ORGANIZAN
cedia · UCUENCA

ticec.cedia.edu.ec

TIC 2023
ec 11TH

SÉ PARTE DE MÁS

EXPLORING THE PERFORMANCE OF DEEP LEARNING IN HIGH-ENERGY PHYSICS

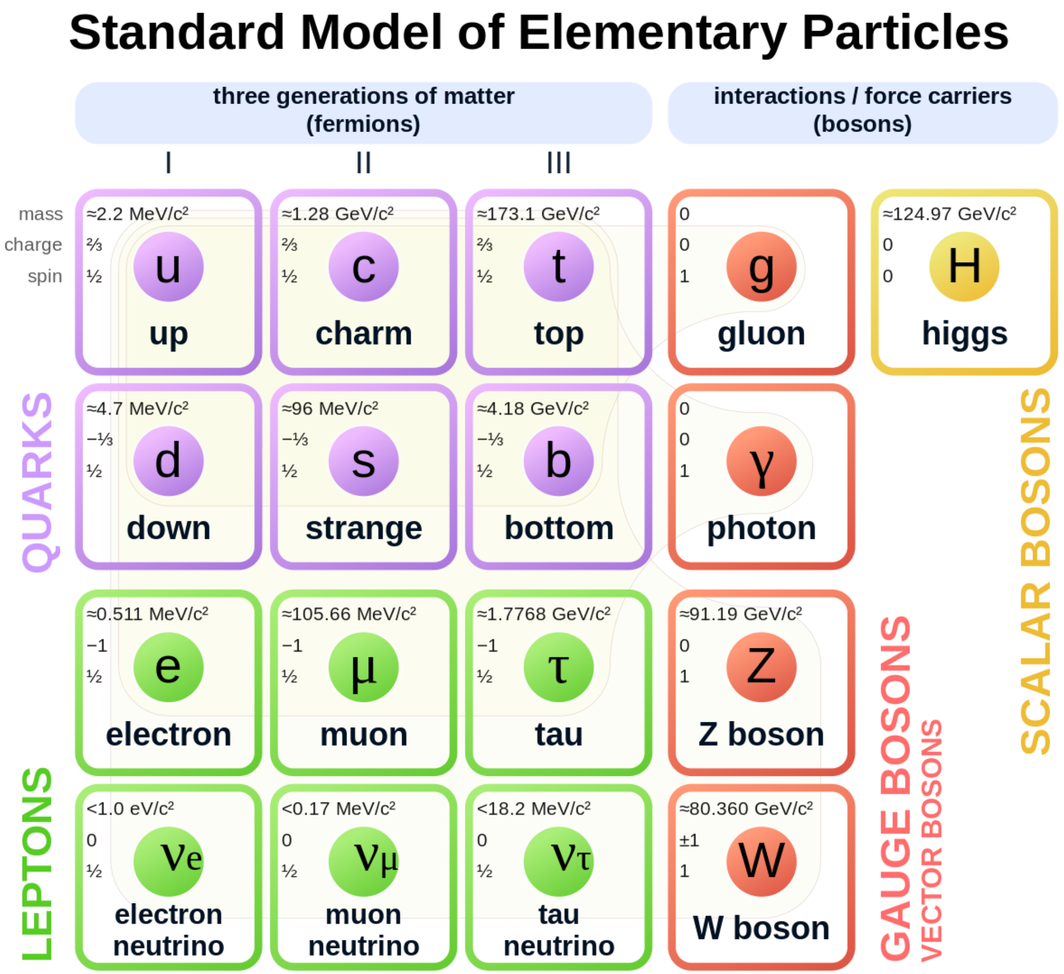
José Ochoa, Daniela Merizalde, Xavier Tintin, Edgar Carrera, Diana Martínez, David Mena

