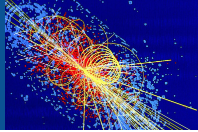
Deep Learning Based Tagger for Highly Collimated Photons at CMS

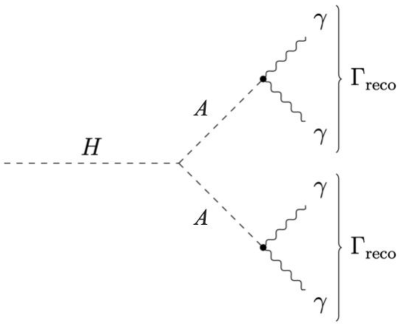
Kyungmin Park, Manfred Paulini
On Behalf of the CMS Collaboration

May 14, 2024 @ Pheno/DPF 2024

Introduction



- **Motivation**
  - Search for exotic Higgs boson (H) decaying into two pseudoscalar A, much lighter than H.
    - Each boosted A decays into **highly collimated two photons**.
    - Angular separation between the collimated photons is too small.
      → reconstructed as one artificially "*merged photon*".
  - Major background of the analysis: QCD jets with photons.
    - Neutral mesons, i.e. pion, in jets can decay into photons, producing additional photons, or hadronic "*fakes*".



Mass for A [0.1, 1] GeV

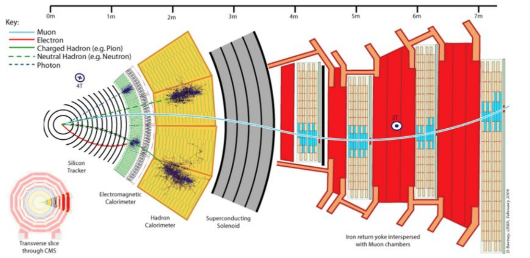
- Analysis from CMS RunII **[1]** used the standard CMS photon identification algorithm, based on Boosted Decision Tree trained to classify between photon and fakes.
  - Standard photon identification is not optimal for merged photons.

➢ *Develop a dedicated tagger to optimally identify the merged photon signature using **deep learning**.*

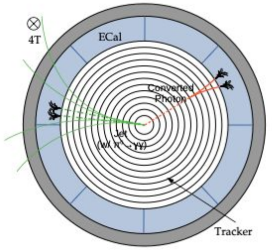
[1] 10.1103/PhysRevLett.131.101801

Deep Learning with Detector Images

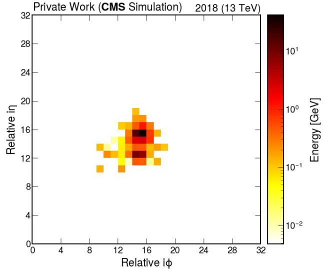
- **Photons in CMS detector**
  - Interact with the detector material and "shower", depositing their energy over the range of crystals in the **electromagnetic calorimeter (ECAL)**.
  - Converted photons can leave their tracks in the **tracker**.



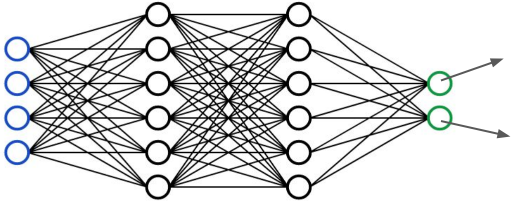
- **Identifying highly collimated photons**
  - Build the **tagger** for the highly collimated photons from A decay using **deep learning.**
  - Input: images of signal and background photons' trace in ECAL and tracker.



