Searching for long-lived particles at the LHC: Fourth workshop of the LHC LLP Community

23 Oct 2018, 08:00 → 25 Oct 2018, 21:00 Europe/Zurich

MACHINE LEARNING TECHNIQUES FOR DARK PHOTONS AT ATLAS

Iacopo Longarini
Cristiano Sebastiani

Stefano Giagu

Luigi Sabetta
Simone Francescato

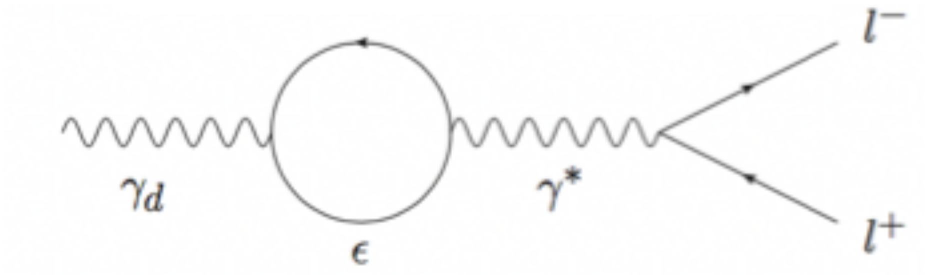
INFN

ATLAS ROMA1

SAPIENZA Università di Roma

dark sector models @ATLAS

Dark-QED U(1): $\mathscr{L} \propto \varepsilon e \gamma_d^{\mu} J_{\mu}^{em}$

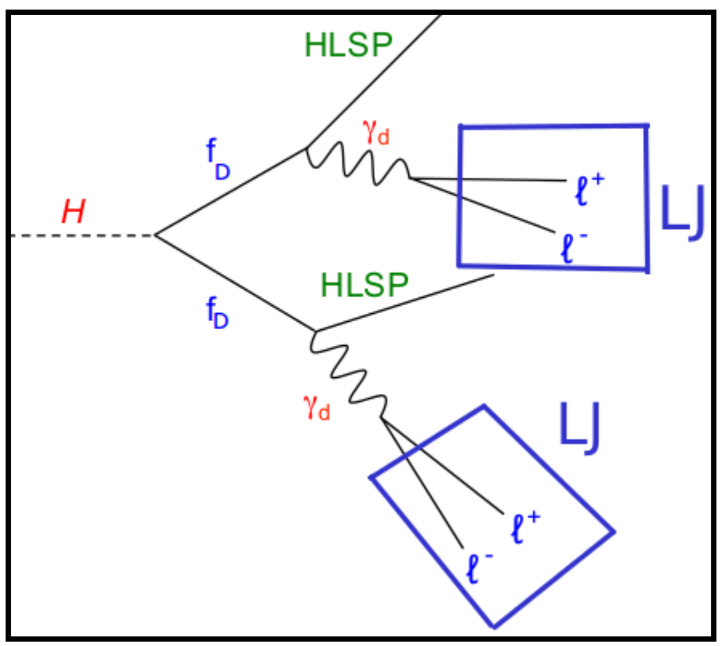


- dark photons from FRVZ simplified model: broad spectrum of signature both leptonic and hadronic

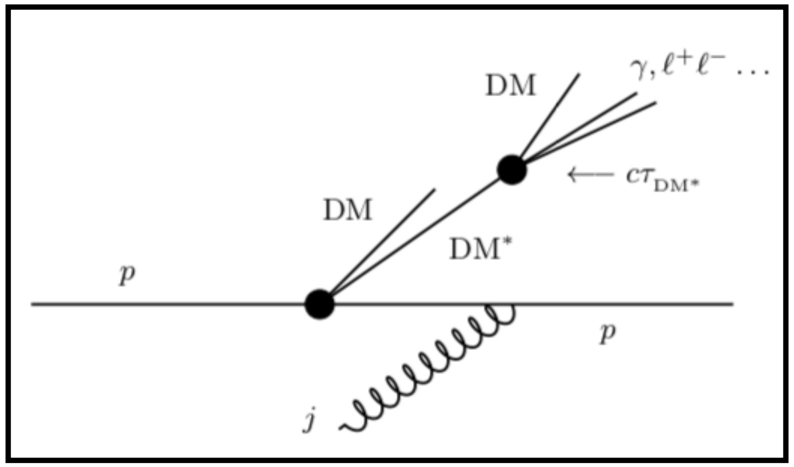
- same signature for inelastic Dark Matter model with a possible addition of an ISR jet

for small epsilon very displaced decays

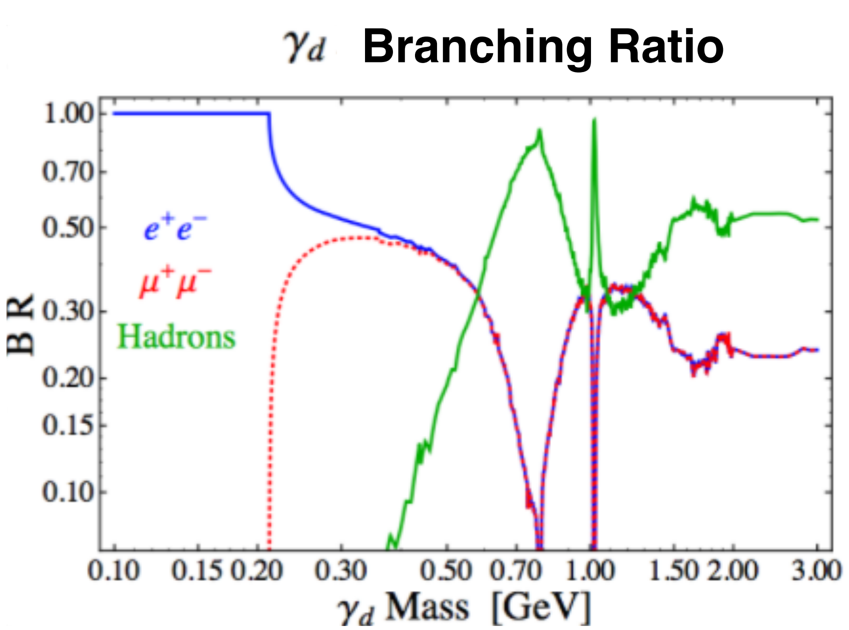
$$c\tau = \frac{1}{\Gamma_{tot}} \propto \frac{1}{\epsilon^2}$$



Falkowsky, Ruderman, Volansky, Zupan [FRVZ] arXiv:1002.2952

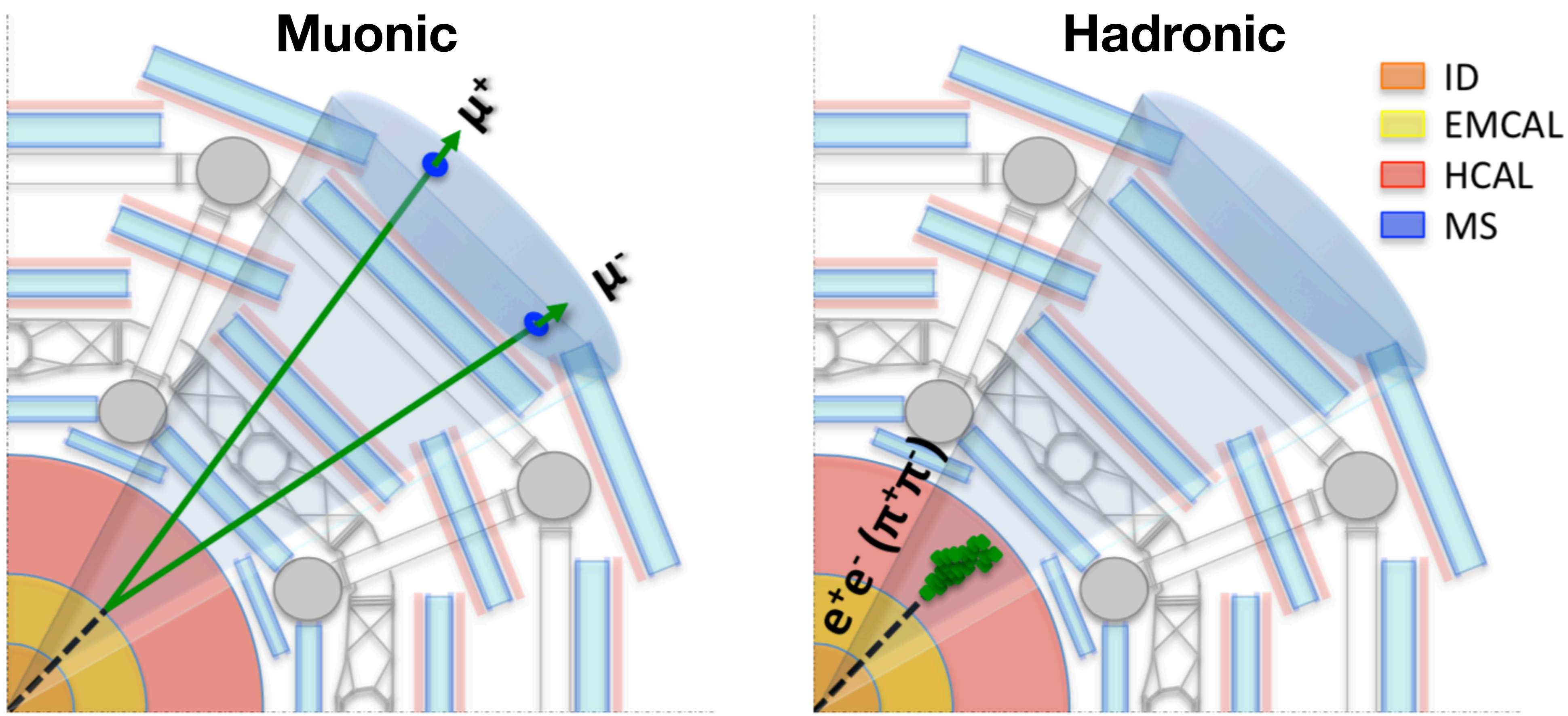


Izaguirre, Krnjaic, Shuve arXiv: 1508.03050



dark photon reconstruction

Schematic picture of the dark photon according to the decay final state



Muonic decay

Collimated bundle of muon without track in the inner detector

Difficult to trigger: low-pt muon and/or non-pointing to primary vertex

Hadronic decay

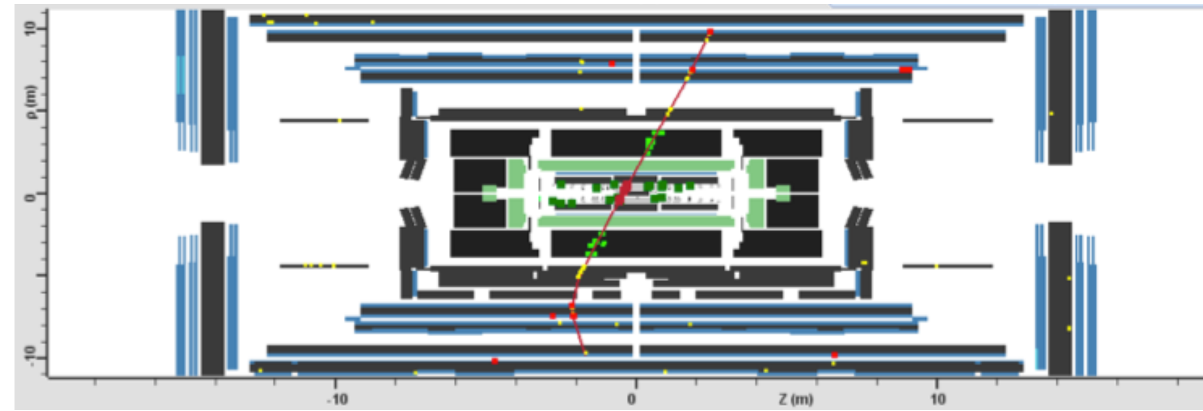
Displaced jet with most of energy deposit in the HCAL