10-18-2023

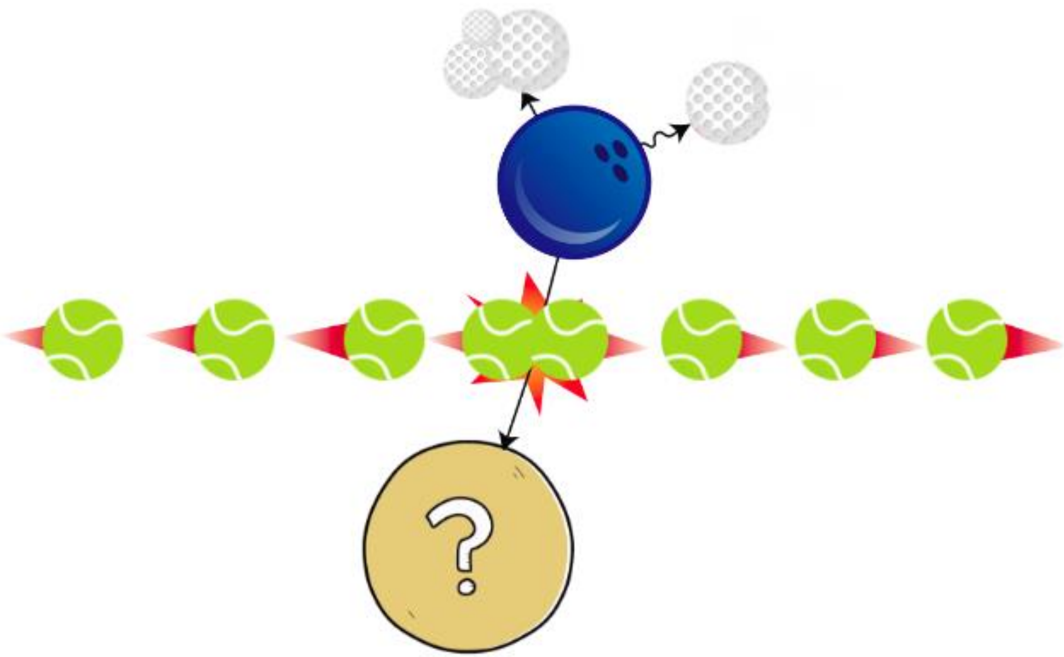
Table of contents

Introduction

The universe we see and how to "break" it



$$E = mc^2$$



LHC and CMS

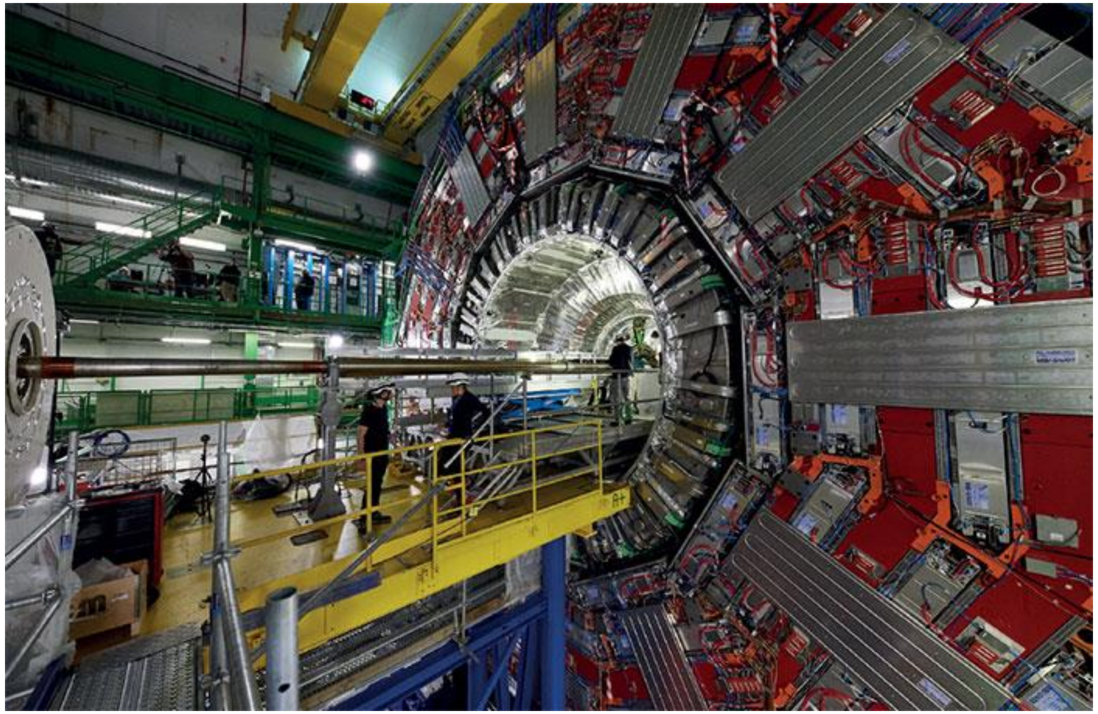
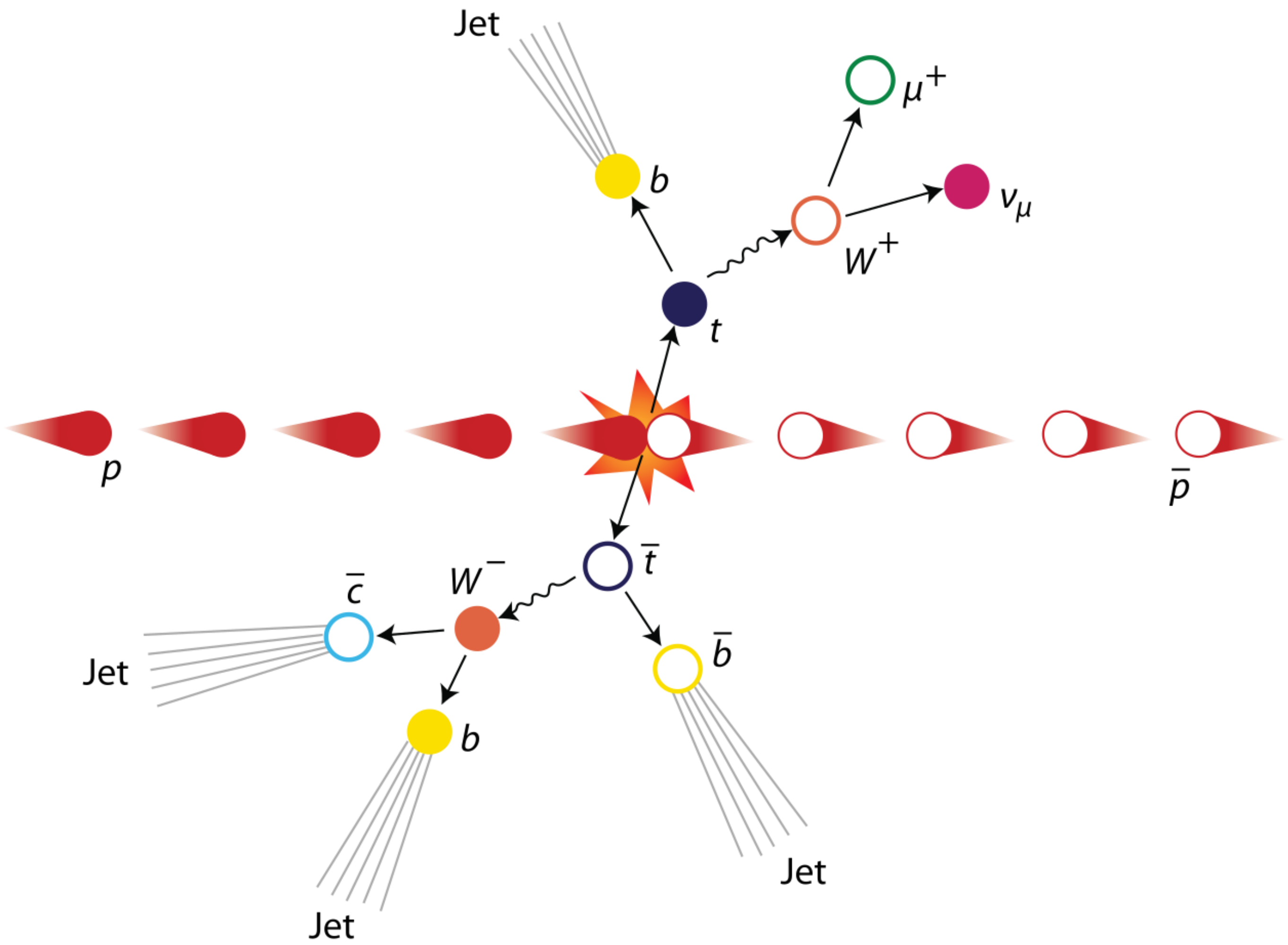


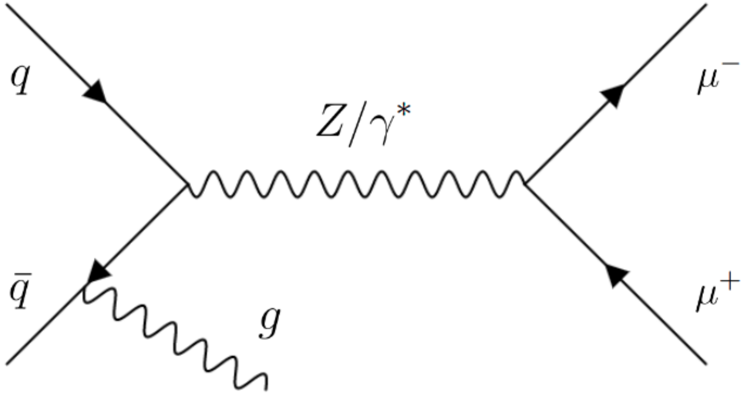
Fig 1: Distribution of the different detectors at LHC and the CMS detector[1]

