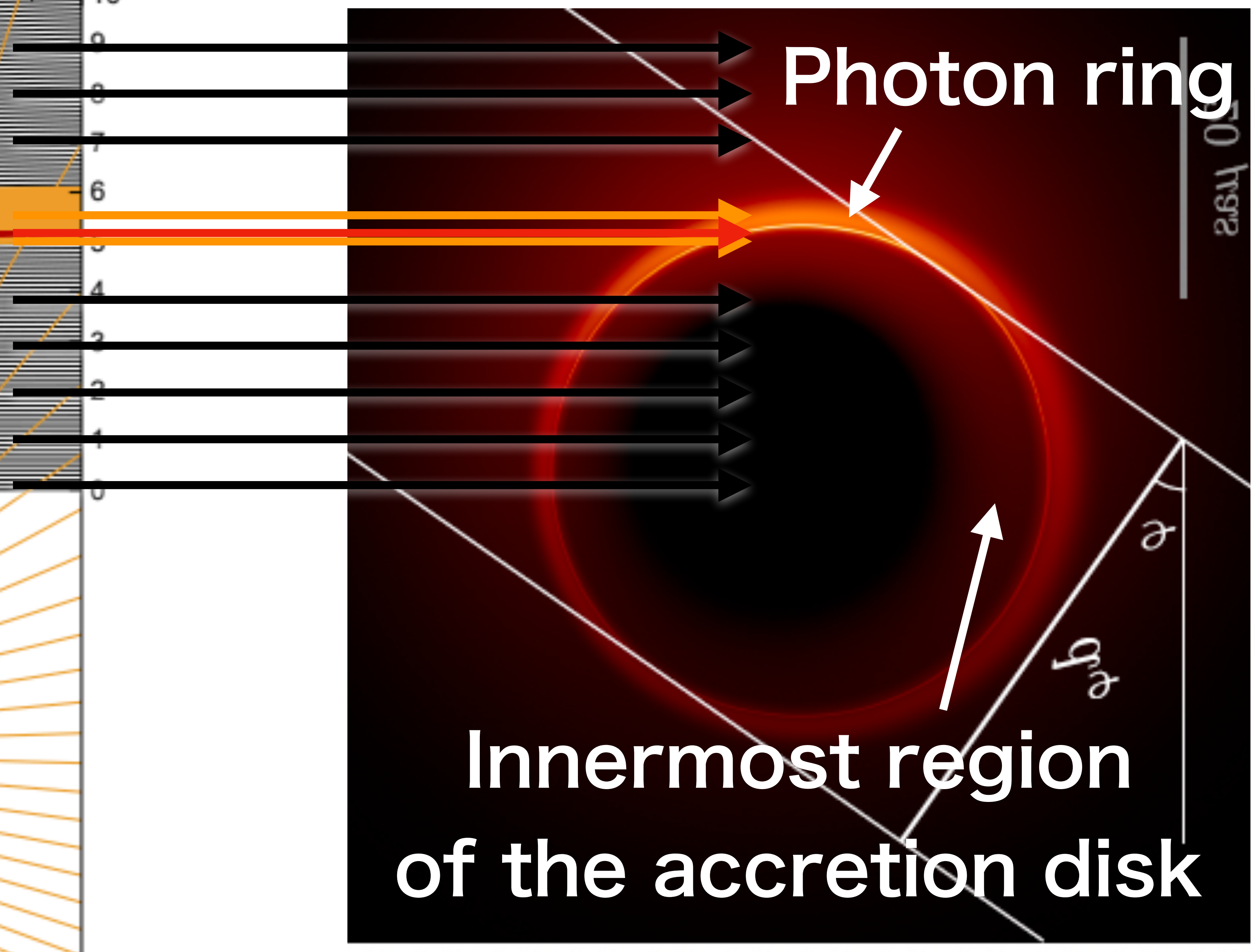
Important Observable Feature: Photon ring

- **Photon ring is thin ring-like image caused by the light bending**
  - photons passing near the BH reach observer after orbinting around the BH (→ path length from the emitter to the observer becomes large)
  - intensity is proportional to the path length in the optical thin medium

➡ thin ring-like image is created

- The shape of the photon ring is **mainly determined by the BH spin and inclination angle**
- It is expected to observe by the **Black Hole Explorer (BHEX)**

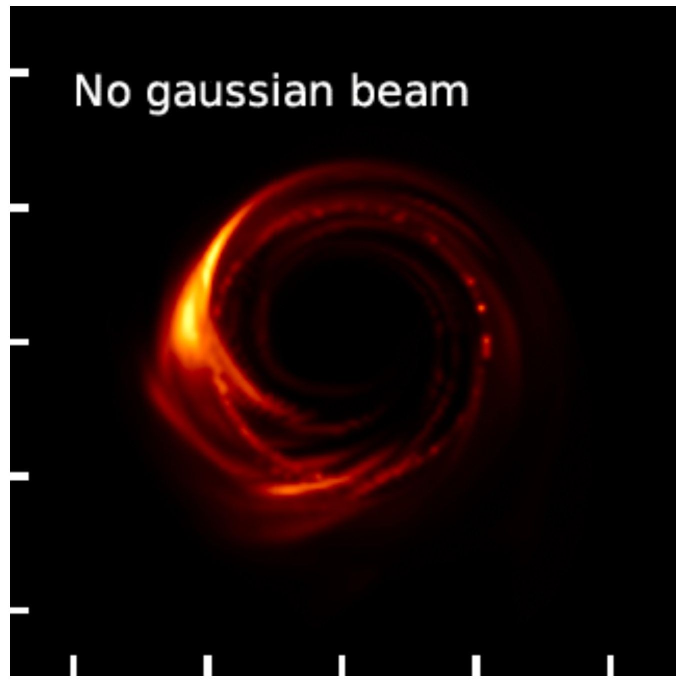




One method to determine the BH spin is machine learning

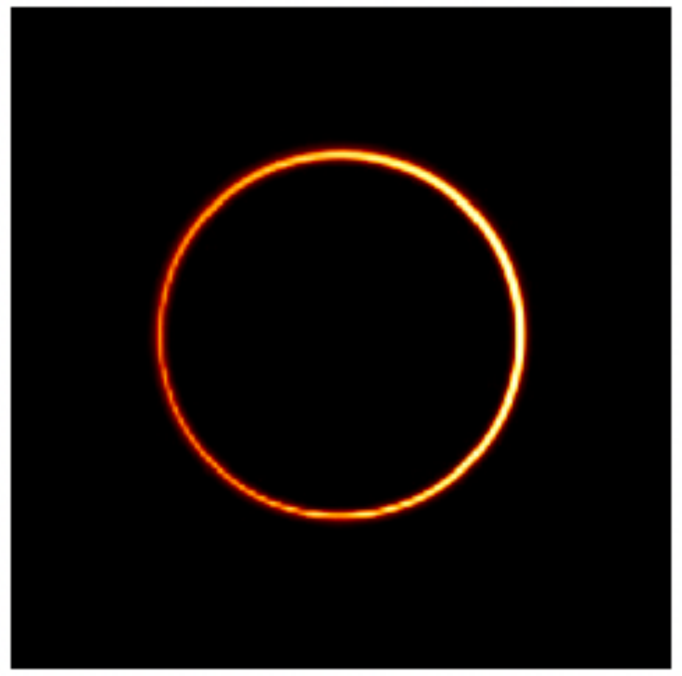
- **Based on the simulated images** (ven der Gucht+2020)
  - Convolutional Neural Network (CNN) is directly trained by simulated images
  - Due to large computational cost of simulations, few patterns of the BH spin and inclination angle is included