Very high background and few handles to play with (especially in mono-dark photon channel)

Backgrounds

COSMIC

https://arxiv.org/abs/1011.6665



Cosmic muons

In Run2 estimated through triggered events in empty bunches

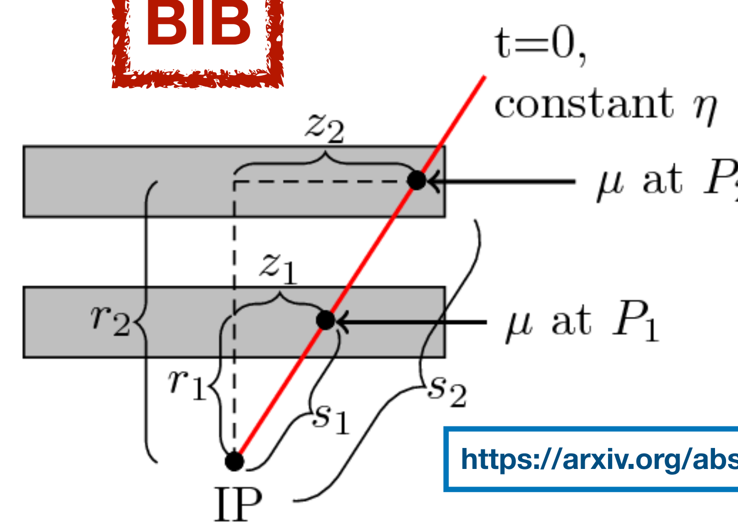
—> Very few events survive selection

Beam induced background (BIB)

Complicated non collision background. In Run2 we are able to remove them.

At HL-LHC with high pilup…

BIB



https://arxiv.org/abs/1810.04450

QCD

Main source of background for both prompt and displaced dark photons.

Displaced jets are rare QCD events nevertheless can outnumber the expected signal

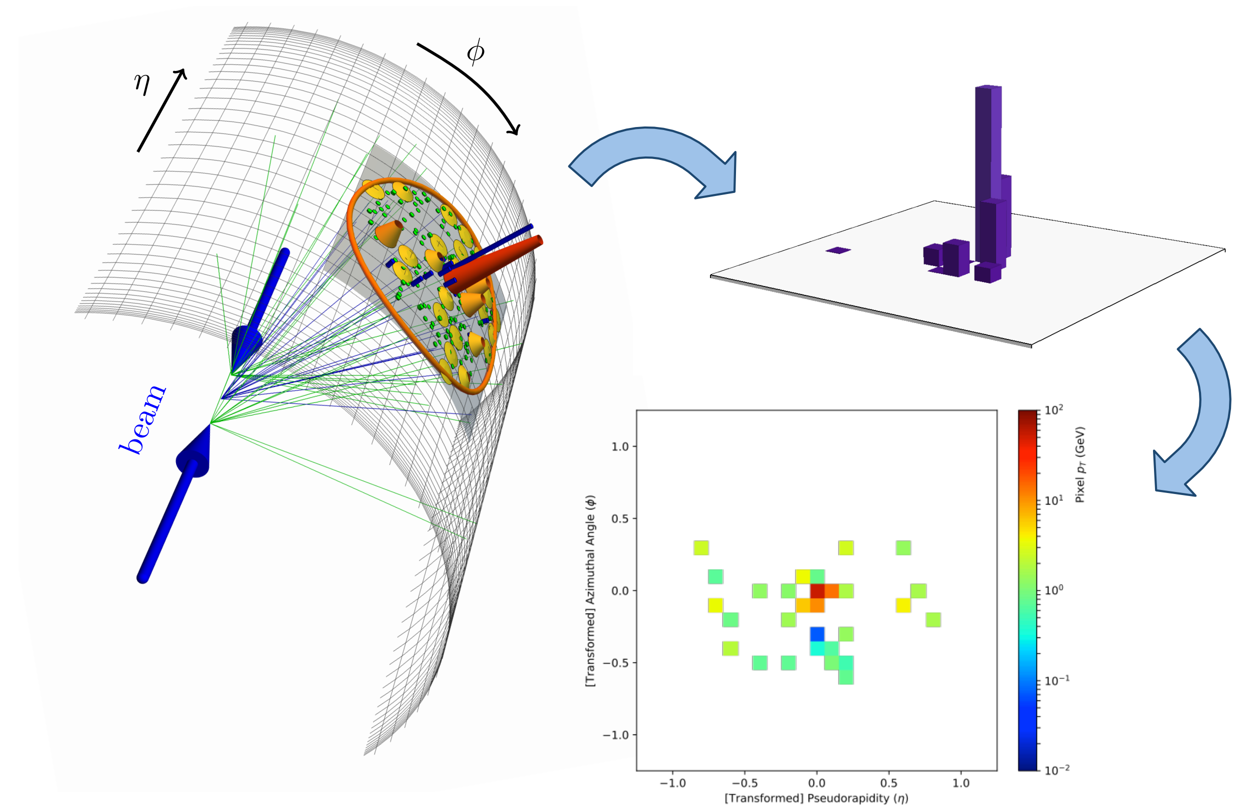
<u>Today's presentation</u>: exploit Deep Learning methods to better discriminate standard vs dark photons jets and to improve triggers for dark-photons decaying in muons

How Deep Learning can improve Dark Photon search

Deep neural networks are a powerful tool to exploit all the peculiar characteristics of displaced dark photon decays for identification and triggering

Hadronic dark photon decay:

- small jet width due to long decay lengths
- anomalous pattern of energy release in the different calorimeter's samplings
- substructures in the transverse energy deposit (e.g. two close-by charged pions)



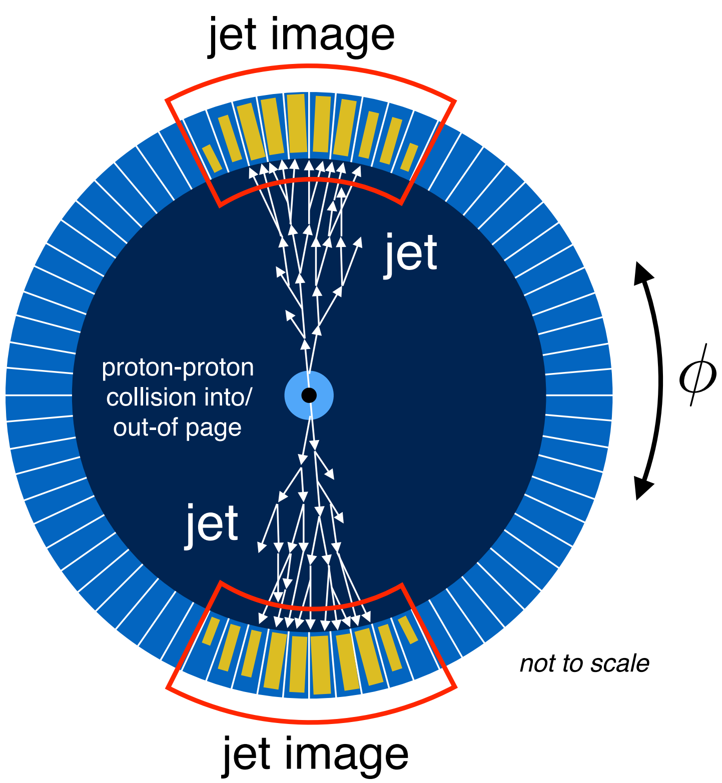
Low level input for the NN: energy deposits distribution in calorimeters