Detector images



Neural Network



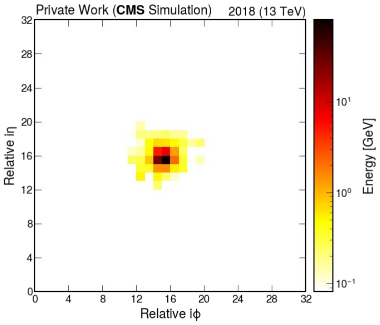
Output

Signal photon from A decay

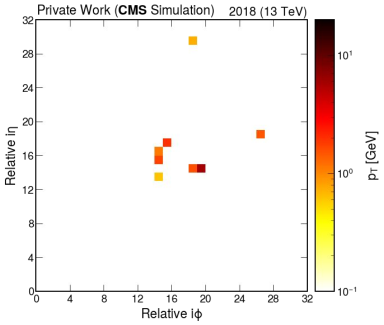
Background from QCD jets

Input for the Tagger

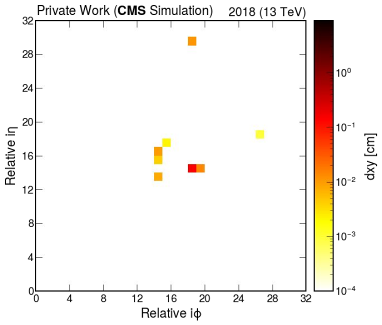
- **<u>Input images</u>** for the tagger



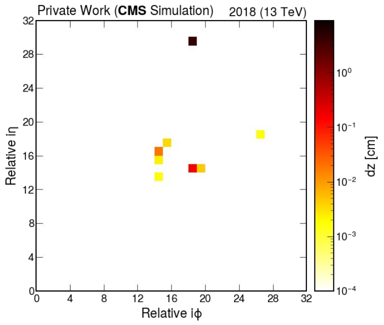
[1] *Signal* (1 GeV):
ECAL shower shape



[2-1] *Background*:
Tracks p_T



[2-2] *Background*:
Tracks *dxy*



[2-3] *Background*:
Tracks *dz*

[1] ECAL shower shape

- Make an image of 32x32 in ECAL crystal grid of azimuthal angle ϕ and pseudorapidity η $(i\phi, i\eta)$, centered around shower seed.

[2] Track "structure"

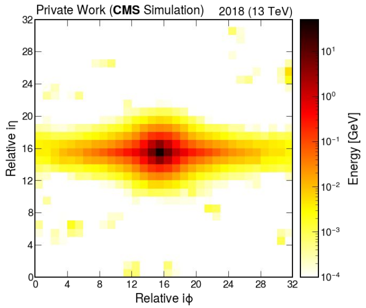
- For the associated tracks, get their transverse momentum (p_T) and impact parameters (dxy, dz).
- Each track is projected onto the 32x32 $(i\phi, i\eta)$.

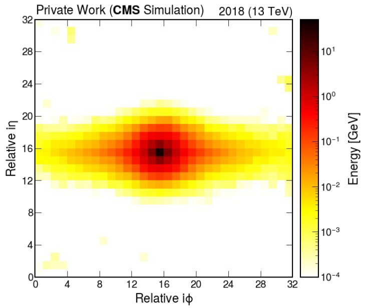
Input for the Tagger

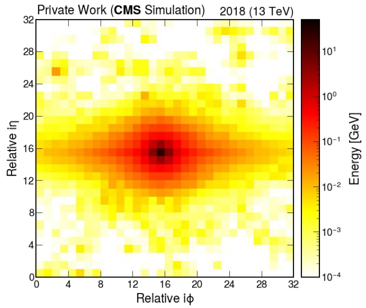
- <u>Inputs averaged over 10k images</u>
  - **Background** hadronic fakes have ECAL shower and tracks more spread out.
  - **Signal** merged photons have narrower shower shapes as A mass gets lighter.