example training data



- **Based on the geometric features of the photon ring** (Farah+2024)
  - Gradient Boosting Regression Tree (GBRT) is trained by feature values defined manually
  - the performance may depend on how to choose the feature values that are needed to defined manually
  - It is unclear whether it can be applied to real observation because the model is too simple

example training data



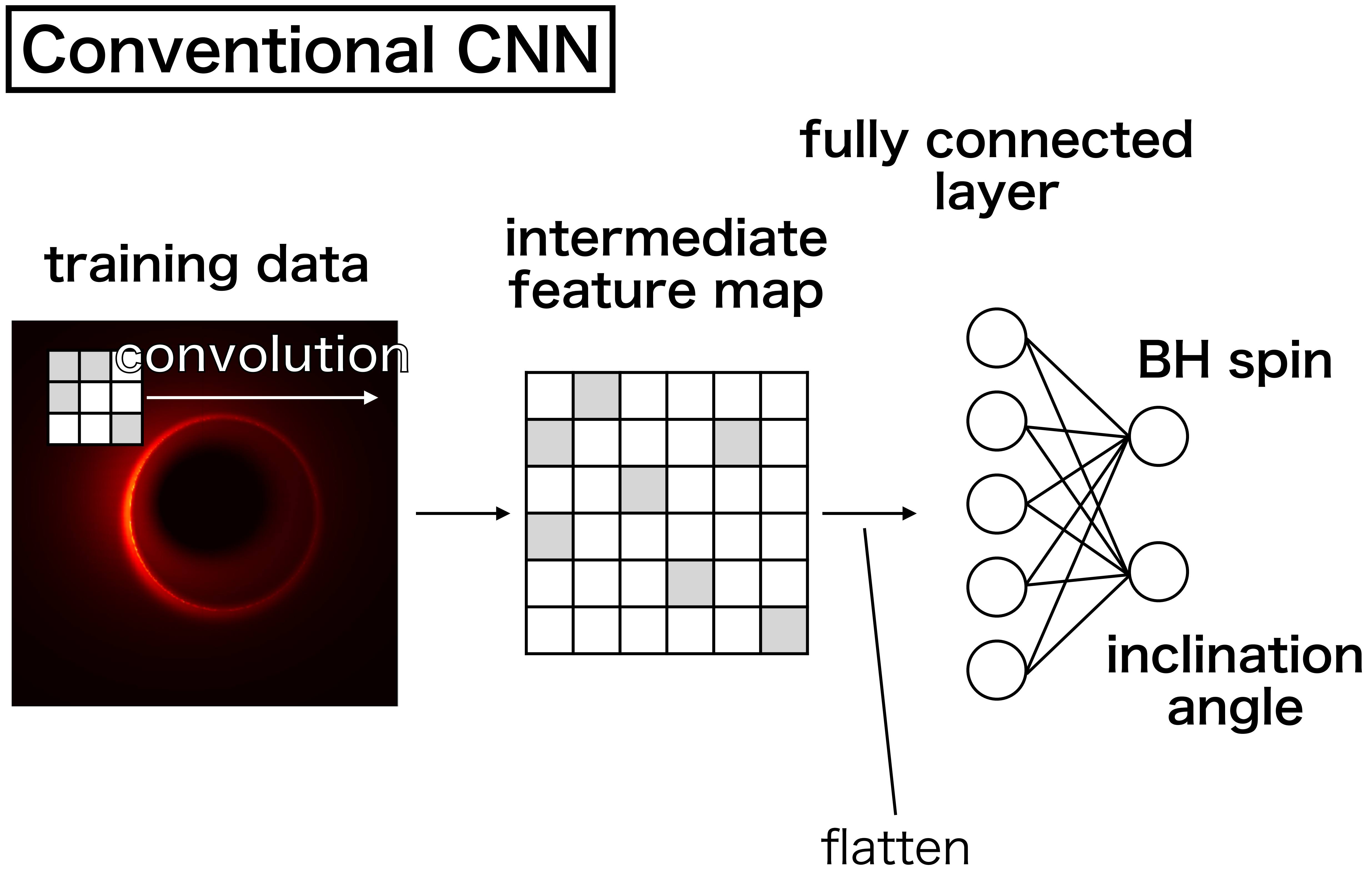
A general-purpose machine learning model that learns the realistic images has not yet been established