A physics process

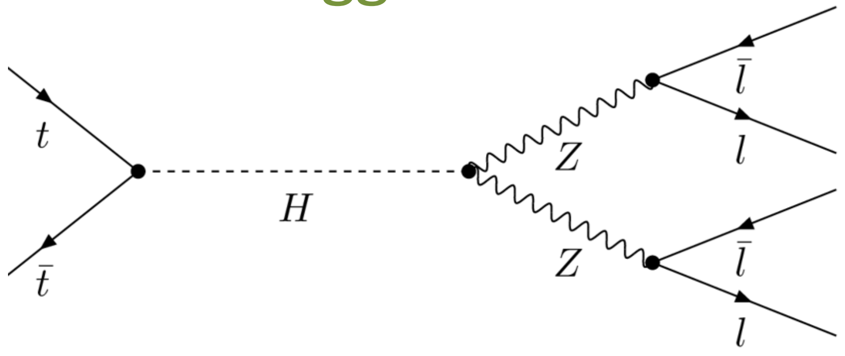


The DHJW scenario

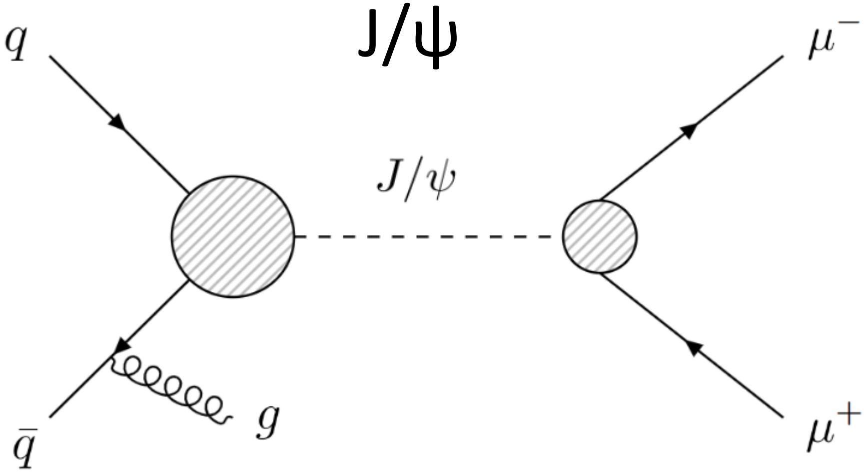
Drell-Yan



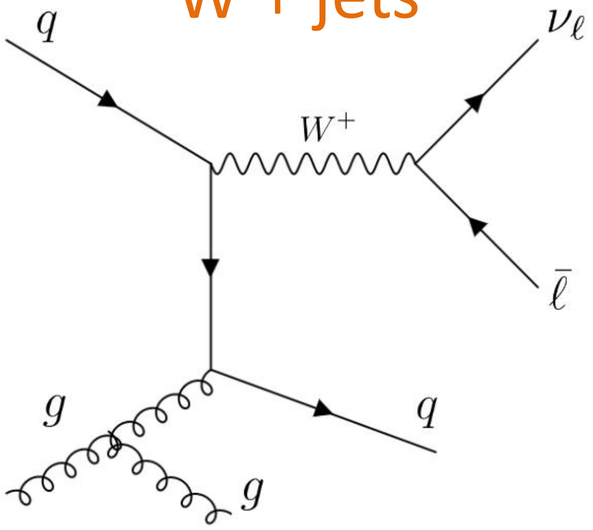
Higgs



J/ψ

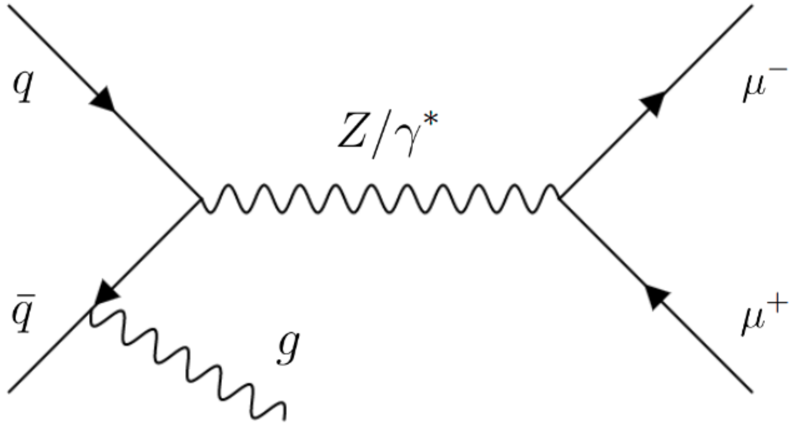


W + jets

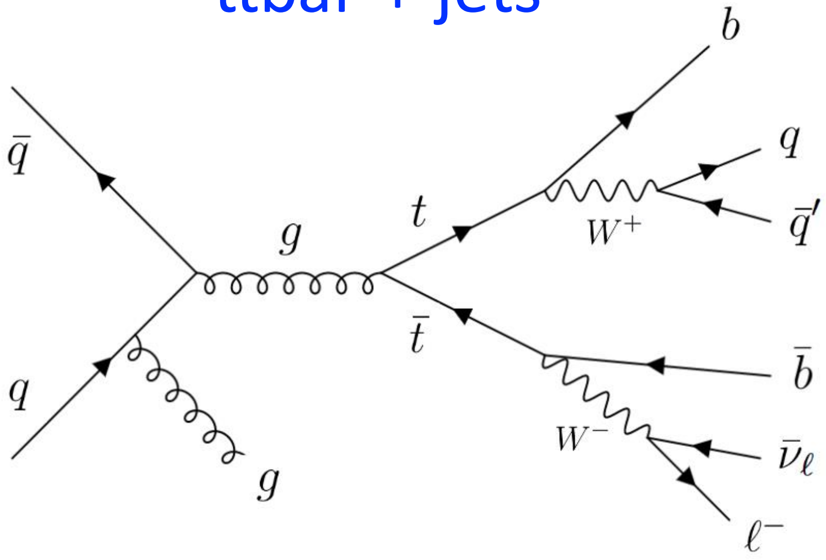


DTW scenario

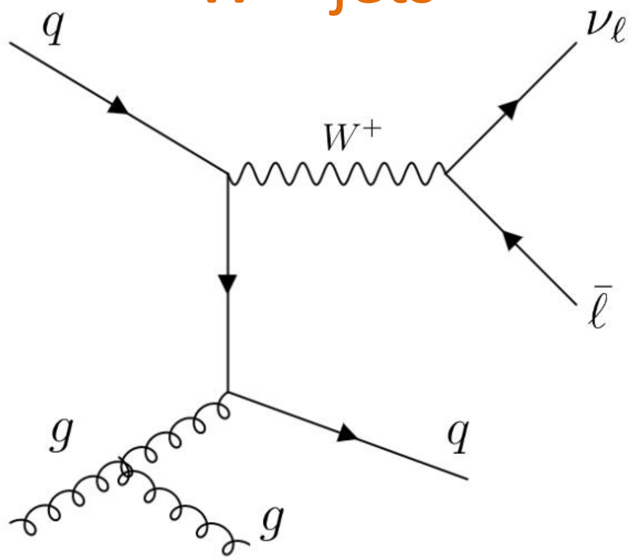
Drell-Yan



ttbar + jets



W + jets



General Approach

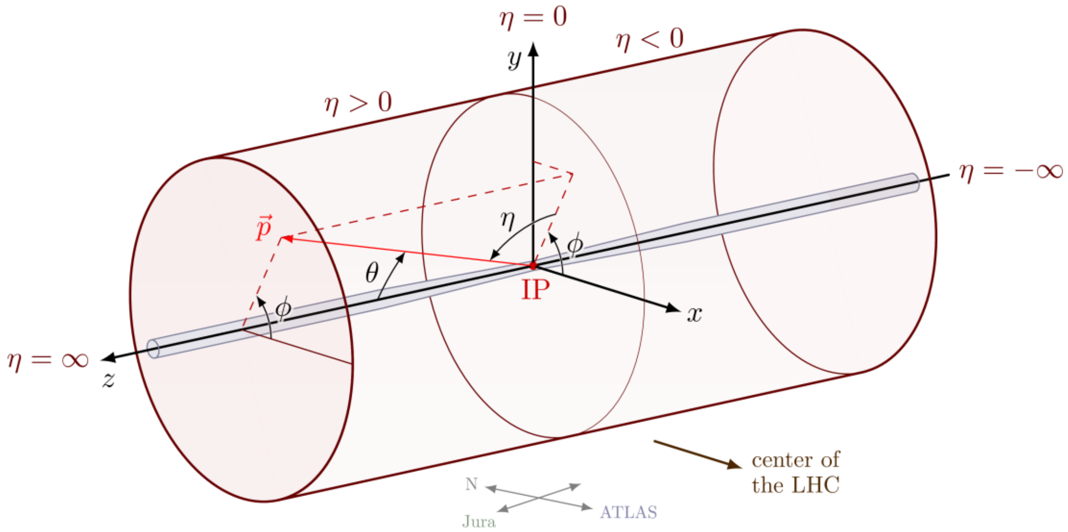
1. Extract relevant information from CERN Open Data Portal
2. Use that information to generate images (two scenarios)
3. Use the images to train various CNN's architectures
4. The trained neural network with the best performance metric is employed to classify real collision data