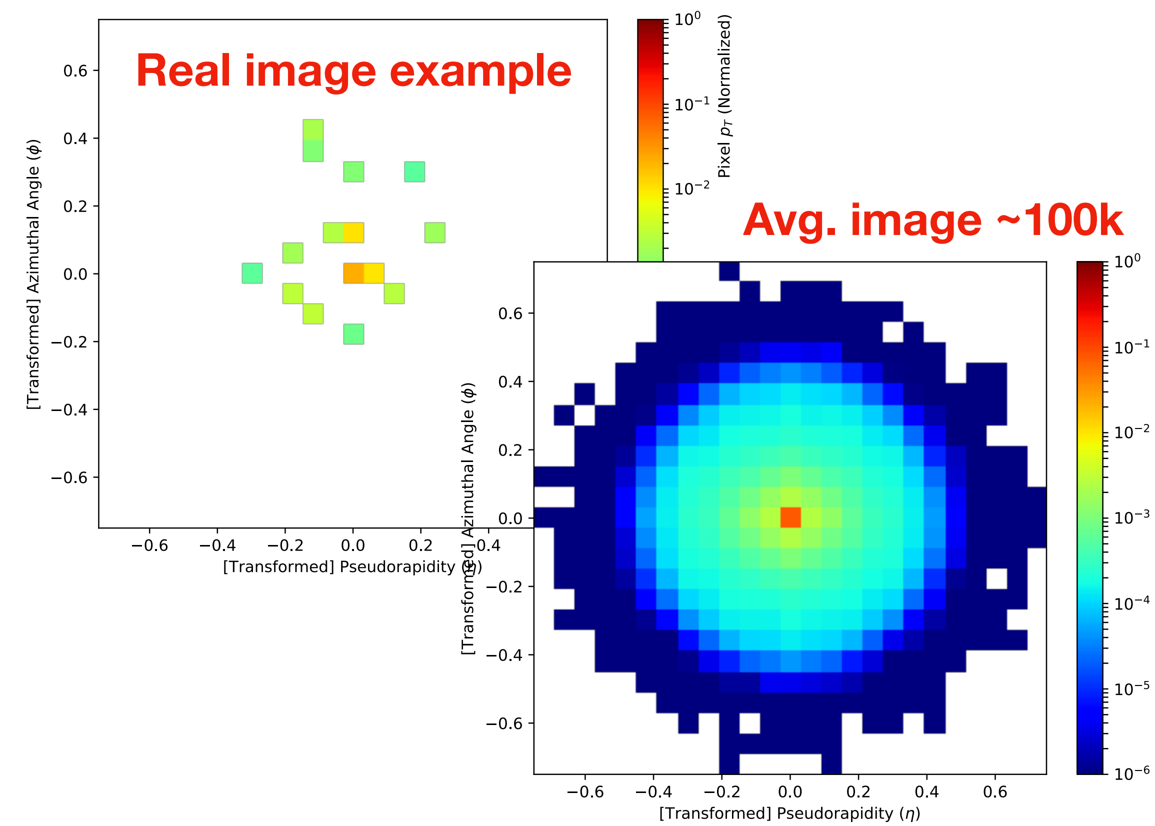
Today's presentation

Jet discrimination with a multi-layer convolutional neural network (CNN) trained on a mix of full-simulated examples and data augmented samples

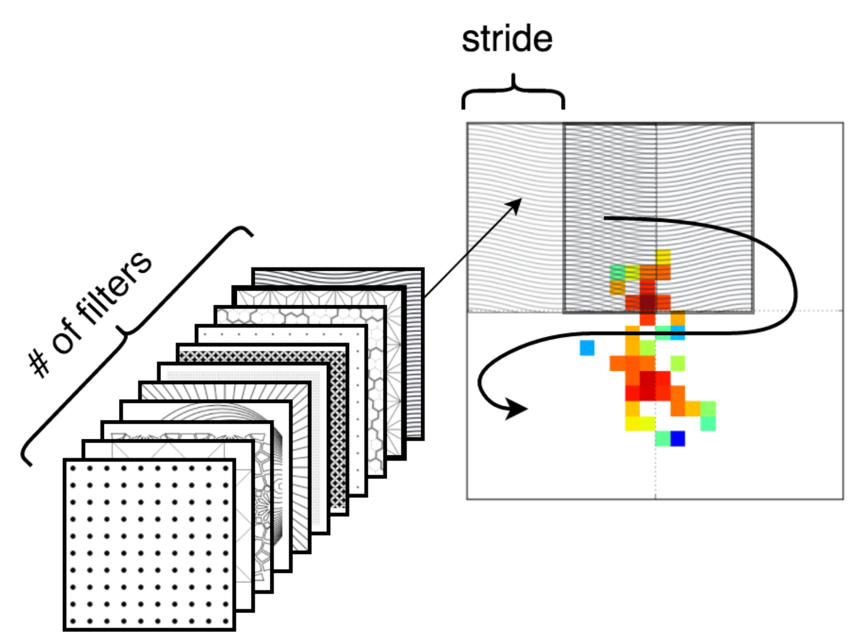
Jet images

Jet images (25x25 pixel) built from CaloCluster informations in (η x ϕ) plane
Two samples: QCD dijet and dark photon benchmark model



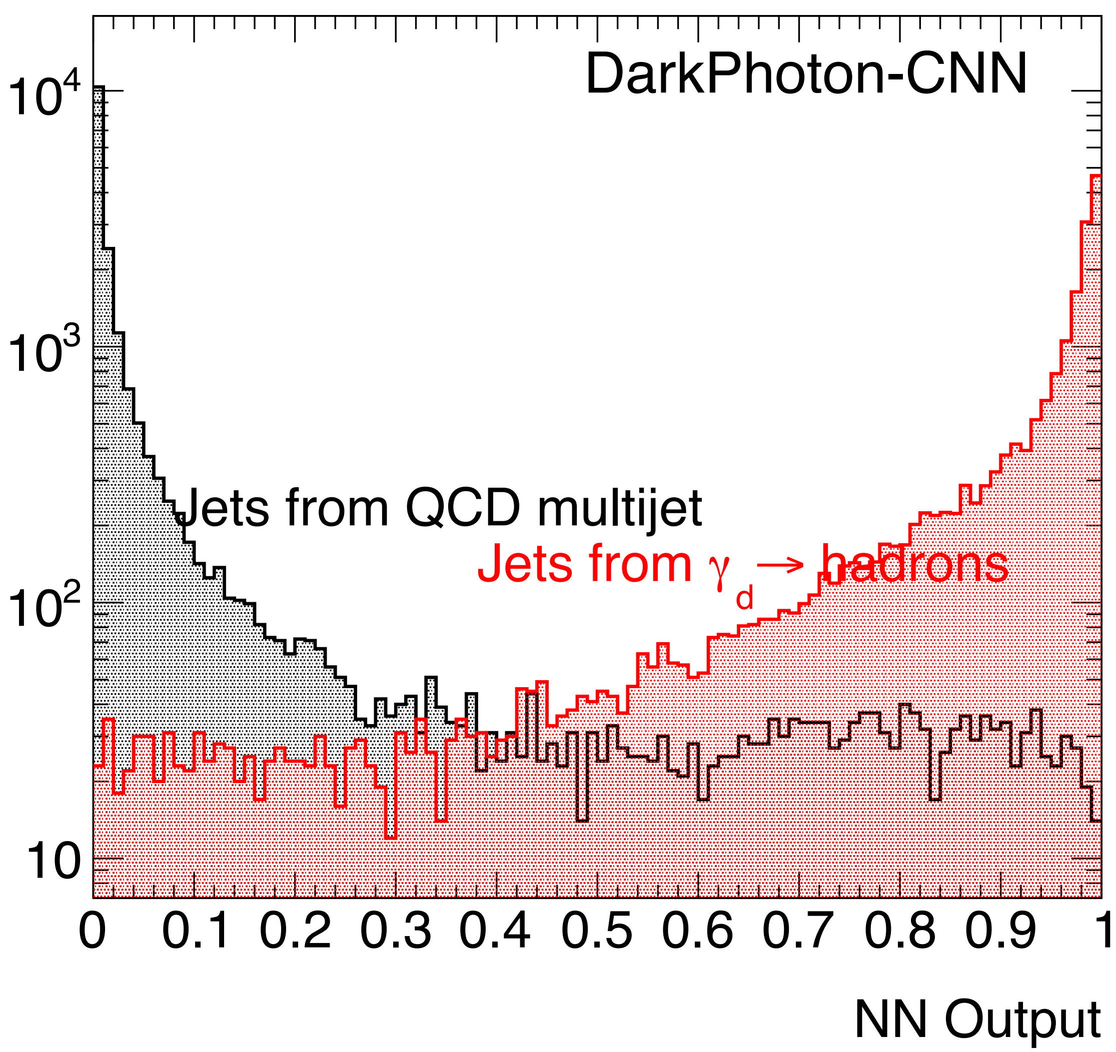


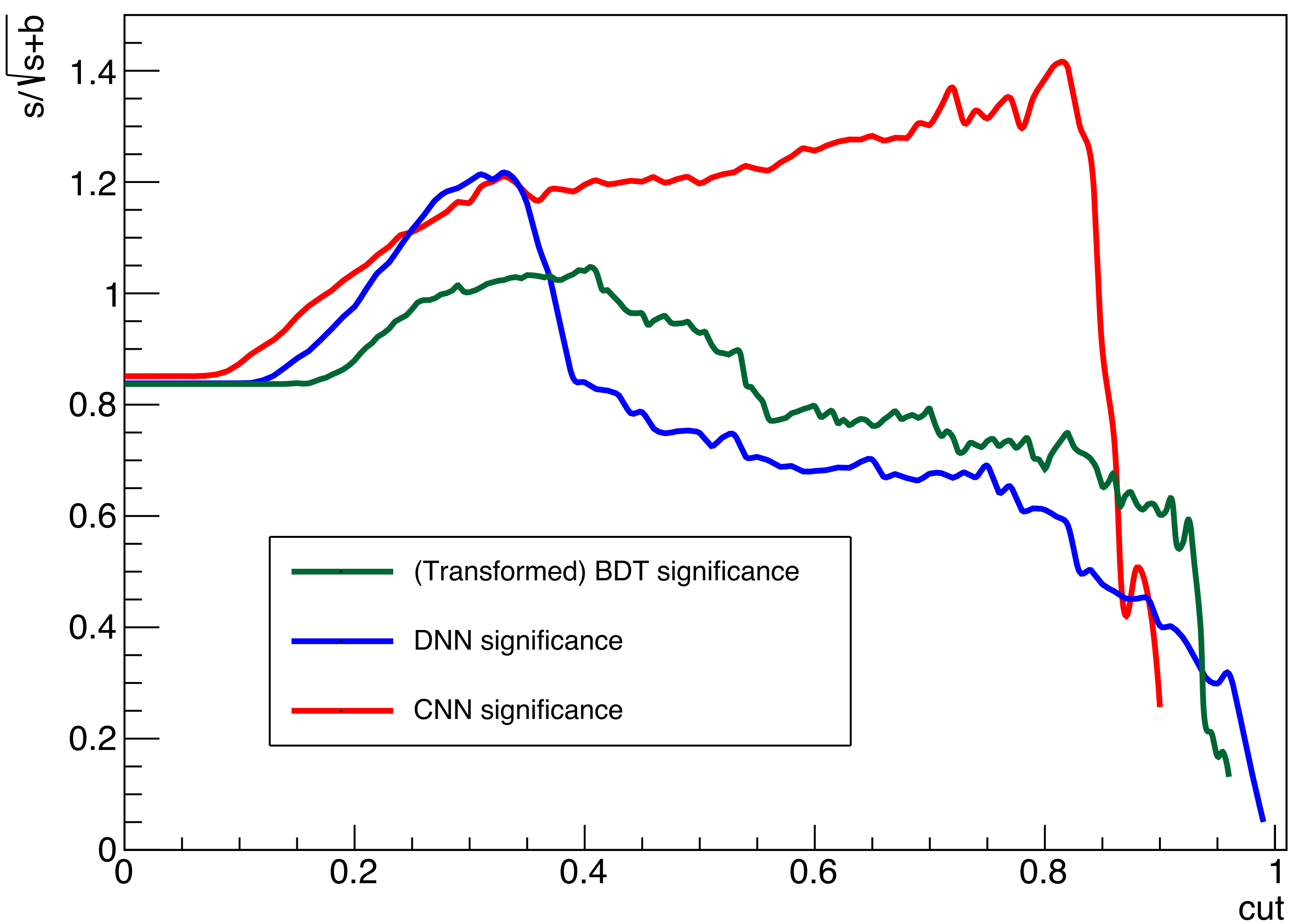
We first run a standard multi-layers CNN setup to test discrimination power of QCD from dark photon deposits in calorimeter



CNN output

Already with a very simple CNN and a limited training set obtained an improvement of 40% in sensitivity for dark photon identification with respect ATLAS conventional BDT based analysis





Data augmentation

To further improve the CNN discrimination power a very large training sample is needed.
We are computationally limited since we are interested in rare multi-jet QCD events!