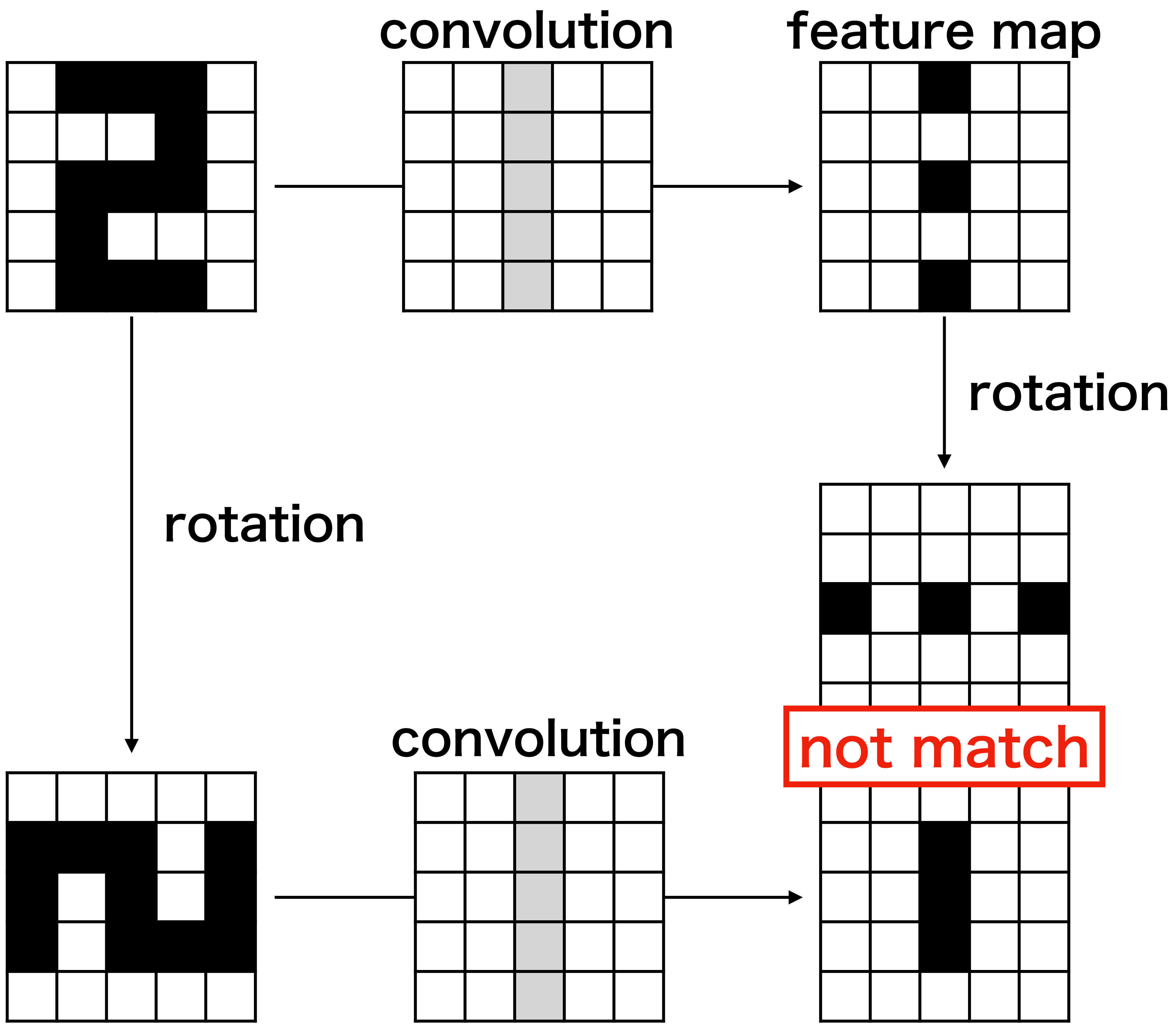
Purpose of This Work

equivariance: 同変性, 等変性, 동변성

- **Constructing the general-poupose machine learning model to estimate the BH spin and inclination angle for the future observations**

➡ We first apply the **E(2)-Equivariant Steerable CNN** to the BH images





Purpose of This Work

equivalance: 同変性, 等变性, 동변성

- **Constructing the general-poupose machine learning model to estimate the BH spin and inclination angle for the future observations**

➡ We first apply the **E(2)-Equivariant Steerable CNN** to the BH images

Method : Generating $\sim 10^4$ Training Data

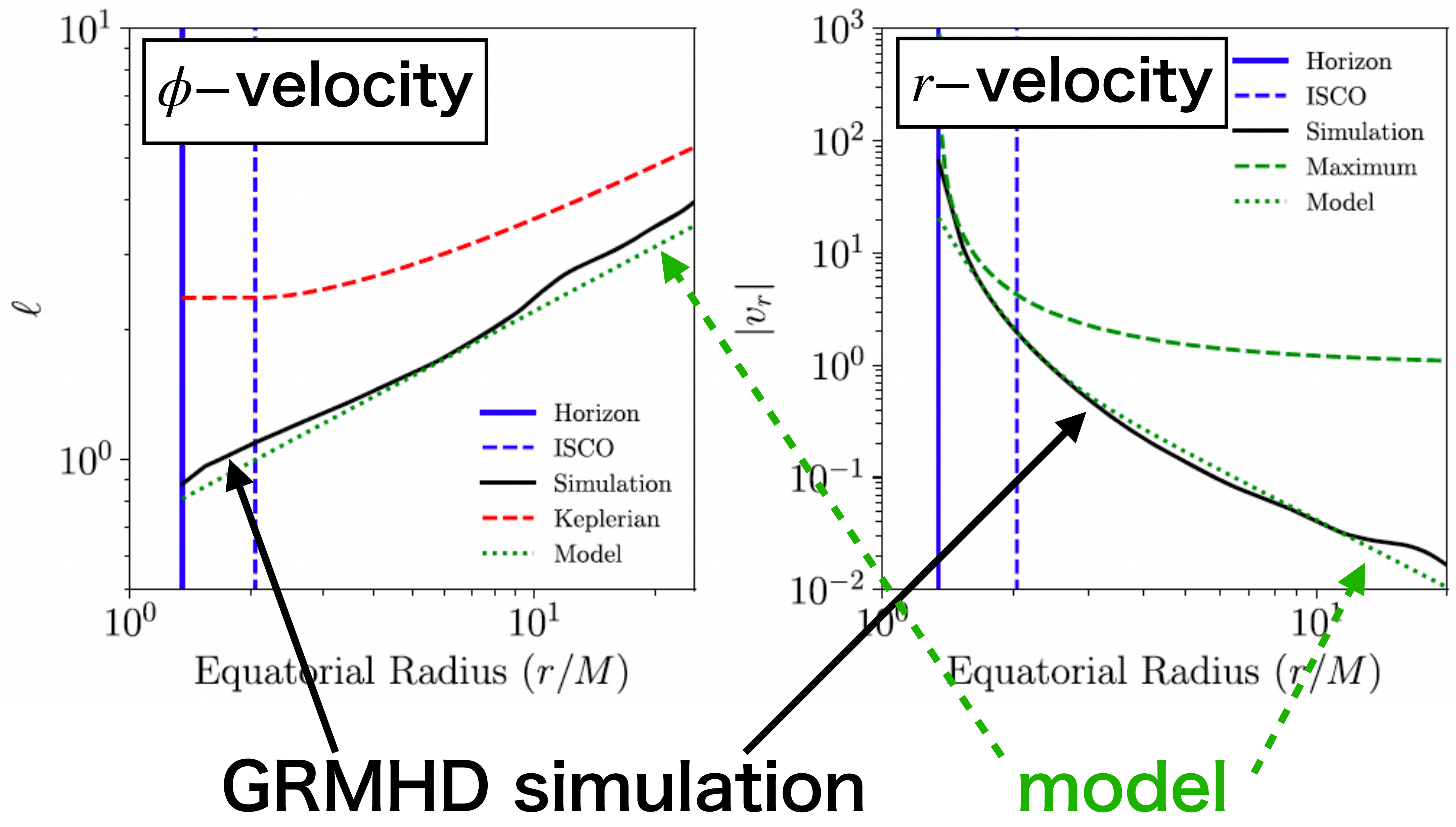
- **Equatorial emission model** (Chael+2021)
  - fluid velocity, emissivity (limited to $\theta = \pi/2$)

emissivity ($\theta = \pi/2$) $\log[j(r)] = p_1 \log[r/r_+] + p_2(\log[r/r_+])^2$