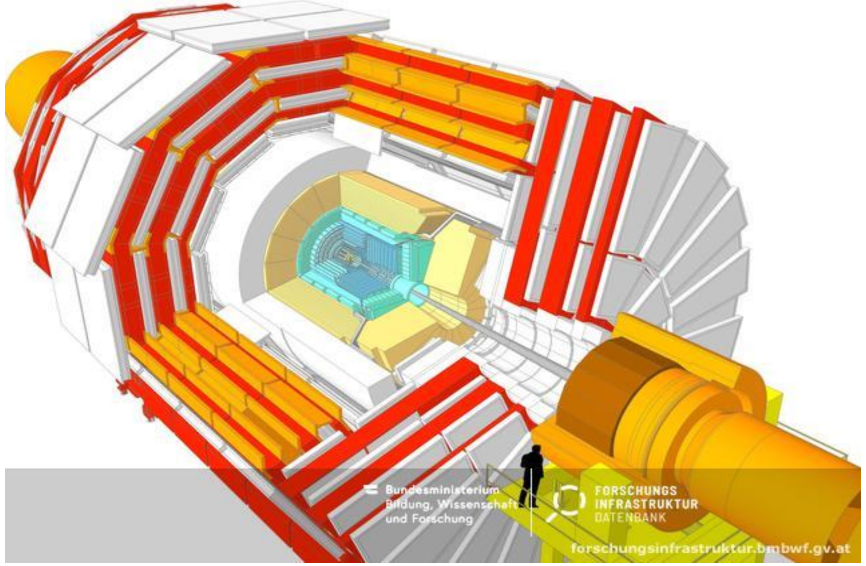
Methodology

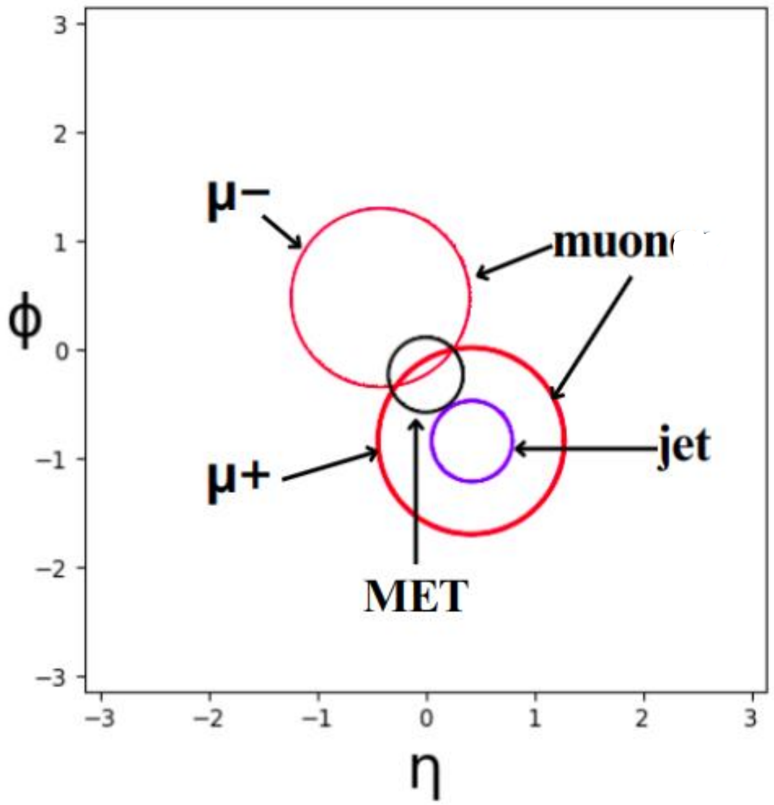
Data collection and information extraction

- The datasets used correspond to the simulation and real data obtained in 2015 during CMS Run II at 13 TeV. Open Data
- Muons, jets, MET

Image Generation







$$p_T = \sqrt{p_x^2 + p_y^2} \qquad R = \alpha \cdot \ln p_T$$

DHJW

Dataset	Number of Jets	Number of Images
A	0	110796
B	1	110796
C	2	64028

Table 1: Number of jets and number of images presented on Datasets A, B and C.

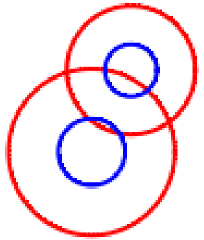
DTW

Dataset	Number of Jets	Number of Images
D	1	83097
E	2	83097
F	3	83097
G	4	83097

Table 2: Number of jets and number of images presented on Datasets D, E, F and G

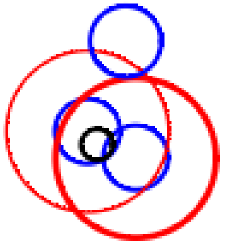
DHJW

- No MET or Muon charge information
- Constant Jets across the images



DTW

- MET and Muon charge information
- Variable number of Jets across the images



DHJW Images

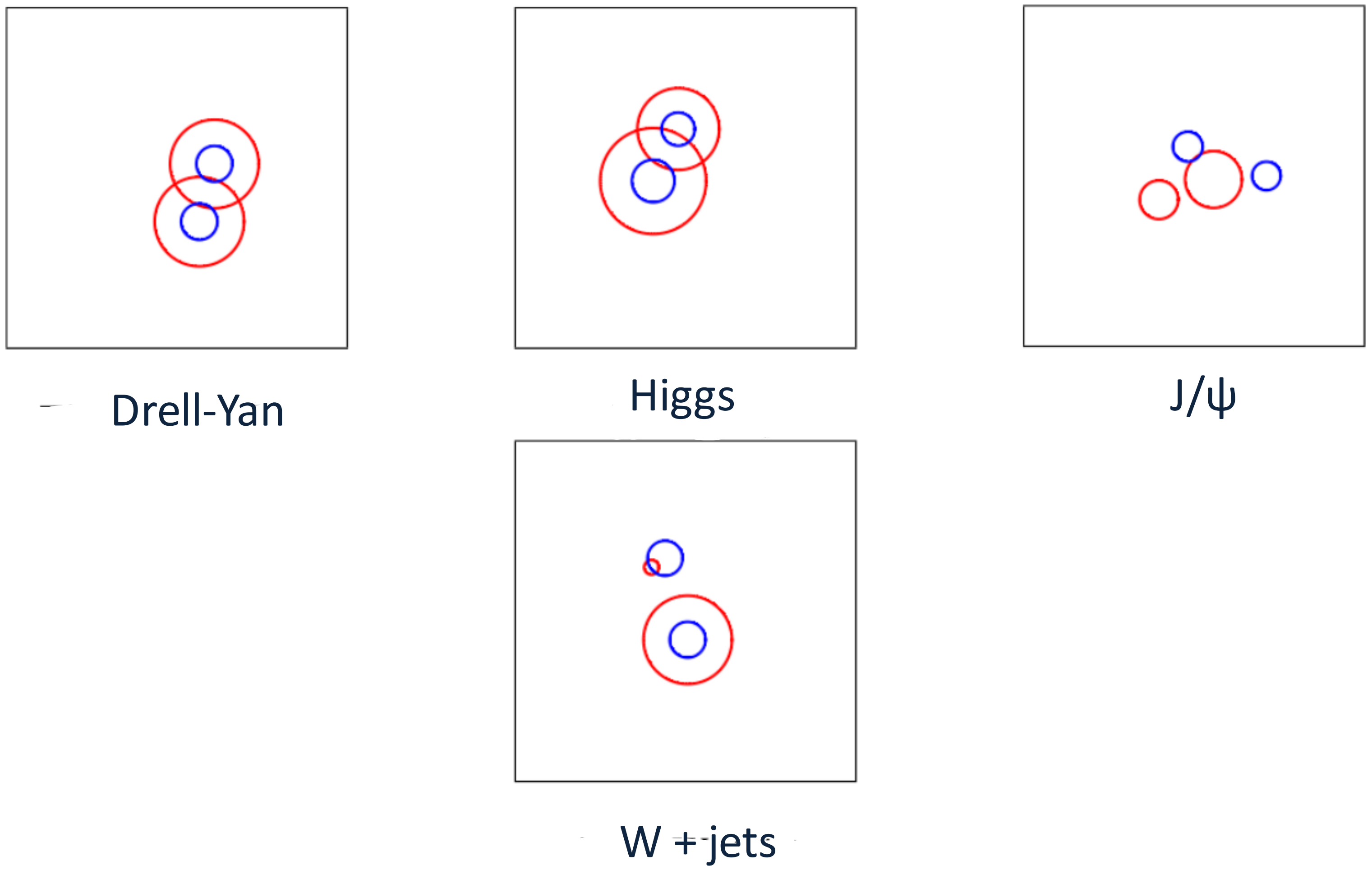
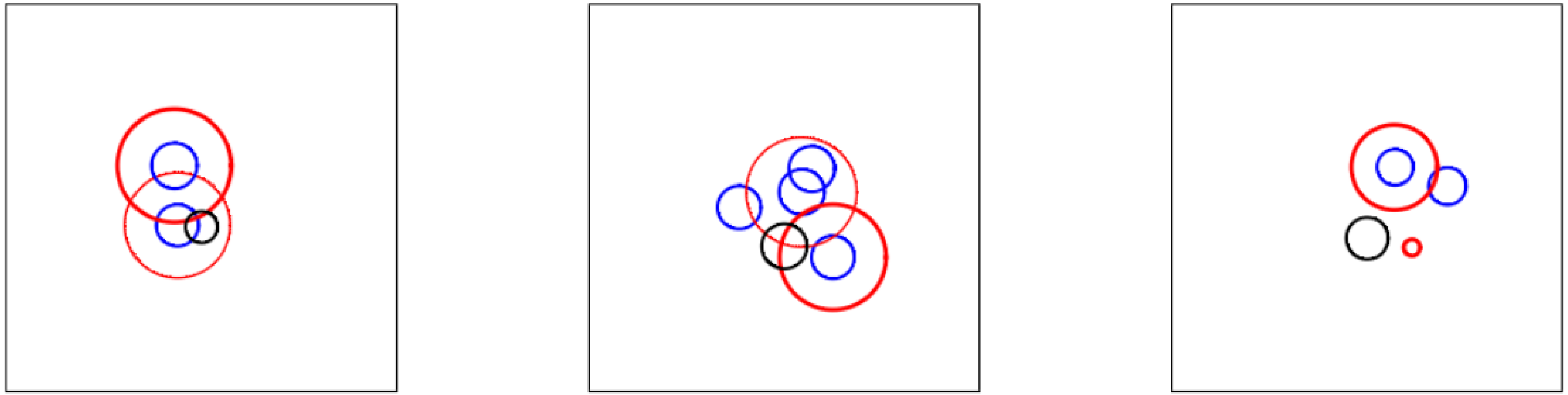


Fig 3: Example of images belonging to dataset C

DTW Images



Drell-Yan | Ttbar + jets | W + jets

Fig 3: Example of images belonging to dataset G

Types of Neural Network

- ResNet 50
- DenseNet
- InceptionV3
- MobileNet V2

Evaluating each model’s accuracy, loss and F1 score sets foreword the most suitable model when classifying high-energy particle collision outcomes.