We can exploit known physical symmetries to enlarge the sample

- jet invariance over azimuthal angle phi
- jet invariance over eta reflection

by doing transformation on images or..

<u>Time</u>
Generate full simulation events can take a significantly amount of time (min/event)

<u>Disk space</u>
All the simulated data need to be stored somewhere

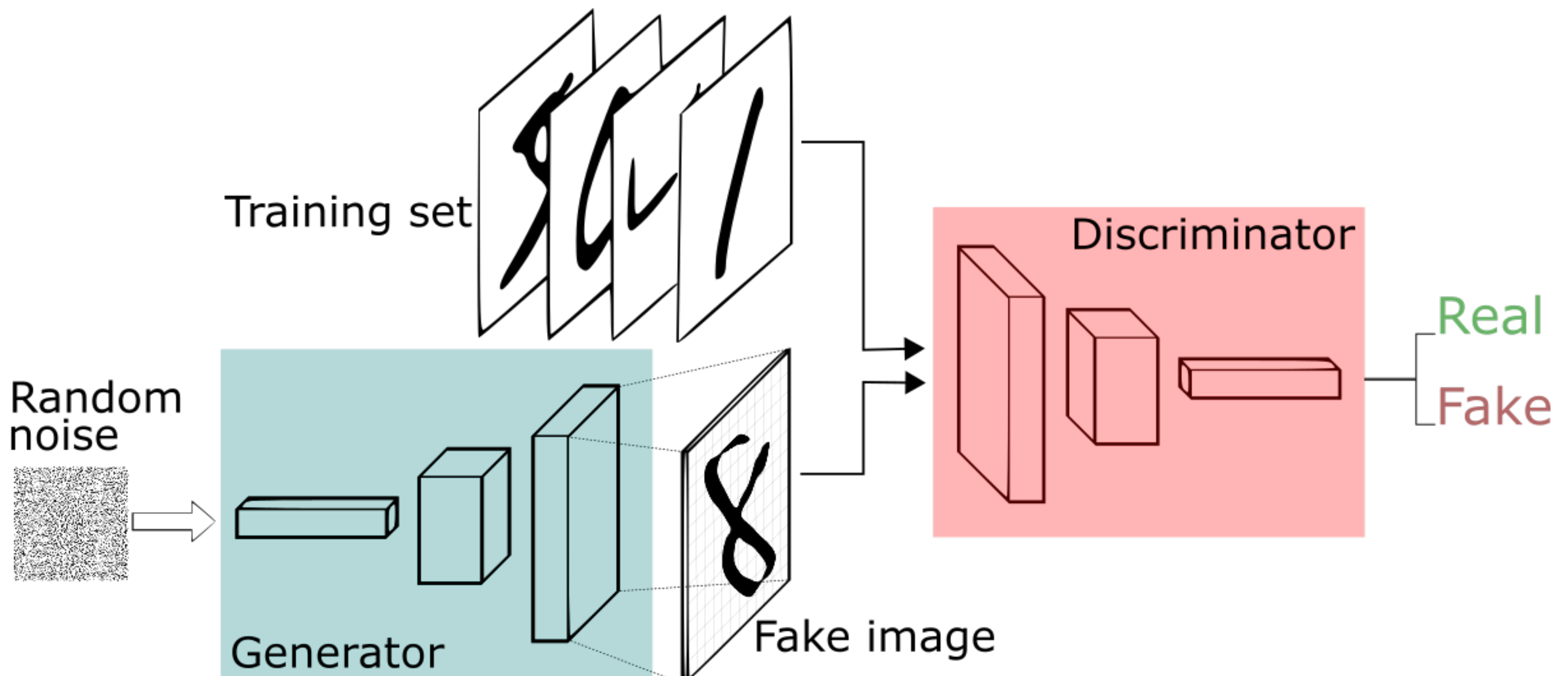
<u>Rare events</u>
To get tail-statistics event need a large number of events to be simulated

We can exploit generative DNN (GAN) to produce <u>very</u> large training samples! (paying attention to instabilities from which these algorithms suffer, plan to adopt new generation algorithms: WGAN-GP, SAGAN, RGAN..)

GAN for fast generation of jet images for CNN training

Exploit generative DNN (GAN) to produce very large training samples for hyper parameters tuning and optimal training of deep CNN

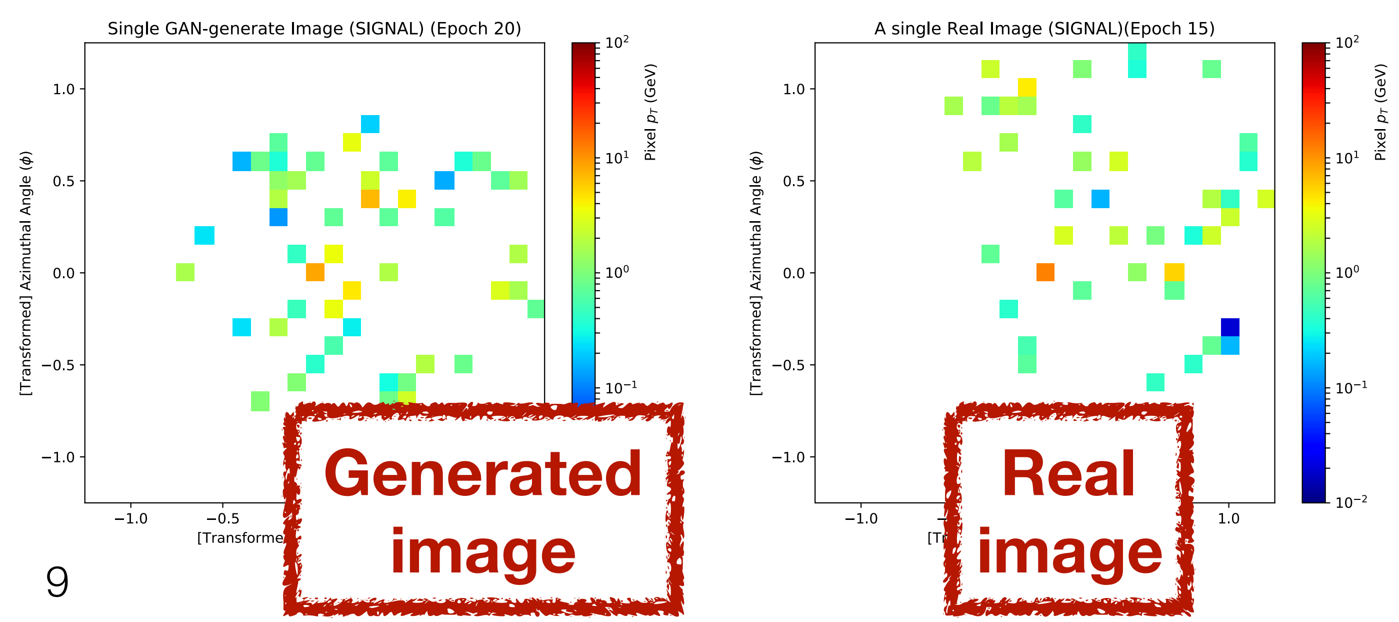
Training aims to make the system not able to distinguish fake from real images



Discriminator tries to determine generated from real images

Generator turns noise into sample images

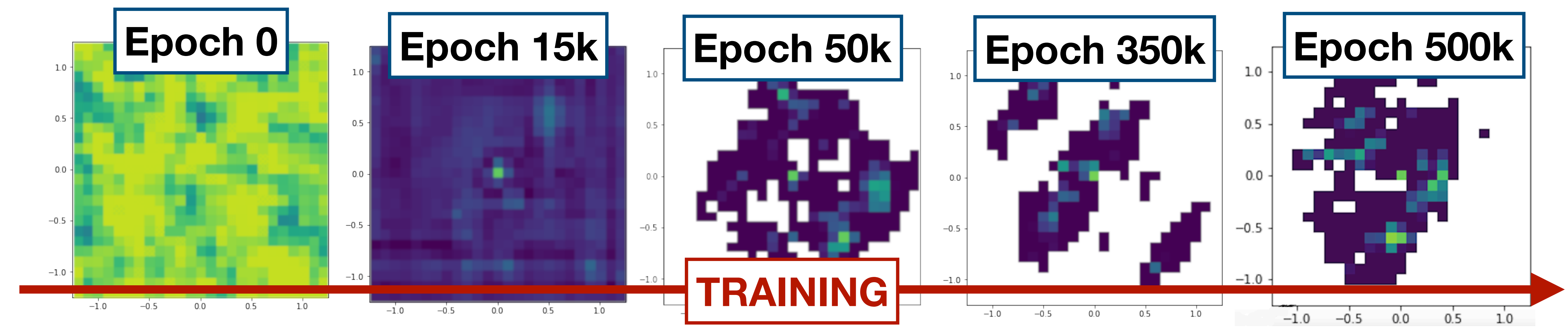
Output example for signal event trained generator