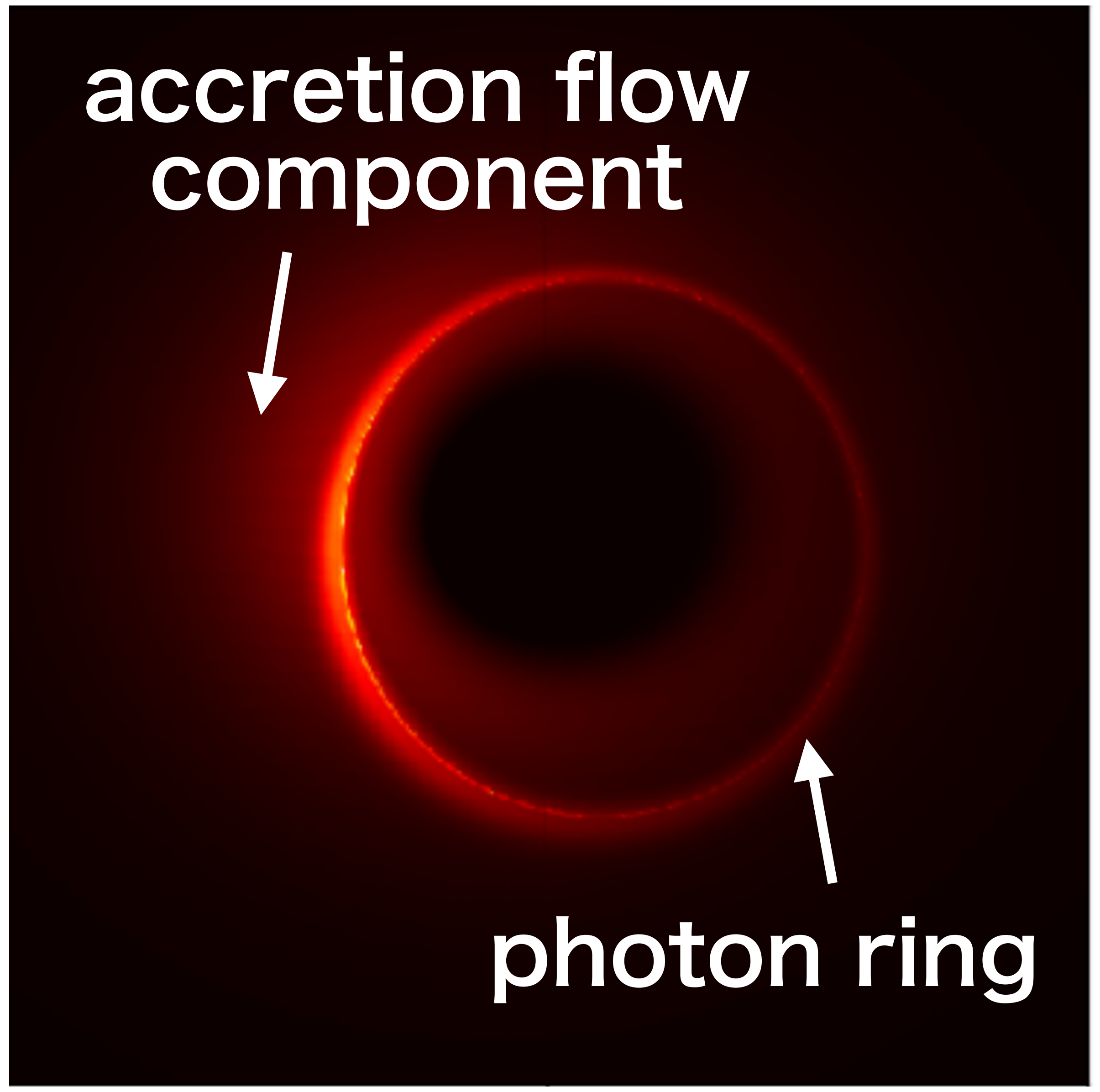
emissivity at (α, β) on screen

$$I(\alpha, \beta) = \sum_{n=0}^{N_{\max}-1} f_n j(r_n) g^3(r_n, \alpha, \beta)$$

n: number of times that photons cross the equatorial plane,
g: redshift factor,
r_+: event horizon radius, f_n, p_1, p_2: free parameters



- g and n are calculated by ray-tracing general relativistic radiative transfer code: CARTOON (Takahashi+2022)
- **BH spin:** $0 \leq a \leq 0.998$, **inclination angle:** $0^\circ \leq i \leq 89^\circ$



Method : Network Architecture

- Regression model based on Cesa+2022 (https://github.com/QUVA-Lab/escnn)

Input	512x512 image
Block 1	Convolution (4x4) → ReLU
Pooling 1	Average pooling (5x5)
Block 2	Convolution (4x4) → ReLU
Pooling 2	Average pooling(5x5)
Group pooling	$N_\theta = 8$ **(=every 45°) feature maps are summerized**
Fully connected	Fully connected layer (50 nodes) → ReLU → Fully connected layer (50 nodes)
Output	BH spin and inclination angle

$N_\theta = 1$ **means conventional CNN**

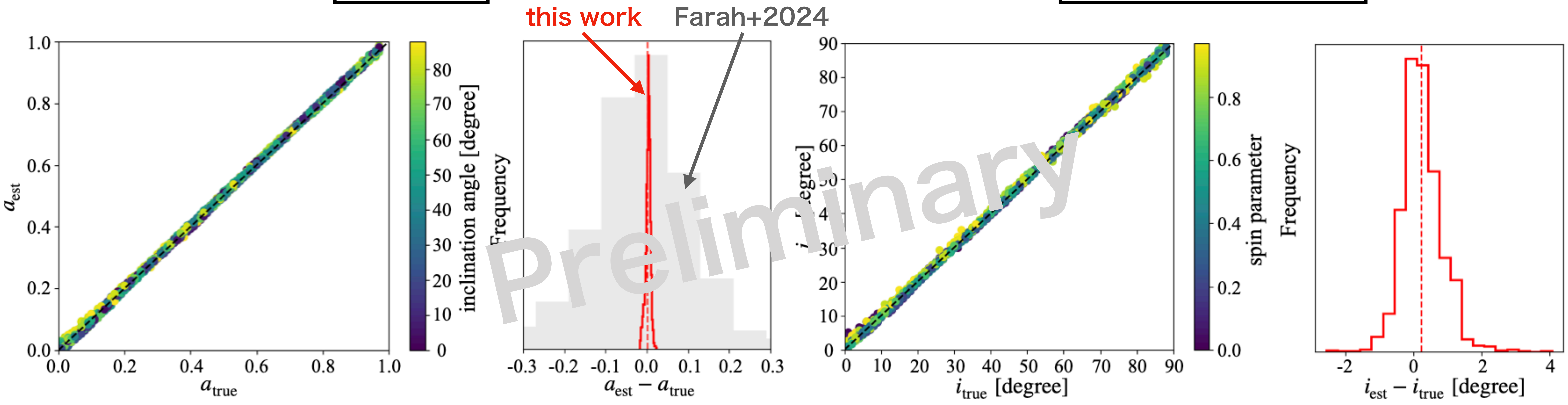
optimizer: adam, loss: mean squared error,
epochs: 200, batch size: 50

Result for Validation Data

- Absolute error: less than ~0.04 (BH spin), less than ~4° (inclination angle)
- Our absolute errors are 1/10 to 1/5 of those of previous work (Farah+2024) even though the number of the training data is 1/10 to 1/100