Training process and Evaluation Metrics

TICec

- 40 epochs with early stopping.
- Adam optimization algorithm and Softmax Loss
- The training and testing of the CNN models were conducted on Google Colaboratory using A100 GPU hardware accelerators.
- All the code can be found in https://github.com/jose8af/cnn-hep-thesis [4]

ticec.cedia.edu.ec

Results

DHJW

Dataset	Test Acc	Test Loss
A	0.7209	0.7463
B	0.7322	0.6944
C	0.7956	0.5148

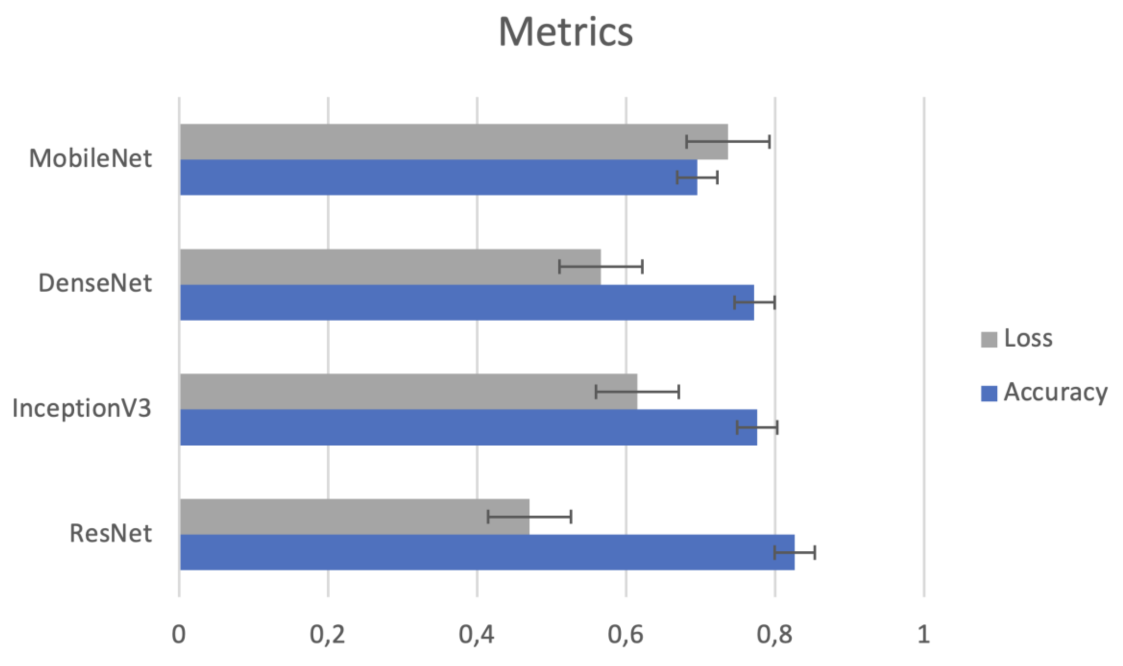
Table 3: Accuracy and loss value of the DHJW datasets

DTW

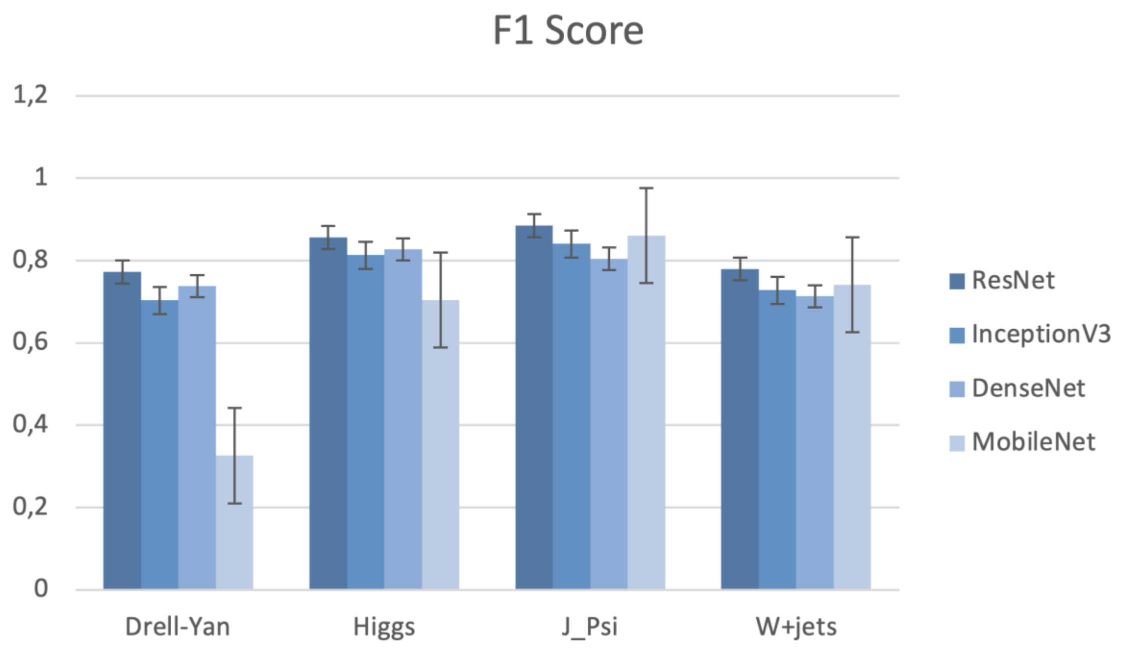
Dataset	Test Acc	Test Loss
D	0.8012	0.4508
E	0.8208	0.4185
F	0.8355	0.3875
G	0.8416	0.3681

Table 4: Accuracy and loss value of the DTW datasets

DHJW

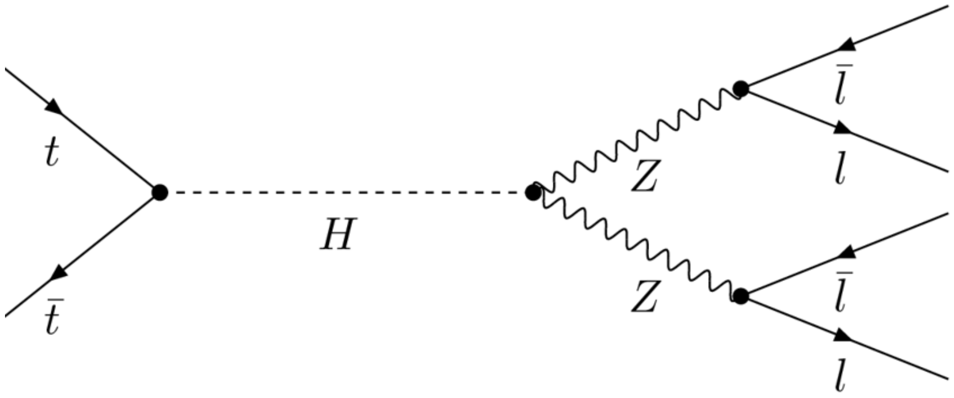


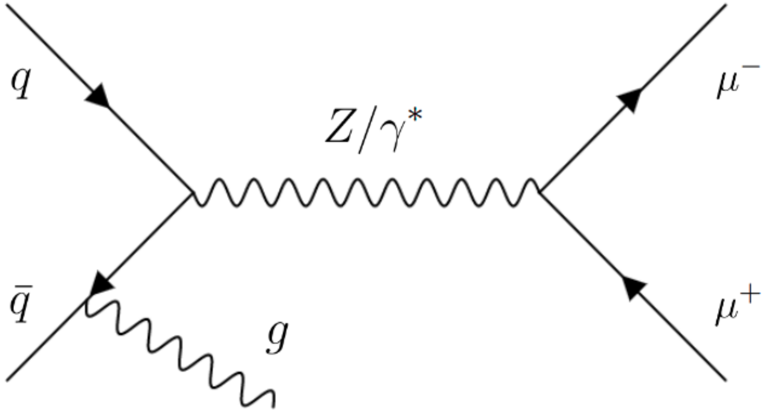
Metrics
MobileNet
DenseNet
InceptionV3
ResNet
0
0,2
0,4
0,6
0,8
1
Loss
Accuracy