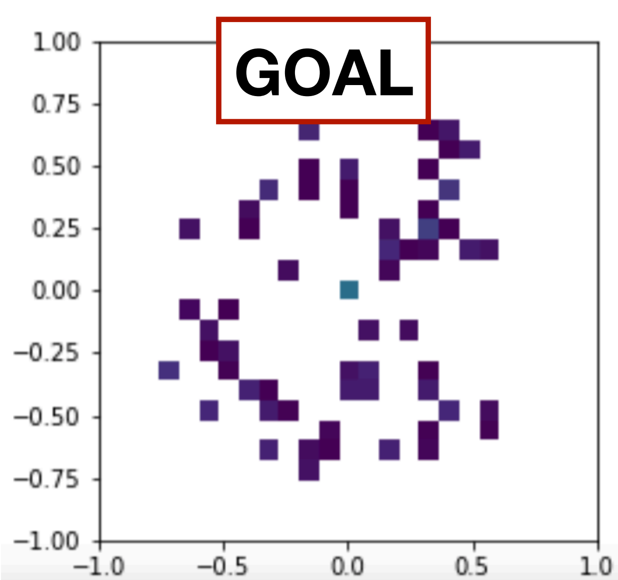


GAN outputs

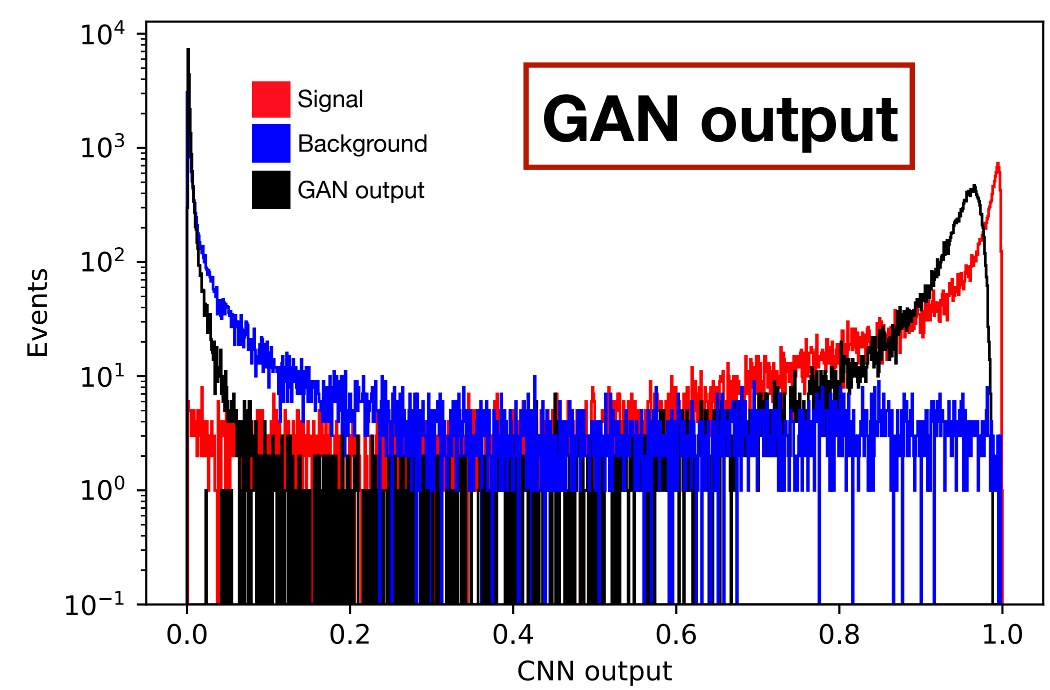
Basic setup for Generative DNN has been developed and used to generate realistic images of different physics processes. Very promising initial results…

New generation algorithms under development (Gradient-Penalty Wassertrain-GAN: WGAN-GP, Relativistic-GAN:RGAN) and post-processing image corrections algorithms, to improve stability and convergence of the system

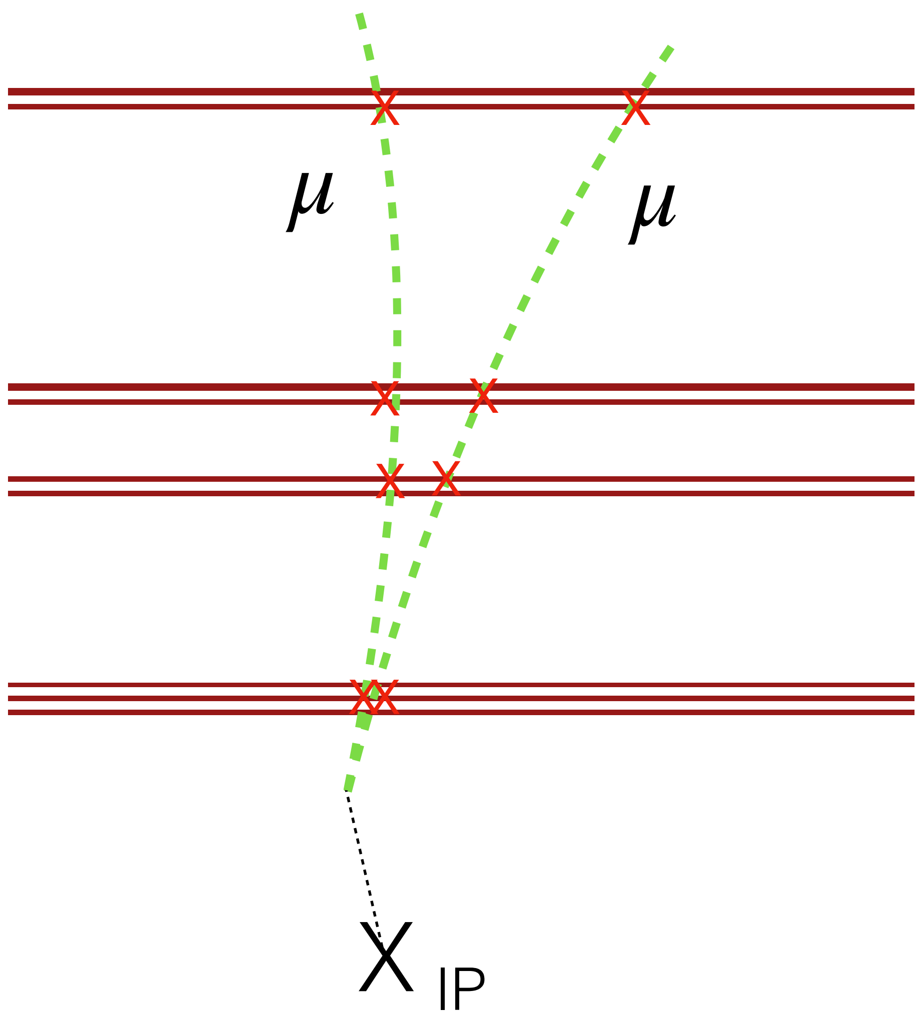




Model optimisation in progress, need large processing power/time. Able to reproduce main features of the events w/o specific tuning



How Deep Learning can improve Dark Photon search



Muonic dark photon:

- close-by muon pairs pointing to a common very displaced vertex (light dark photon)
- non pointing muons from not boosted decay (heavy dark photon >~few GeV)

Low level input for the NN: hit patterns in the Muon Spectrometer RPC chambers

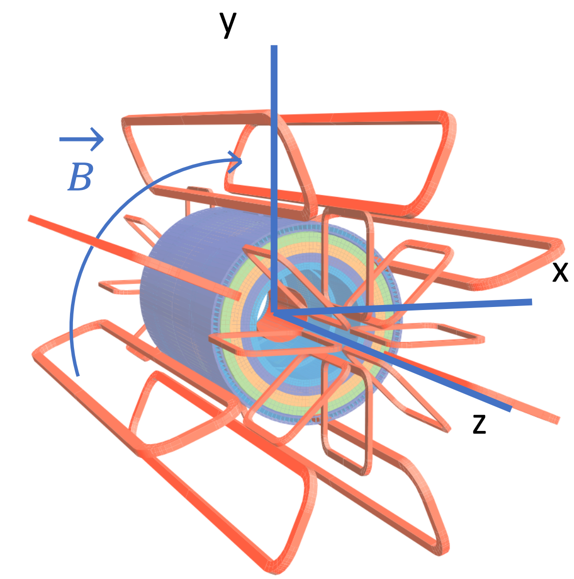
Today's presentation

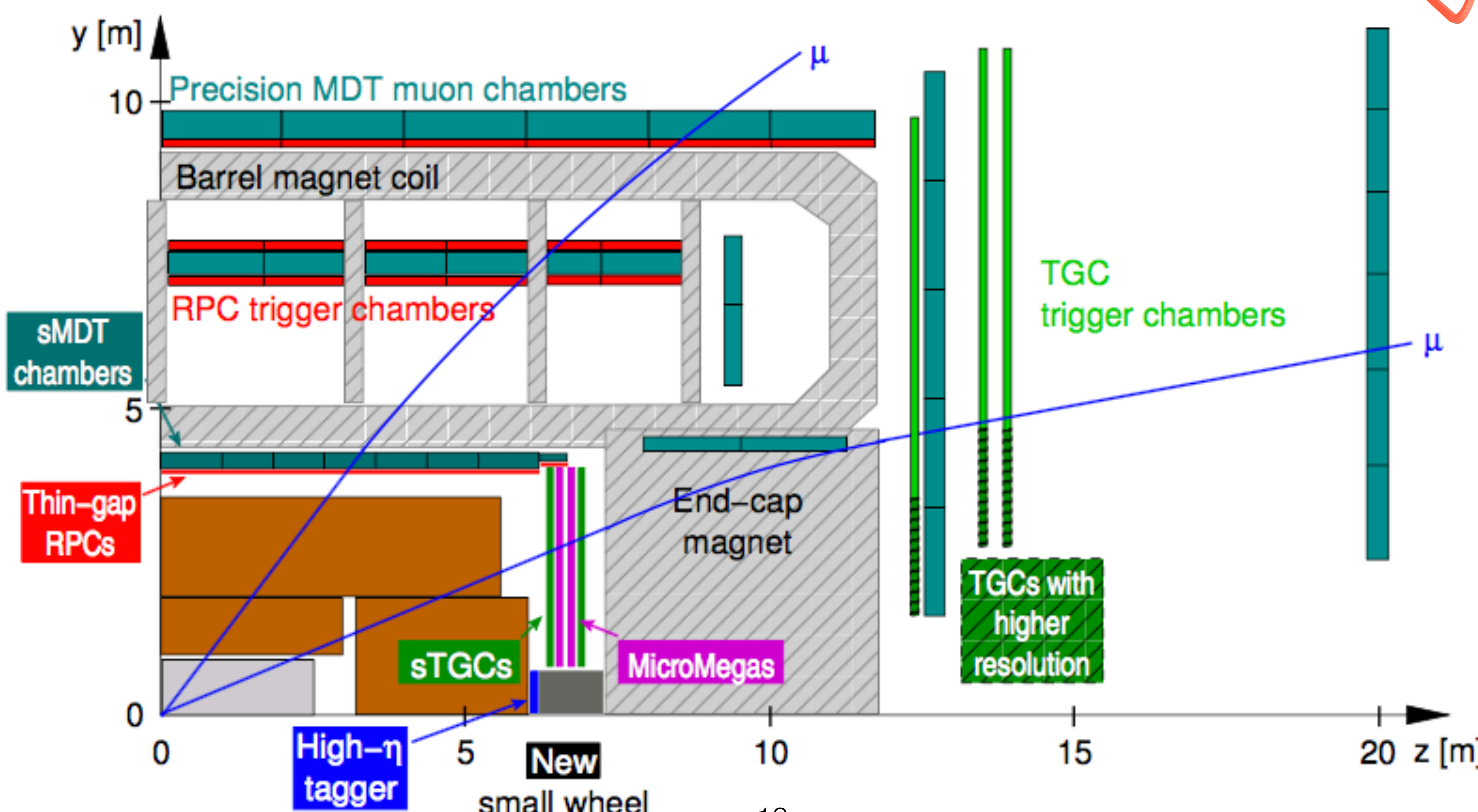
L0 muon trigger based on CNN to be implemented on FPGA @HL-LHC