BH spin

inclination angle

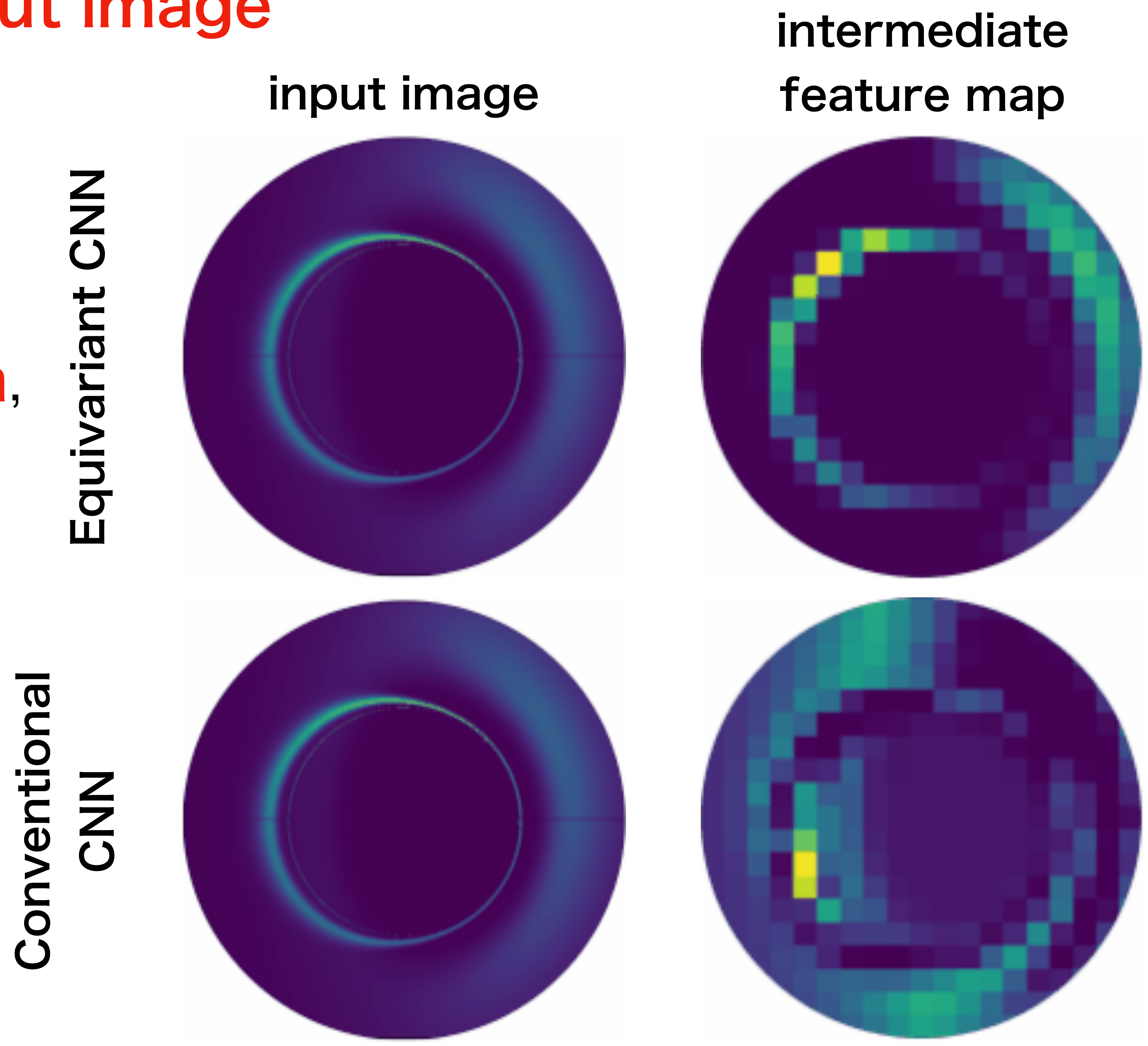


Comparison of E(2)-Equivariant Steerable CNN and Conventional CNN

- **The structure of the intermediate feature map of equivariant CNN does not change with the rotation of the input image**

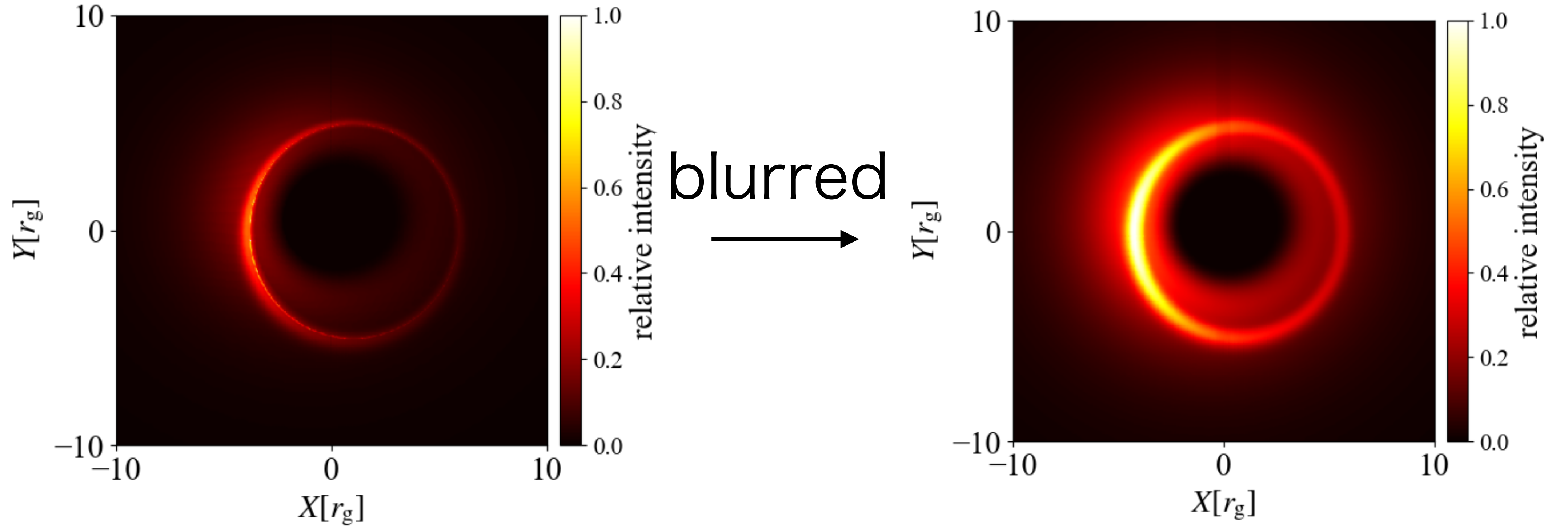
- From intermediate feature maps, it can be seen **the shape of the photon ring is important for estimating the BH spin**, and **the accretion flow component is important for estimating the inclination angle**