F1 Score
1,2
1
0,8
0,6
0,4
0,2
0
Drell-Yan
Higgs
J_Psi
W+jets
ResNet
InceptionV3
DenseNet
MobileNet

DHJW





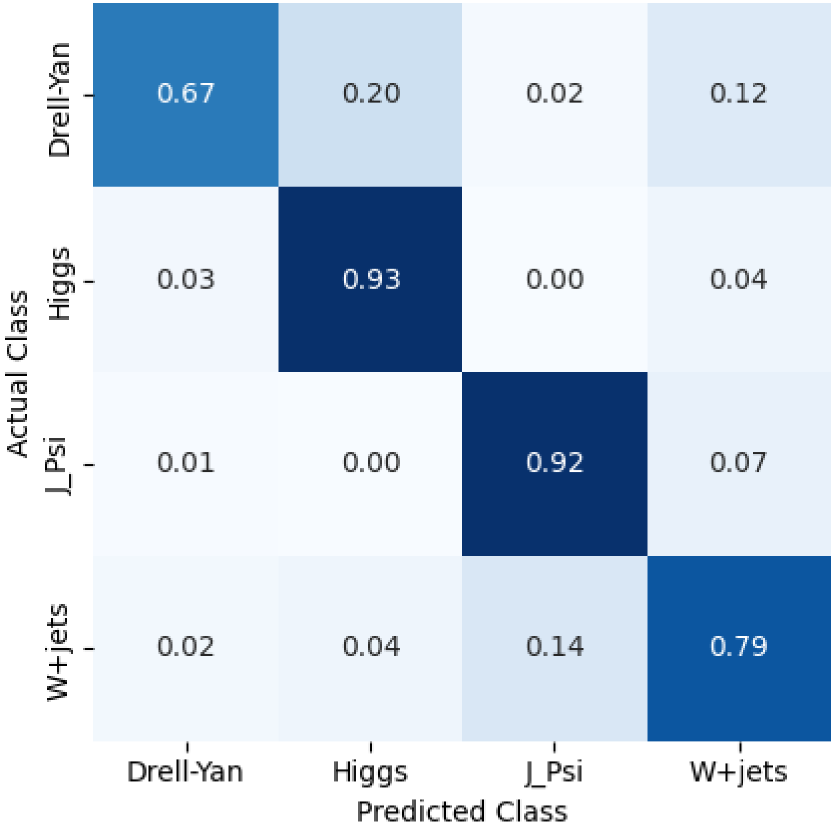
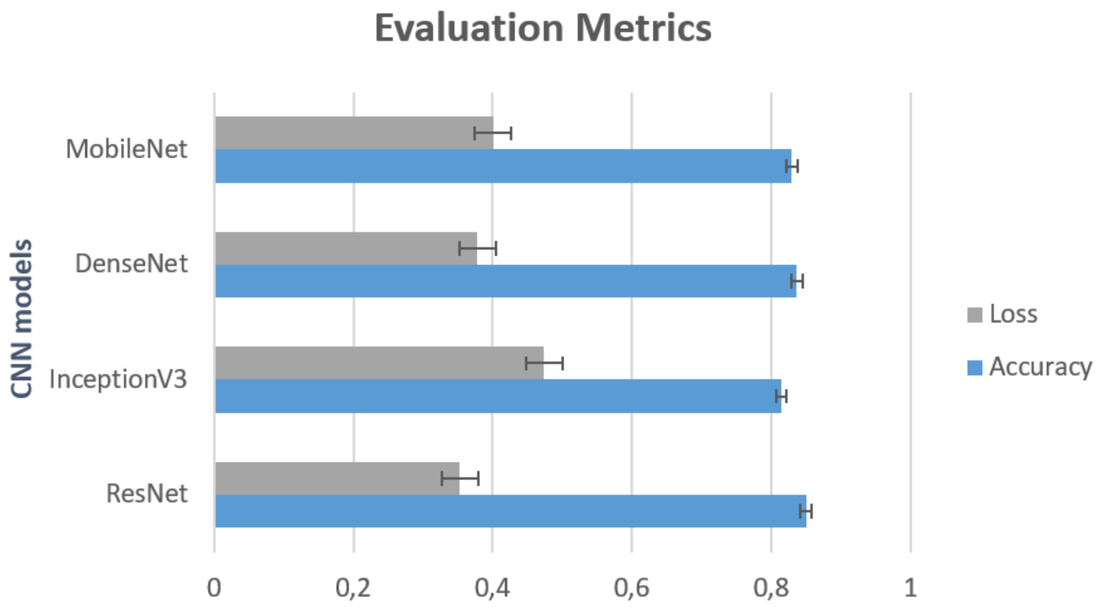
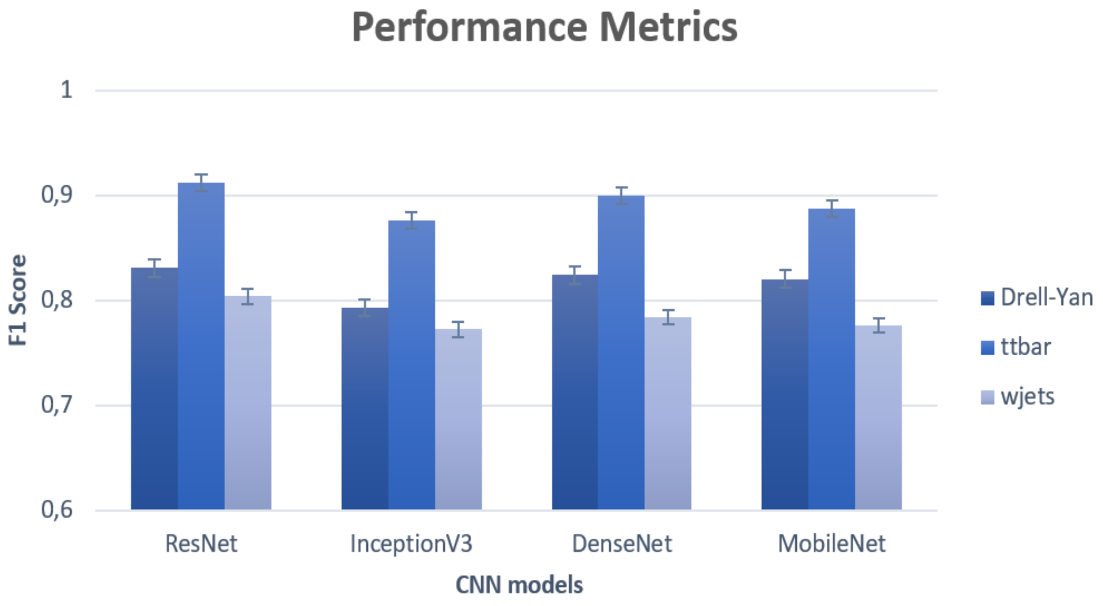