- **Pointing and non pointing muons reconstruction**
- **Decay vertex reconstruction of long lived neutral particles decaying in two muons**

ATLAS Muon Spectrometer @HL-LHC

- New layer of RPCs to close acceptance gaps of barrel muon trigger
- New electronics for new trigger architecture based on FPGA Xilinx UltraScale+ boards

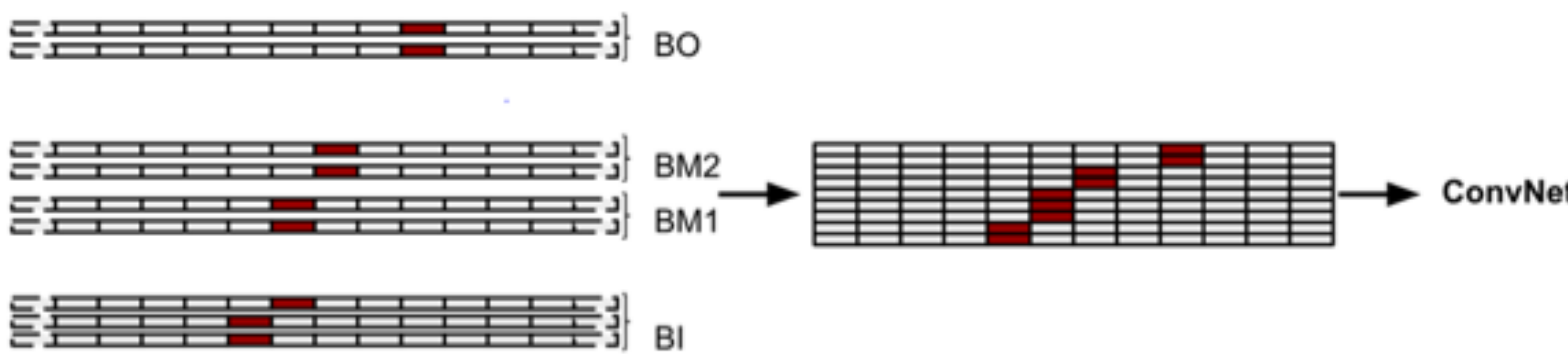


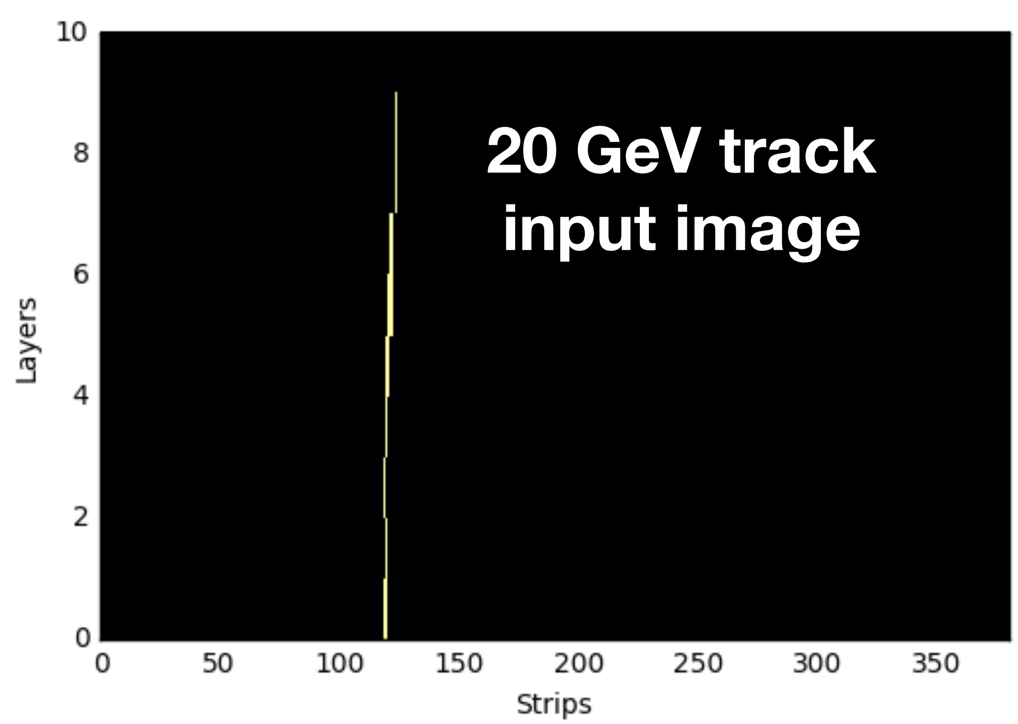


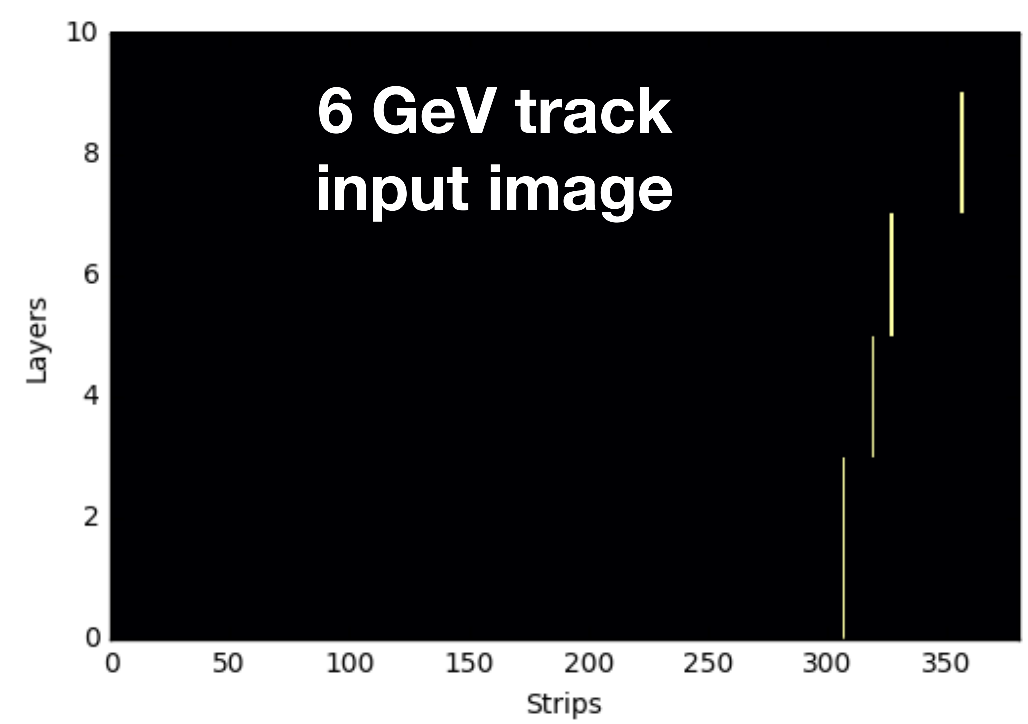
L0 muon trigger

Interpret RPC hit patterns as images and exploit ConvNet, to be run on new FPGA boards, to trigger events