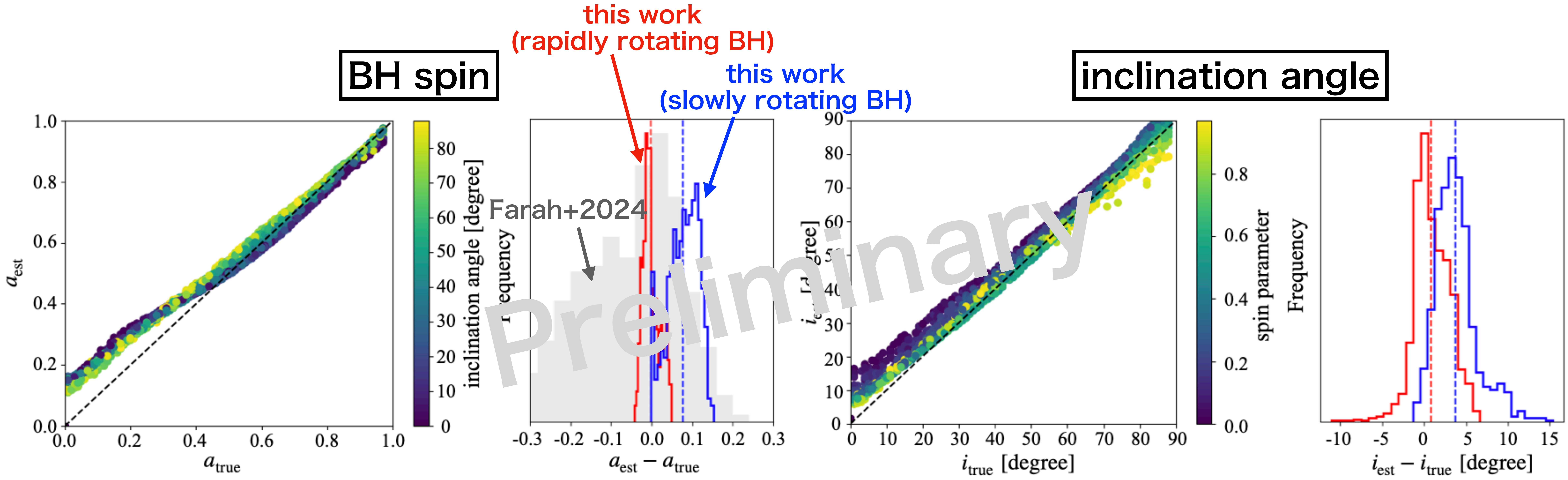
- The E(2)-Equivariant Steerable CNN is quantitatively better than the conventional CNN



Application to Blurred Images

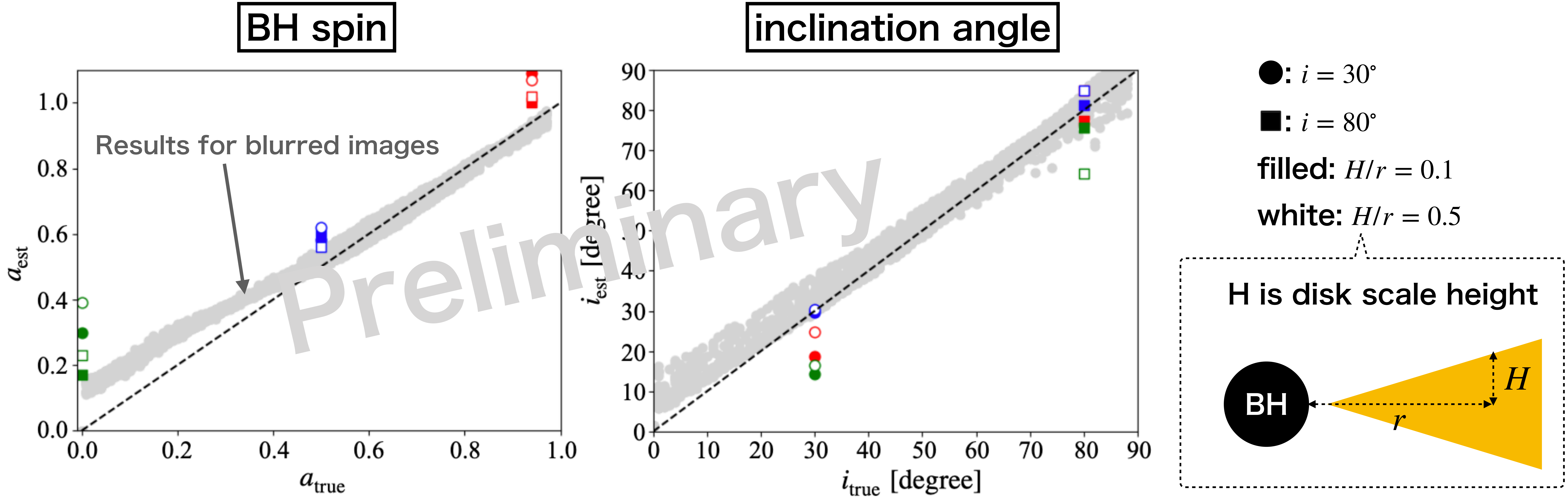
- Blurred images with 6 uas Gaussian beam (not included in the training data)
- **Our absolute errors are ~1/5 of those of previous work (Farah+2024)** even when the input images are blurred





Application to Unknown Images

- Blurred images of the accretion flow **with finite scale height**
  - train data is generated by the model whose emissivity is limited to the equatorial plane
- **Our network also shows relatively good performance for unknown images**
  - The assumption of ignoring the disk scale height does not significantly affect the results



Summary

- We **first apply the CNN with the equivariance under the rotation** of inputed BH images.

- Our network performance is **greatly better than the previous study even with lower number of the dataset**.

- With the equivariant CNN, the orientation angle of the input images is arbitrary.

- **For the blurred images (especially for the images of the ralpidly rotating BH)**, our network performance is also better than the previous study.