Fig 4: Confusion matrices of the best model

DTW

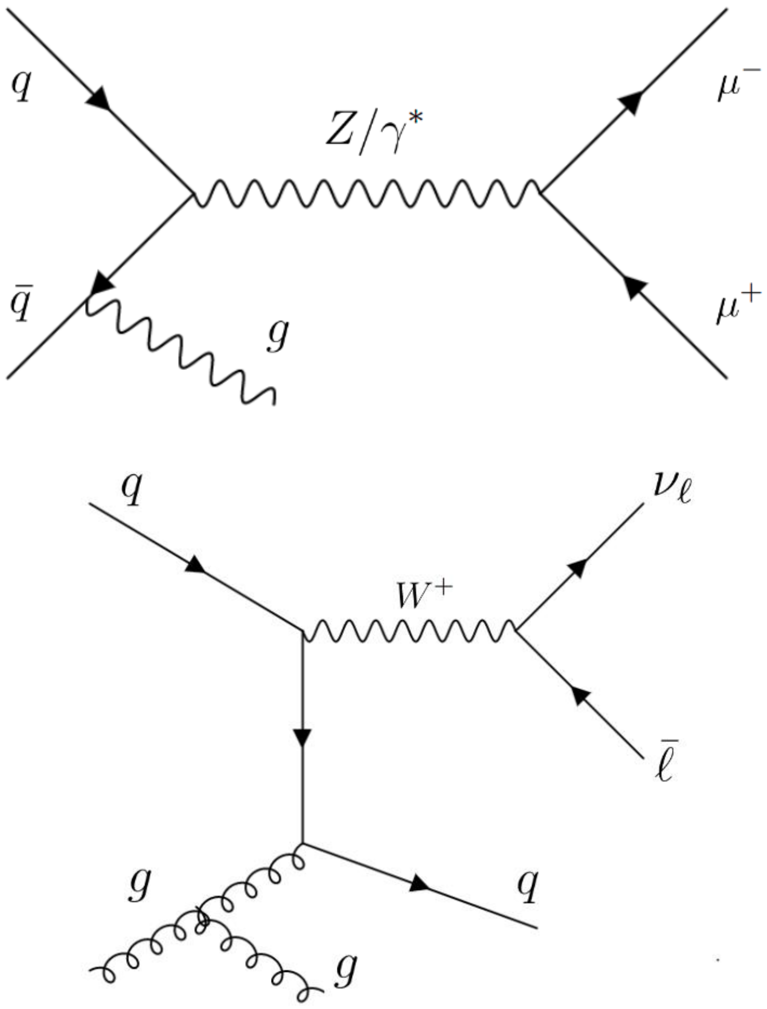


Evaluation Metrics
MobileNet
DenseNet
InceptionV3
ResNet
CNN models
0
0,2
0,4
0,6
0,8
1
Loss
Accuracy



Performance Metrics
F1 Score
1
0,9
0,8
0,7
0,6
ResNet
InceptionV3
DenseNet
MobileNet
CNN models
Drell-Yan
ttbar
wjets