9 RPC-layers X 384 strips each —> 3456 pixels per image!

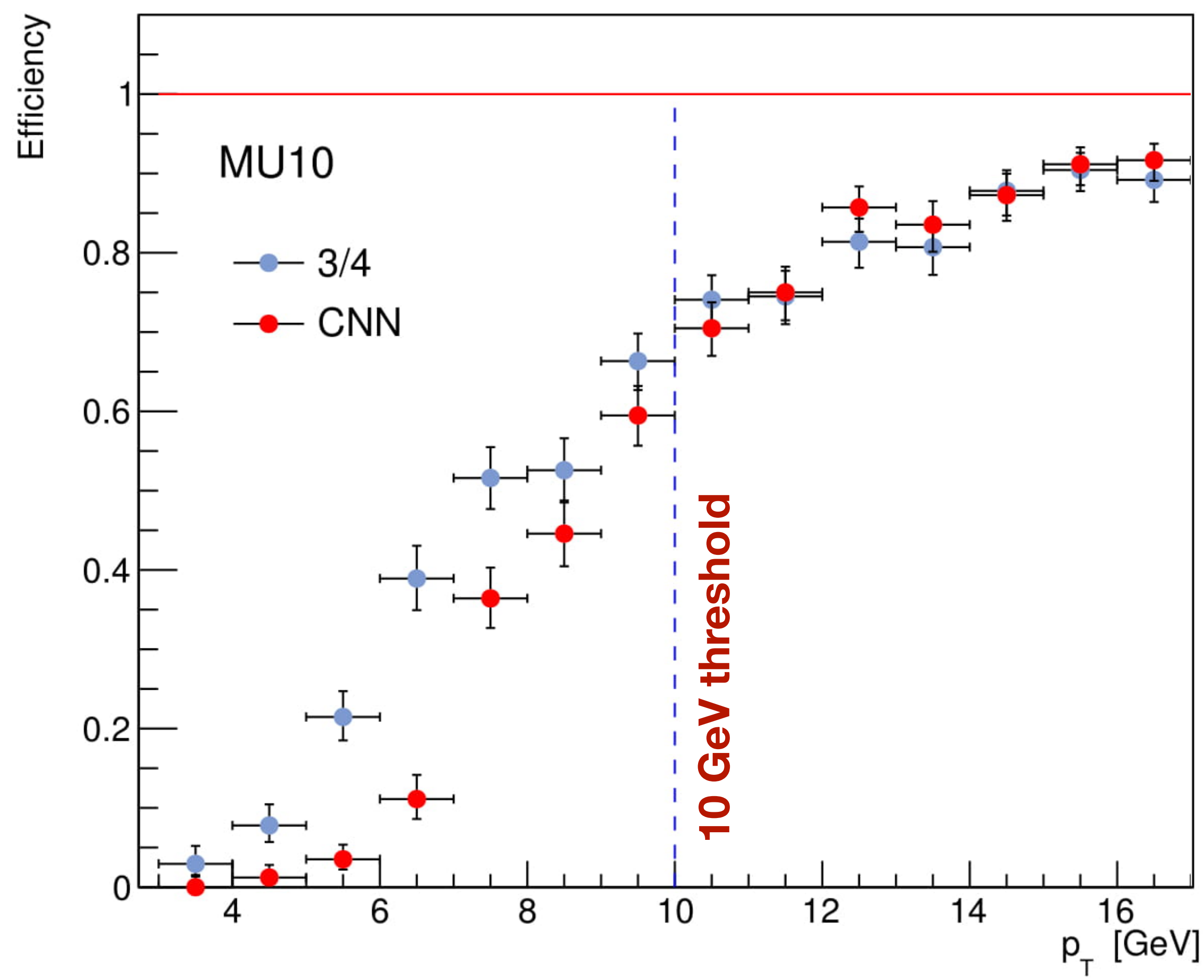






CNN muon trigger

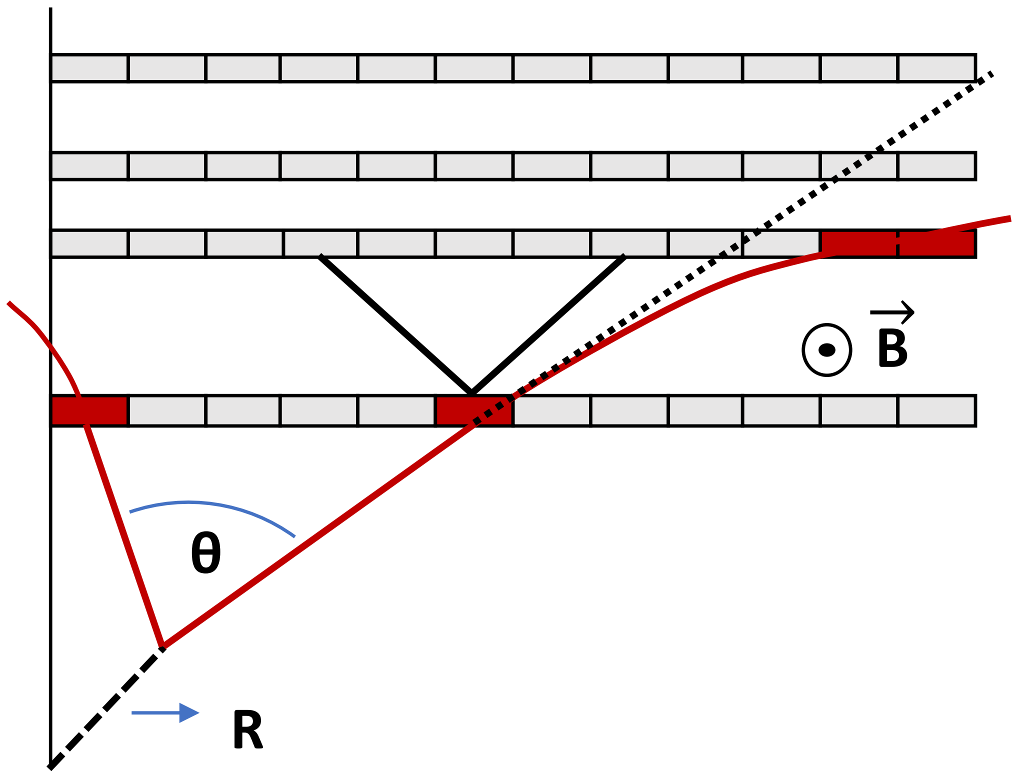
Simple ConvNet trained:
already able to match standard RPC algorithm performance



Next step is to test a newly developed Ternary-CNN to be used on FPGAs and evaluate resource/latency performances

CNN based vertexing in MS at L0

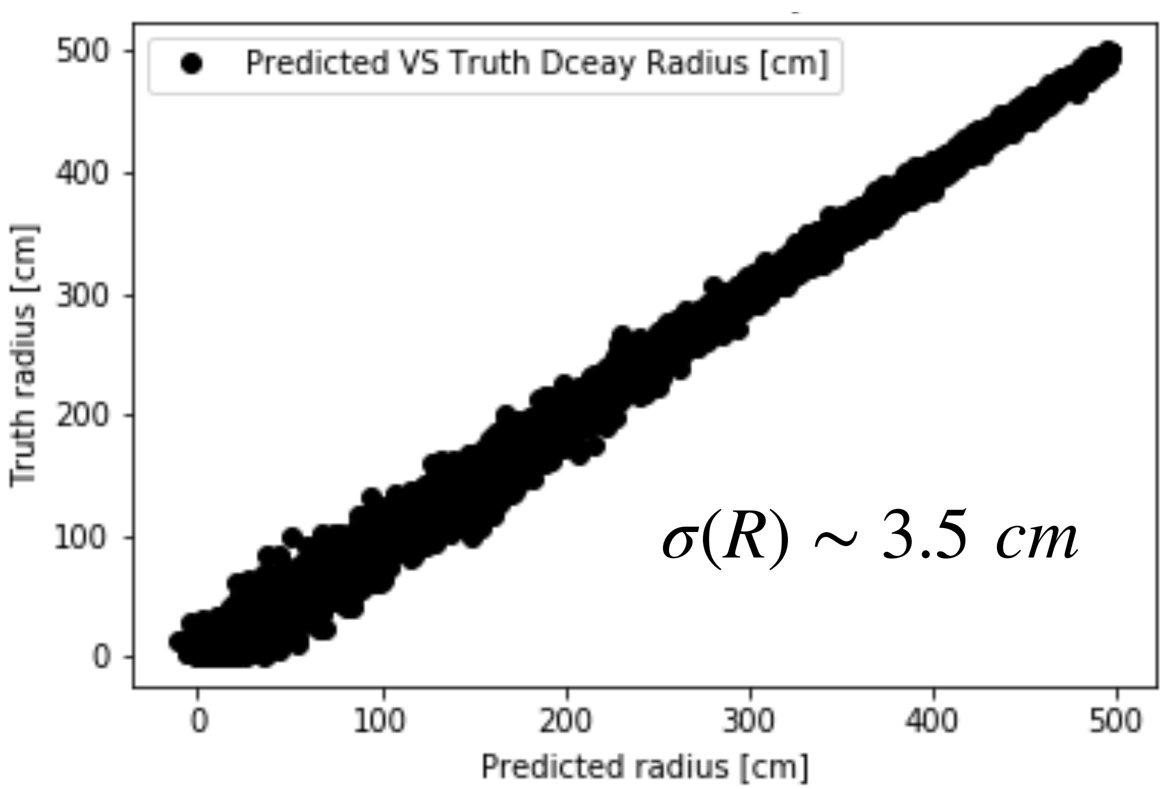
ConvNet trained for regression to find secondary vertex at L0



First test with simplified setup with 30k events only without cavern background noise:

$\theta(\mu,\mu) \in [\mathbf{0.05}, \mathbf{0.2}]\mathbf{rad}$

$\mu\ \mathbf{p_T} \geq \mathbf{20\ GeV}$



Will improve by using realistic sample with noise and high statistics.
Will also test ternary-CNN

Ternary ConvNet

These new CNN triggering algorithms will have to run in new generation FPGA boards and have to be compatible with all the specification of the L0 ATLAS muon trigger @HL-LHC. Have to be designed to be: fast, low latency and light!

Testing ternary CNN with 2 bit output instead of f32:

- Intrinsically runs faster on FPGA, no need for pruning and routing information is simpler thus yielding a lower latency.
- Requires 16 times less memory than f32 (FPGA memory is precious)
- Algorithms are more effective when running on sparse images (where most ‘pixels’ are not significant)

f32 CNN (10.5 Mpars):

- accuracy: 94.5%
- compression: x1

T-CNN w/same expressive power as the simple f32 CNN (0.7 Mpars):

- accuracy: 92.5%
- compression: x16

Currently testing ternary CNN on ‘altera’ FPGA reaching the same accuracy of the reference f32 CNN

Conclusions and future plans

- **Deep NN methods have been tested and have good performance on managing sparse images, we can exploit these tools to search for dark photons at LHC**
- **Initial studies show promising results for generation and discrimination of jet images of calorimetric deposits from exotics object**
- **CNN based triggering and vertexing in MS at L0 for HL-LHC can be achieved and can be used to improve selection of LLP**

Very interesting results… but plenty of room for improvement!