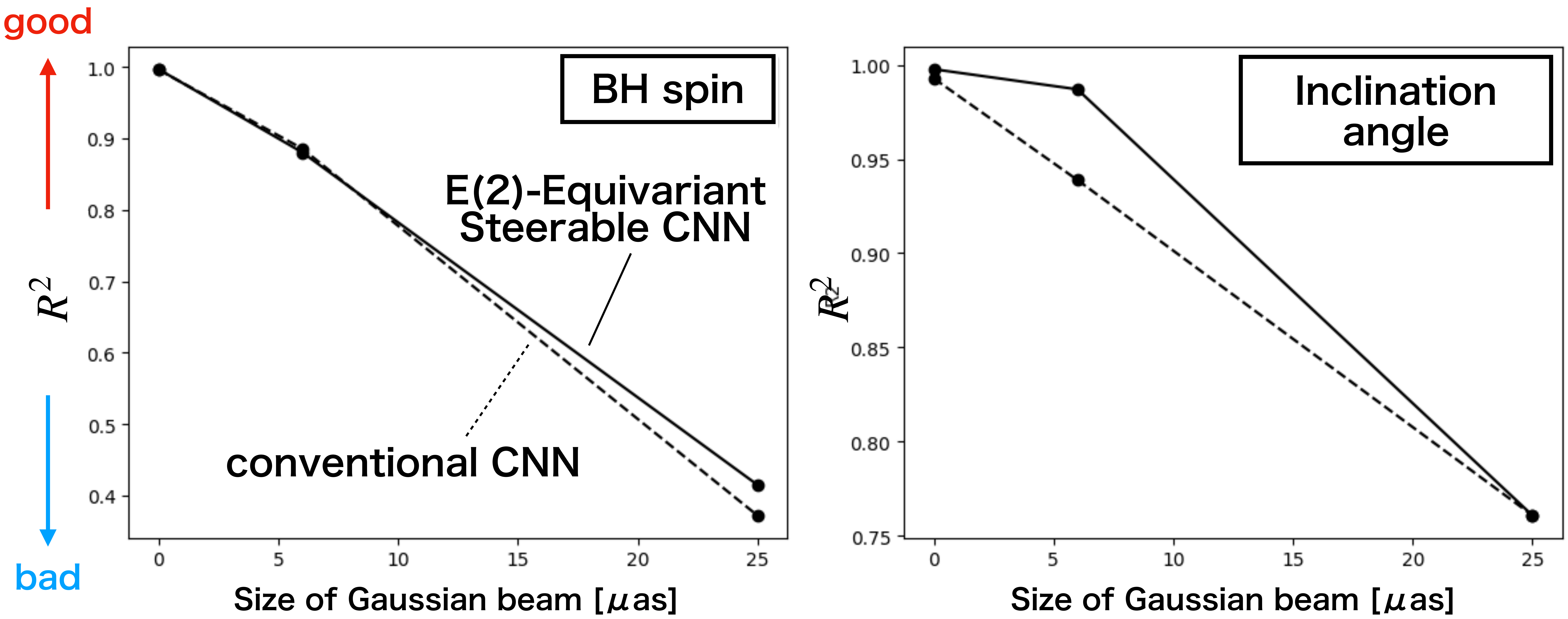
- Our network can predict BH spin and inclination angle **not only for the validation data, but also for unknown images**.

Appendix: Saliency map

- Saliency map shows the degree of importance of each a pixel
- It seems that **the shape of the photon ring is important for estimating the BH spin**, and **the accretion flow component is important for estimating the inclination angle**

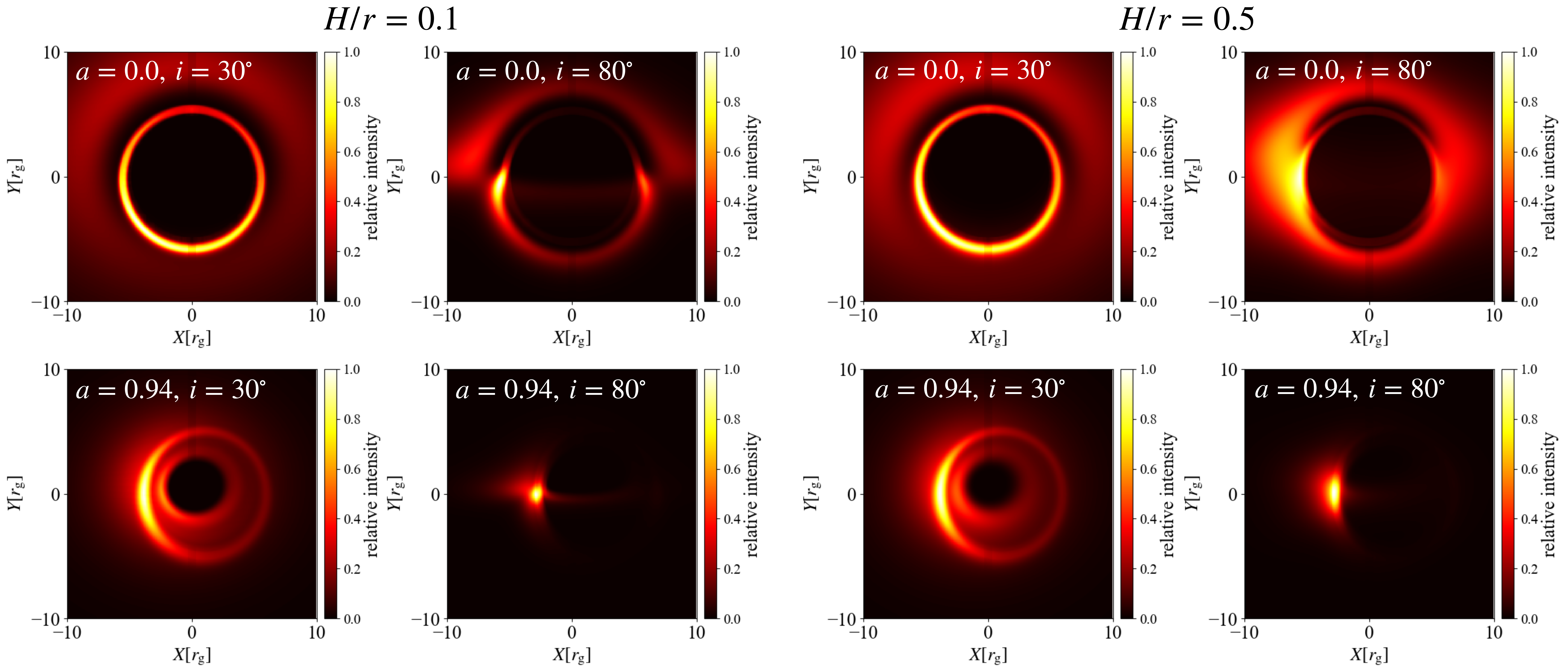
Appendix: Quantitative Comparison

- decision coefficient: $R^2 = 1 - (\sum_i y_{i,\mathrm{true}} - y_{i,\mathrm{est}})/(\sum_i y_{i,\mathrm{est}} - \bar{y})$
- The E(2)-Equivariant Steerable CNN is **quantitatively better than the conventional CNN**