	<u>0.1 GeV A</u>	<u>1 GeV A</u>	<u>QCD</u>

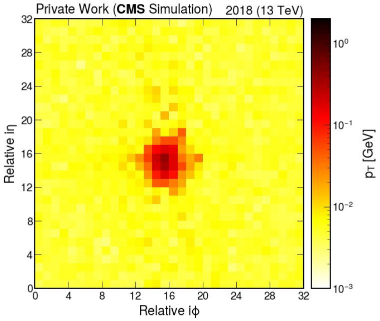
<u>ECAL Shower</u>

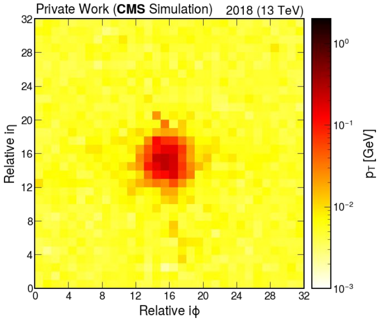


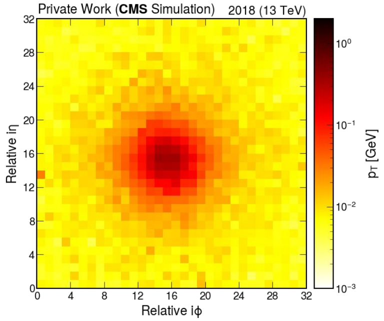




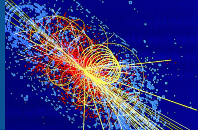
<u>Tracks p_T</u>







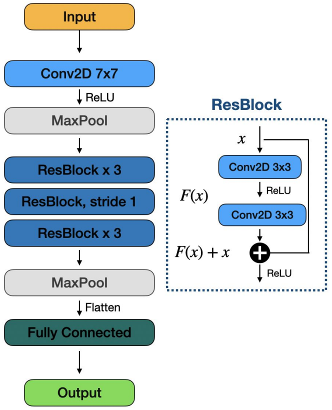
Data Preprocessing and Training



- **Inputs**
  - Dataset: Monte-Carlo simulation for the signal H → AA → two merged photon events and background QCD events.
  - Event selection: require events to pass a trigger with two photon requirements.
  - Multi-layer inputs for CNN: ECAL shower shape and track structure images.

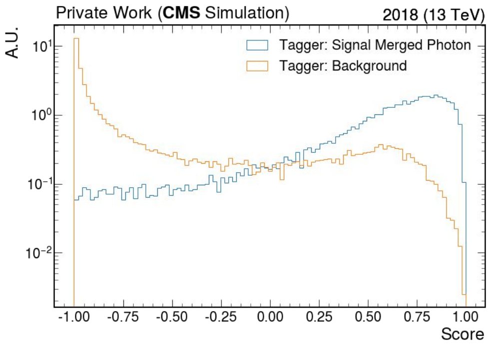
- **Preprocessing**
  - Apply reweighting factor to flatten out the different distributions of p_T and η in signal and background. → Avoid bias in the tagger due to kinematics in (p_T, η).

- **Training**
  - Model: **Convolutional Neural Network (CNN) with ResNet** architecture.
  - Loss: Mean Squared Error
  - 200k images split 8:2 for training and validation.
  - Mass points for training: (0.1, 0.4, 0.6, 1) GeV

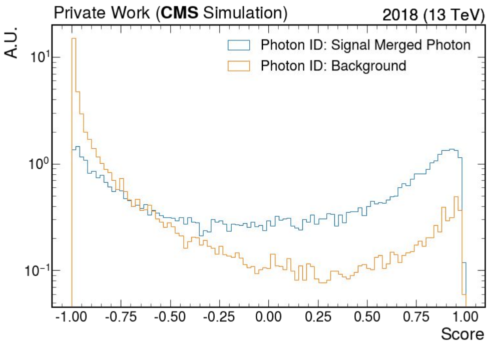


Model Validation and Working Point

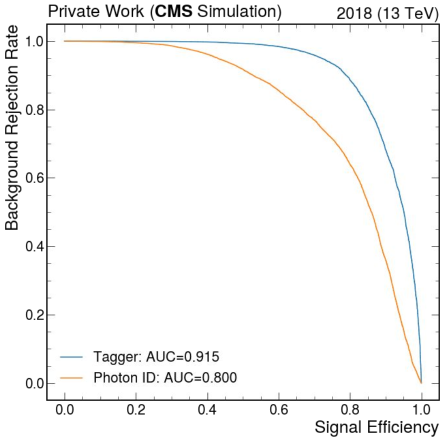
- **Model validation**



Tagger output score



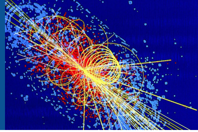
Photon ID output score



- **Working Points (WP)** based on **signal efficiency** ϵ
  - Choose thresholds for each WP.
  - "Loose": ϵ = 0.9, "Medium": ϵ = 0.8, "Tight": ϵ = 0.7
  - **Signal-to-background** ratio in each WP:

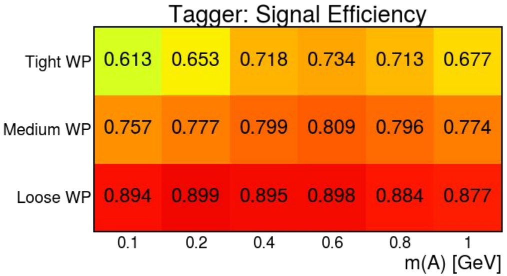
WP	Loose	Medium	Tight
Tagger	4.379	5.420	6.659
Photon ID	1.933	2.441	3.036

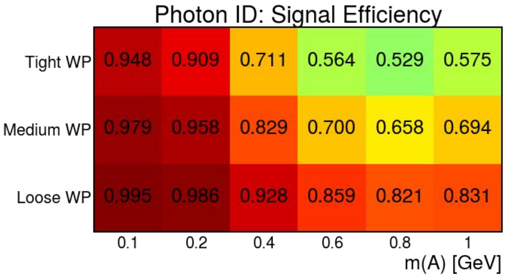
Deployment of the Tagger



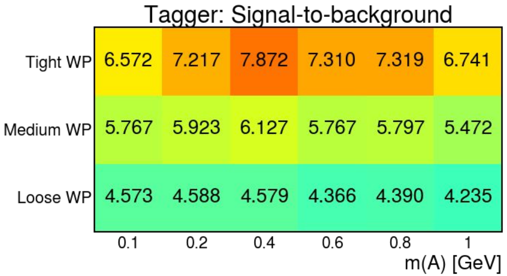
- Testing with each A mass dataset, the tagger
    - Shows flat signal efficiency and signal-to-background ratio across different masses of A.
    - Interpolates well the mass points not used in training (0.2, 0.8) GeV.

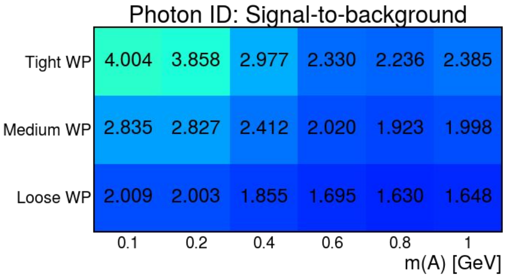
Signal efficiency





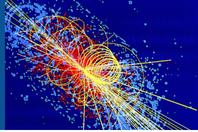
Signal-to-background ratio





- Compared with the standard photon ID performance, one can expect to gain better signal sensitivity for the analysis with the tagger.

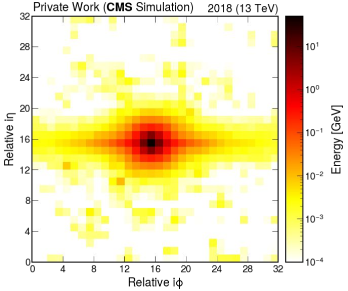
Including Single Photon Background



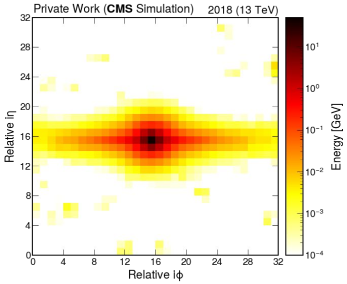
- Target another background of the analysis: **single photon**, *i.e.* H $\rightarrow \gamma\gamma$.

- **Signal merged photon vs. single photon background** classifier
  - Using the ResNet architecture and inputs from ECAL and tracker.