Hadronic Dark Photon

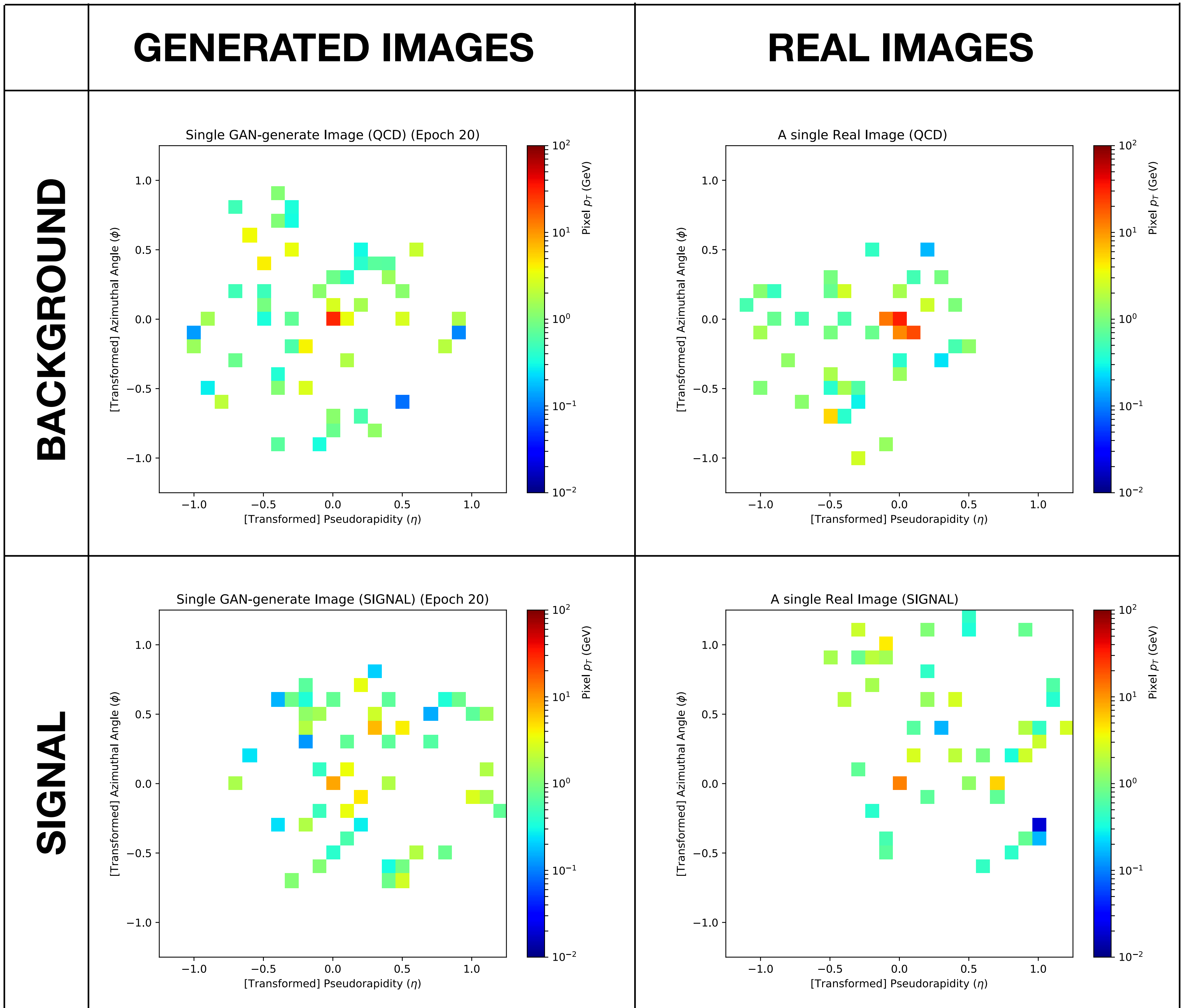
- Optimise WGAN-GP/RGAN data-augmentation for high statistic training of the CNN
- Tune NN for displaced jet identification on GAN generated events
- Implement multi-layer jet images by exploiting calorimeter sampling

Muonic Dark Photon

- Finalise development of Ternary-CNN and test implementation in new generation FPGAs firmware
- Test performances on FPGAs in realistic case scenario to evaluate latency and resource consumption

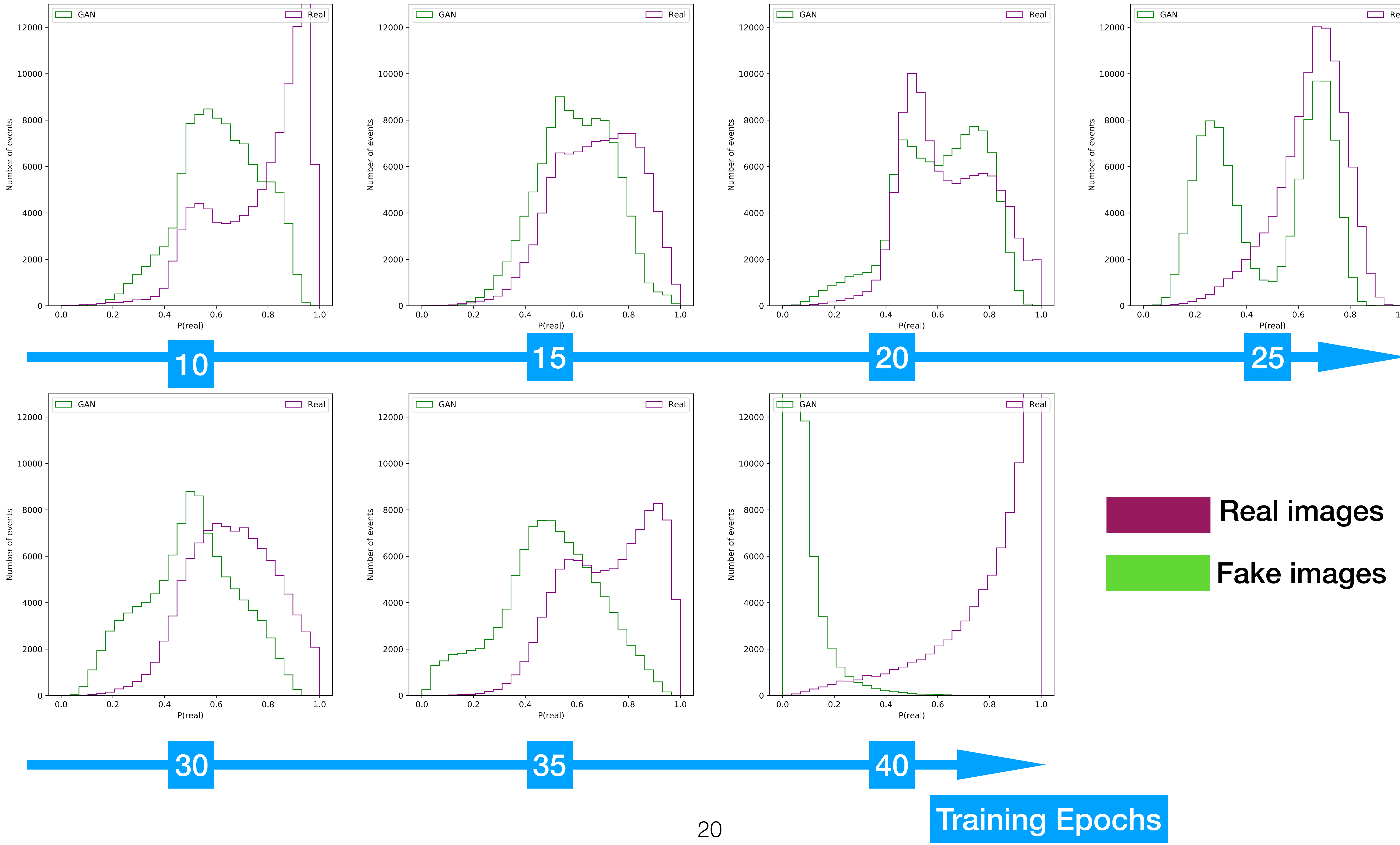
Backup slides

generated images



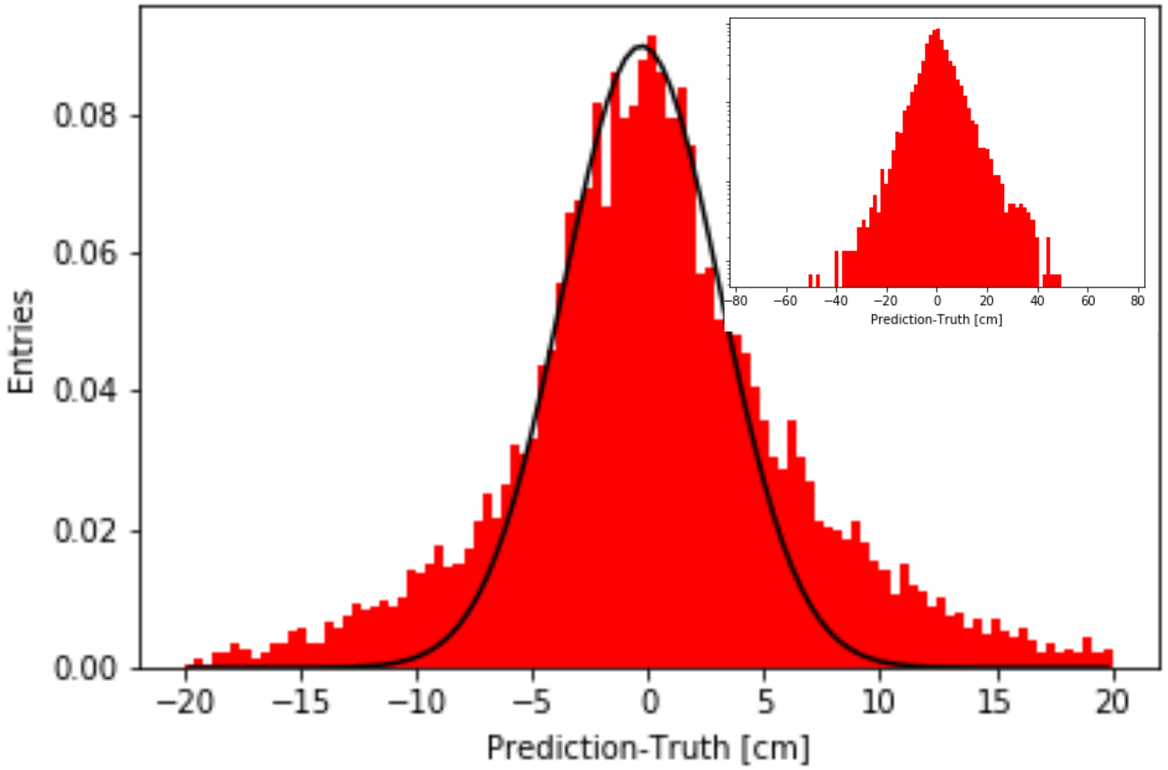
GAN tests

Discriminator output: P(real)



L0 vertexing setup

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d_1 (Conv2D)            (None, 8, 358, 32)        224
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 8, 119, 32)        0
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 7, 117, 64)        12352
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 7, 39, 64)         0
_________________________________________________________________
conv2d_3 (Conv2D)            (None, 6, 37, 128)        49280
_________________________________________________________________
max_pooling2d_3 (MaxPooling2 (None, 6, 12, 128)        0
_________________________________________________________________
conv2d_4 (Conv2D)            (None, 5, 10, 256)        196864
_________________________________________________________________
max_pooling2d_4 (MaxPooling2 (None, 5, 3, 256)         0
_________________________________________________________________
flatten_1 (Flatten)          (None, 3840)              0
_________________________________________________________________
dense_1 (Dense)              (None, 128)               491648
_________________________________________________________________
dense_2 (Dense)              (None, 128)               16512
_________________________________________________________________
dense_3 (Dense)              (None, 1)                 129
=================================================================
Total params: 767,009
Trainable params: 767,009
Non-trainable params: 0
_________________________________________________________________
```



core: σ~ 3.5cm, μ~-0.3cm
tails: A(|Δ|>8cm) < 20%