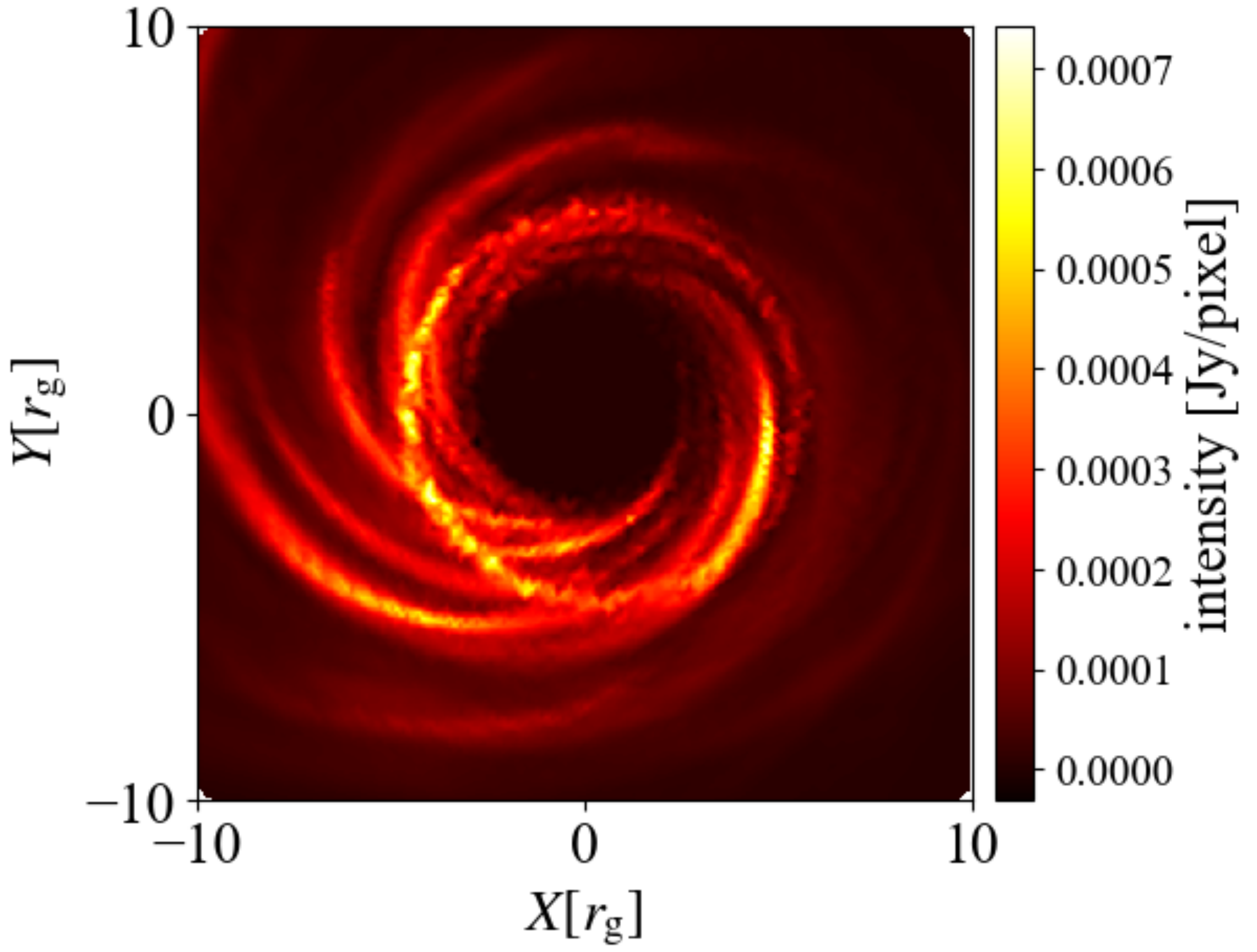
Appendix: Example images of accretion disk with a finite scale height

- A effect of disk scale height significantly affects the images seen from the edge-on view.



Appendix: Application to GRMHD snapshot

- We apply our network to just two GRMHD snapshot.
- **Our network also shows relatively good performance for realistic images.**

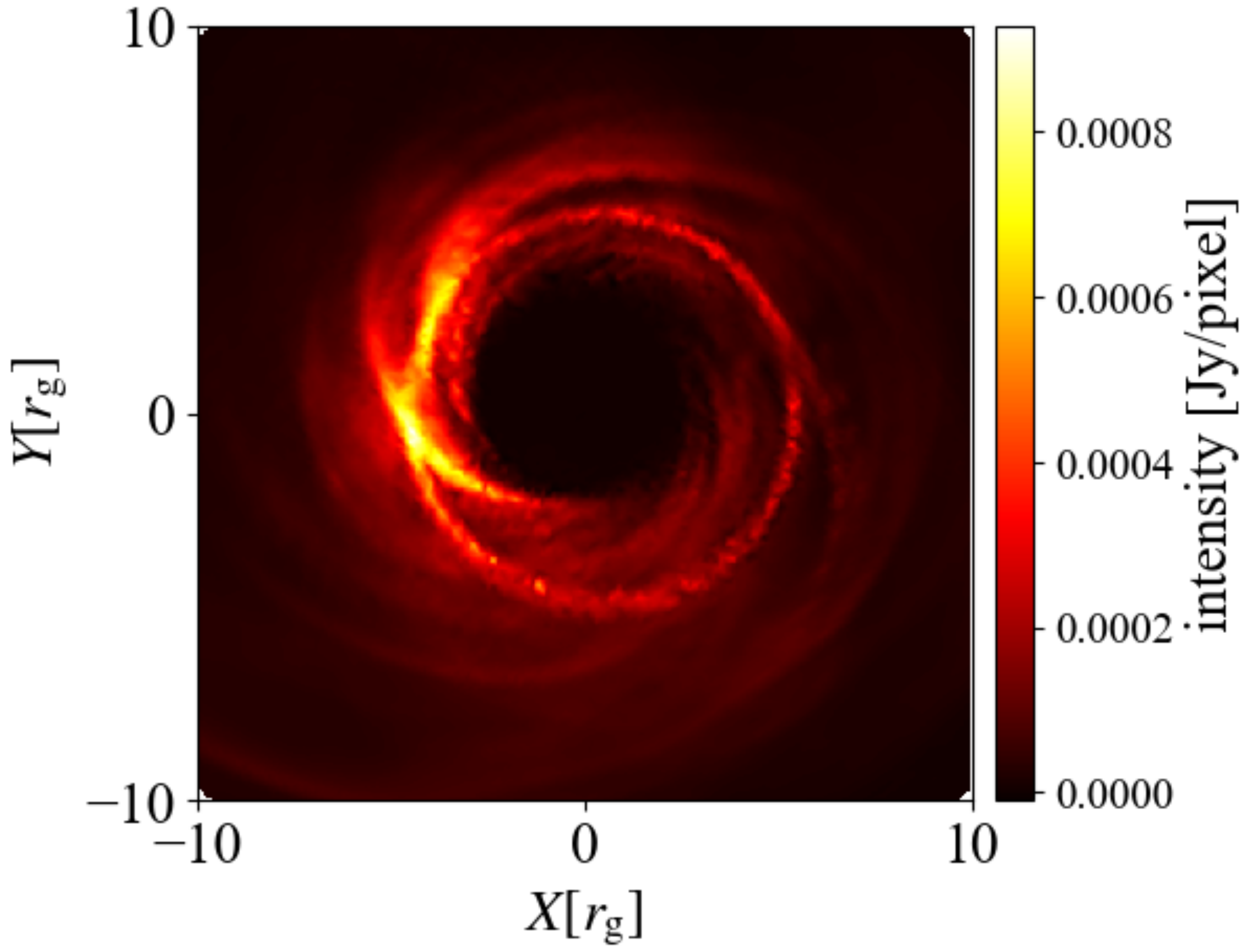


$a_{true} = 0.9375$

$i_{true} = 17^\circ\ (= 0.29\pi)$

$\longrightarrow$ $a_{est} = 0.804$

$i_{est} = 38^\circ\ (= 0.68\pi)$



$a_{true} = 0.9375$

$i_{true} = 17^\circ\ (= 0.29\pi)$

$\longrightarrow$ $a_{est} = 0.840$

$i_{est} = 31^\circ\ (= 0.55\pi)$

BH spin