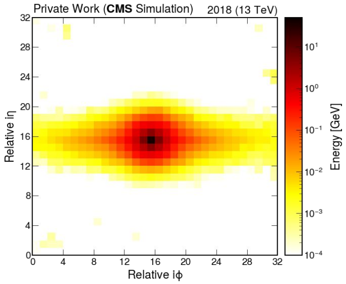
Single photon

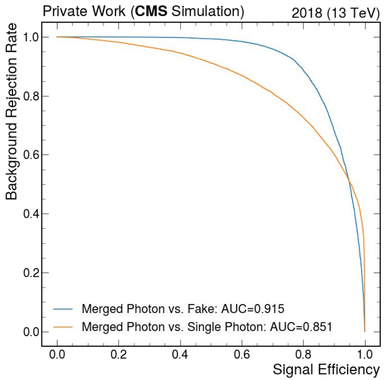


0.1 GeV A

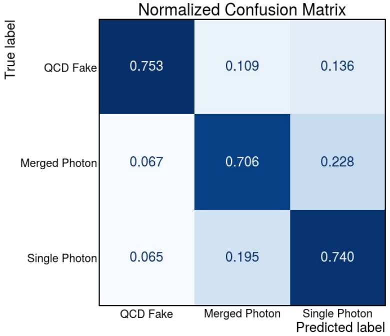


1 GeV A

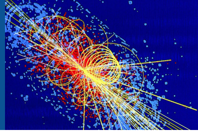




- Expanding to **multi-class tagger**
  - Class: [1] Signal merged photon
    [2] Background hadronic fakes
    [3] Background single photon
  - Loss: CrossEntropy; use *max. probability* for class assignment



Summary and Outlook

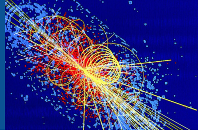


- **Summary**
  - Built a deep learning based tagger to optimally identify highly collimated photons from boosted decay against the QCD background, in search for exotic Higgs decaying to boosted pseudoscalar A’s.
  - Utilized low-level electromagnetic shower shapes and track structures as inputs to the tagger.
  - Obtained good signal-to-background ratio across different masses of A that outperforms the standard CMS photon identification algorithm, allowing an improvement in signal sensitivity for the search.

- **Outlook**
  - Promising results for the multi-class classifier, including the single photon as another background class.

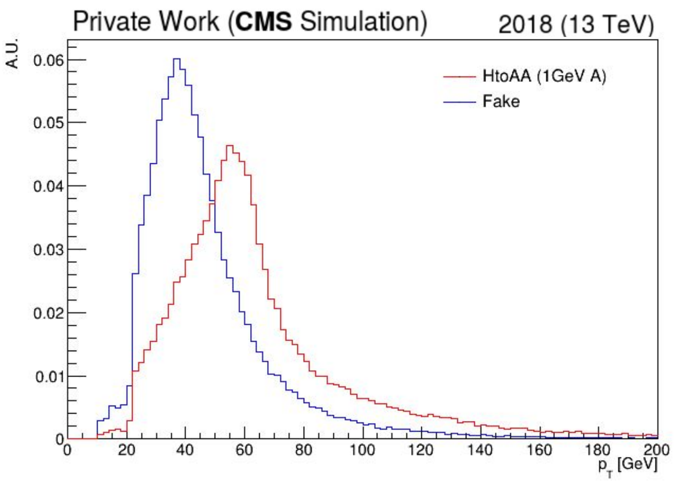
BACK UP

Preprocessing: Object Reweighting

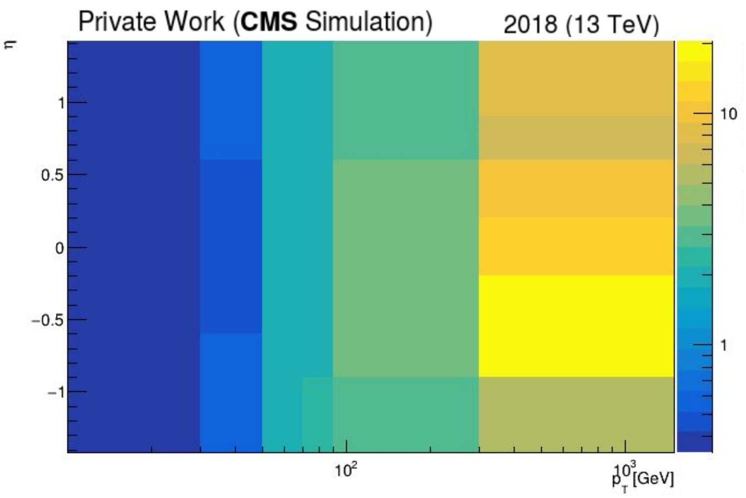


- **<u>Object-level reweighting</u>** based on pT and η
  - pT and η distributions are not identical for signal photons and fakes [1].
  - Reweighting factors taken from ratio of signal and background histograms in each bin [2].