DTW



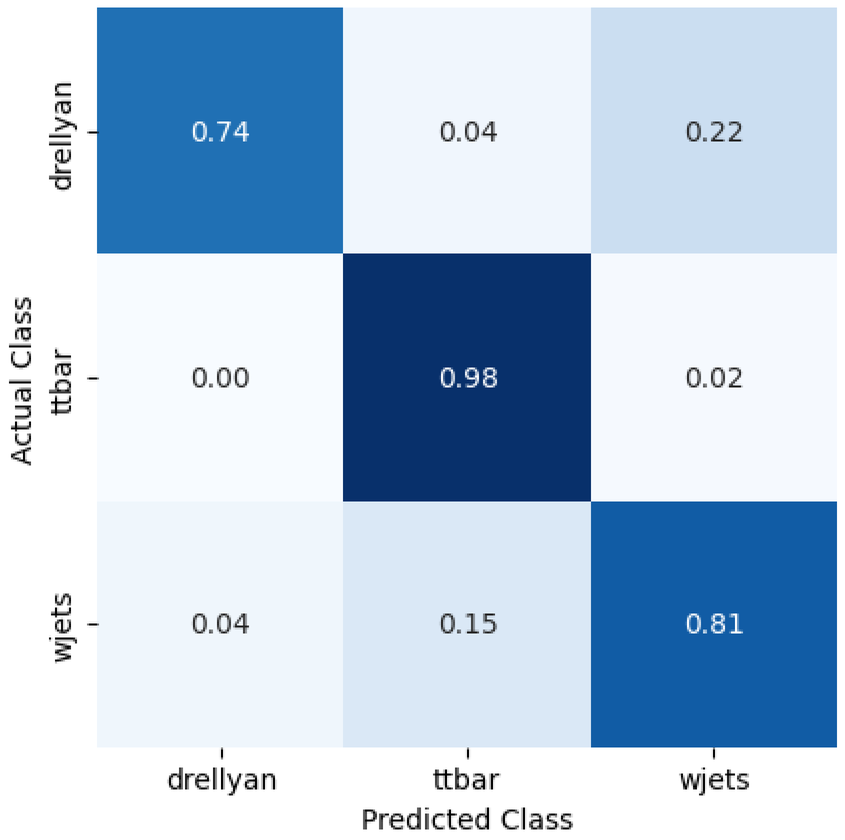


Fig 5: Confusion matrices of the best model

Application of the model in real collision data

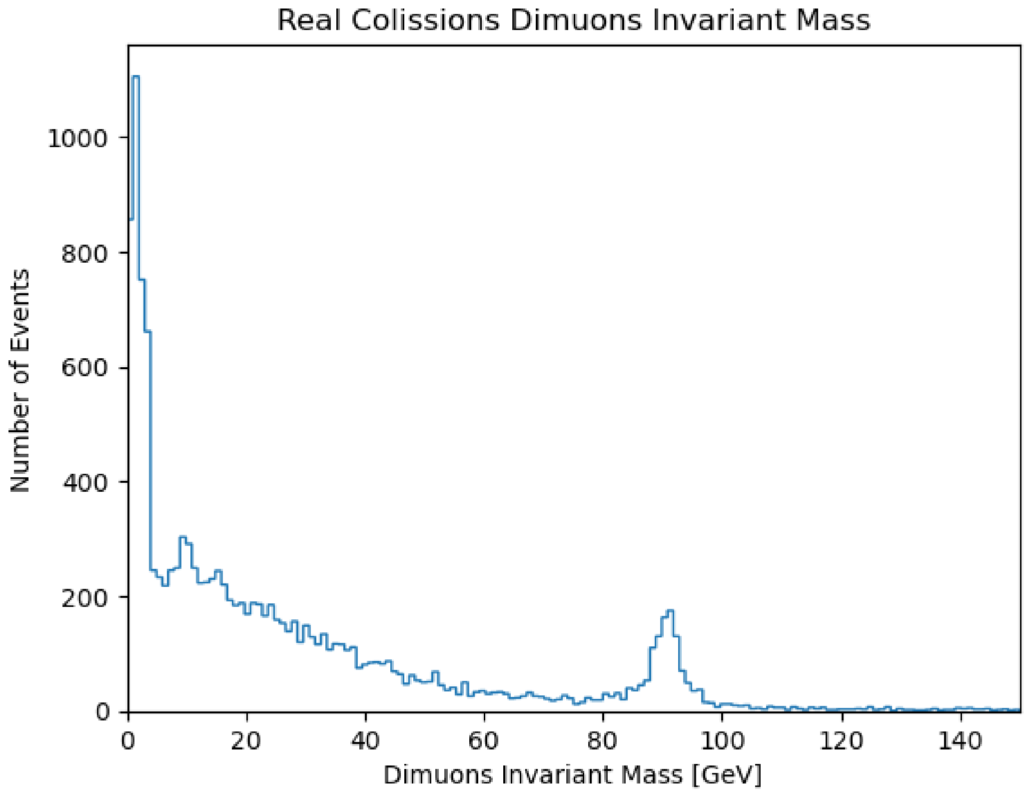
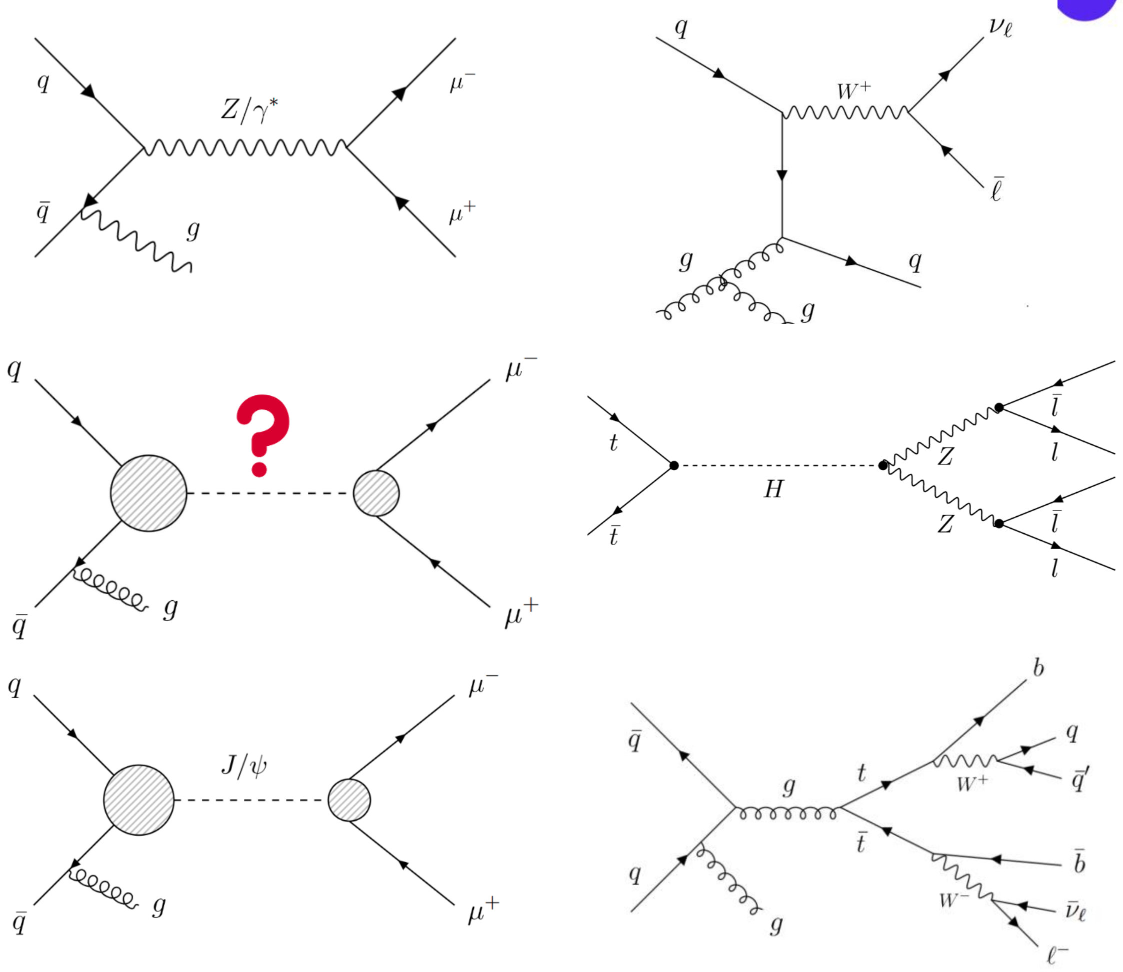
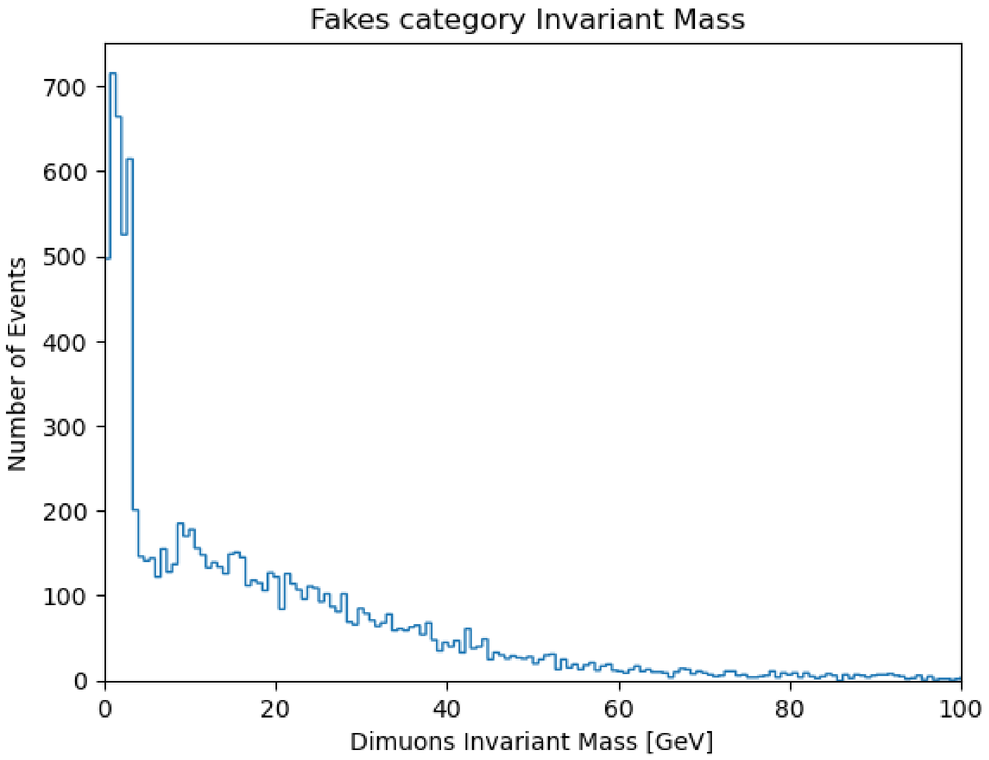


Fig 6: Dimuon invariant mass of real collision data





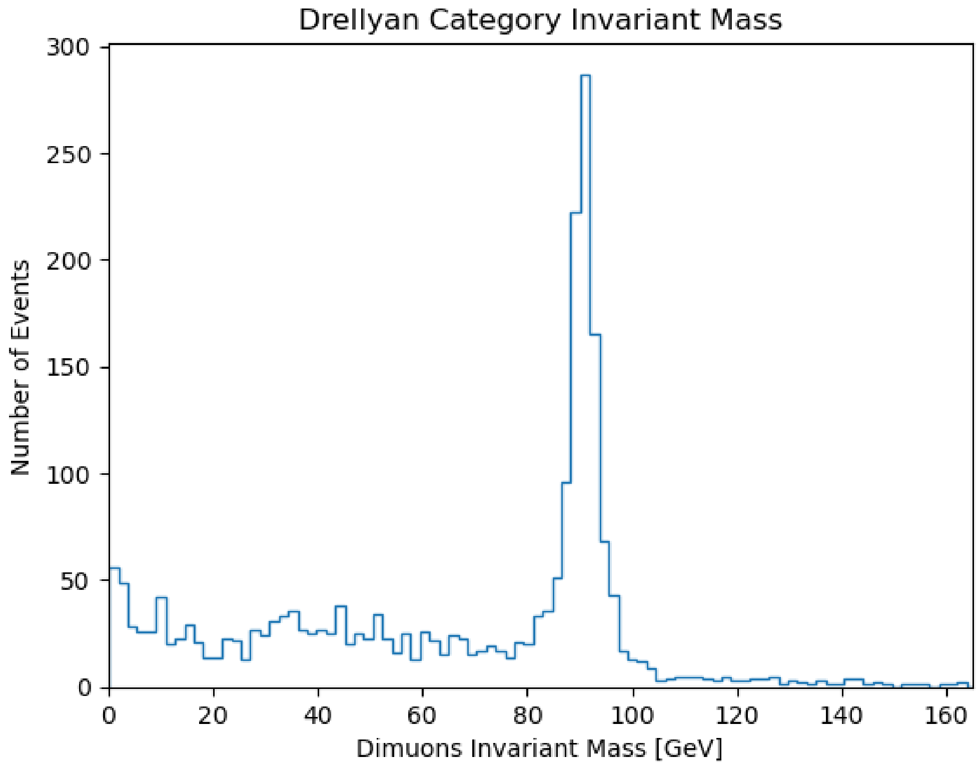


Fig 7: Dimuon invariant mass of the respective predictions

Conclusions

- Promising results were obtained for both DHJW and DTW scenarios with accuracies greater than 80 % in both cases
- ResNet50 has demonstrated to be the best CNN model among all the other popular options
- The DTW model could differentiate the Z boson resonance from a collection of real collision data

Muchas gracias! / Thank you very much!

References

[1] CERN — LHC images gallery. (2023, March 23). Disponible en : https://home.cern/resources/image/accelerators/lhc-images-gallery

[2] Izaak Neutelings.

Cms coordinate system. disponible en: https://tikz.net/axis3d cms/

[3] C. F. Madrazo, I. H. Cacha, L. L. Iglesias, and J. M. de Lucas, “Application of a convolutional neural network for image classification to the analysis of collisions in high energy physics,” CoRR, vol. abs/1708.07034, 2017.

[4] “Ochoa, J. D. CNN-hep-thesis: Undergrad Thesis. using a CNN to classify different HEP processes. GitHub. Retrieved May 5, 2023, from https://github.com/jose8af/cnn-hep-thesis,”