[1] pT distribution

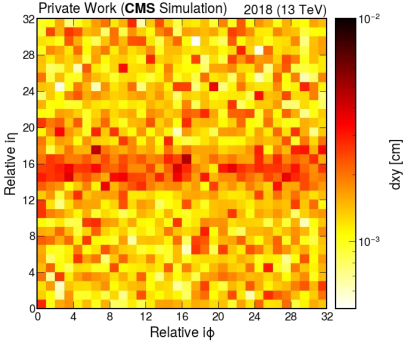
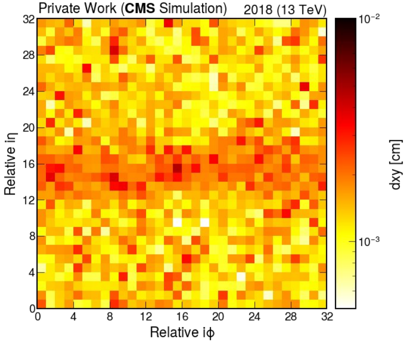
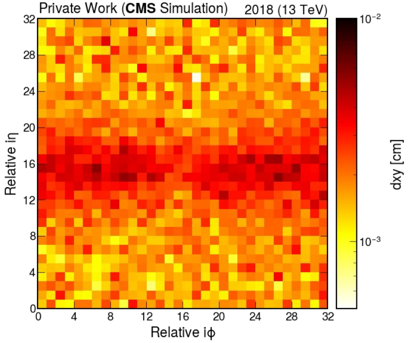
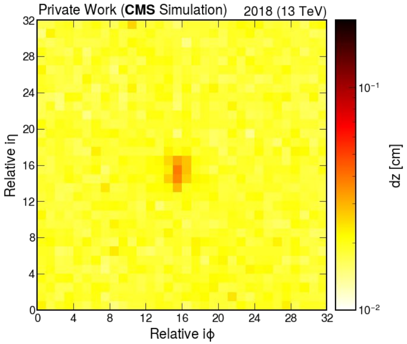
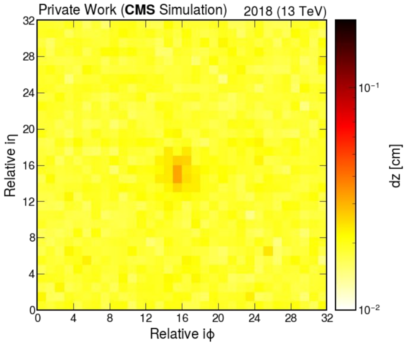
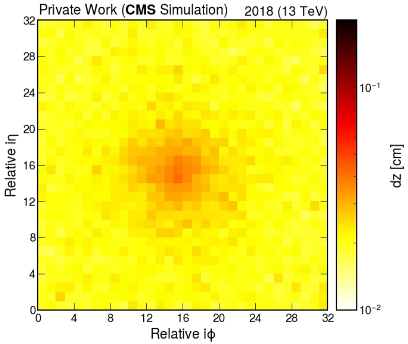


[2] Reweighting factors

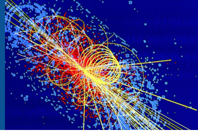


Inputs for Tracks

- Input averaged over 100k images

	0.1 GeV A	1 GeV A	QCD
Tracks dxy			
Tracks dz			

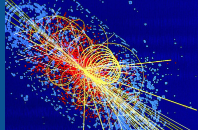
Different Sets of Input Layers



- Determine the optimal sets of input layers
  - Compare AUC scores for models trained for each A mass dataset

Layers \ A mass	**0.1 GeV**	**0.2 GeV**	**0.4 GeV**	**0.6 GeV**	**0.8 GeV**	**1 GeV**
Shower	0.844	0.839	0.849	0.835	0.826	0.832
Track	0.880	0.881	0.879	0.879	0.878	0.877
Shower + Track	0.918	0.916	0.929	0.922	0.913	0.910

➔ Use **shower and track** information.



Deployment of the Tagger: Efficiency

- <u>Tagger and photon ID</u>