Tau identification exploiting deep learning techniques

Andrea Cardini[a],* on behalf of the CMS collaboration

[a]*Deutsches Elektronen-Synchrotron,*
Notkestraße 85, Hamburg, Germany

E-mail: andrea.cardini@cern.ch

The recently deployed DeepTau algorithm for the discrimination of taus from light flavor quark or gluon induced jets, electrons, or muons is an ideal example for the exploitation of modern deep learning neural network techniques. With the current algorithm a suppression of misidentification rates by factors of two and more have been achieved for the same identification efficiency for taus compared to the MVA identification algorithms used for the LHC Run-1, leading to significant performance gains for many tau related analyses. The algorithm and its performance will be discussed.

*Speaker

1. Introduction: tau physics

Tau leptons are the heaviest leptons in the standard model (SM) of particle physics. Their study is crucial in several measurements: tests of electroweak interaction and lepton flavor universality, study of the production [1] and CP properties [2] of the Higgs boson via its Yukawa coupling to fermions, together with several searches for beyond standard model signatures [3]. This talk focuses on the reconstruction of tau leptons in the Compact Muon Solenoid (CMS) experiment [4] at the Large Hadron Collider (LHC) [5]. After an introduction on the properties of the tau lepton and its experimental signatures, the algorithms deployed in CMS for the reconstruction of tau leptons will be described. Particular focus will be given to the DeepTau neural network based algorithm used to reduce the misidentification rate of jets and light leptons as hadronically decaying tau leptons.

2. The tau lepton

With a mass of $m_\tau = (1776.86 \pm 0.12)$ MeV [6] the tau lepton is the heaviest known lepton and the only one which can decay both hadronically and leptonically. Its average lifetime is $\mathcal{T}_\tau \simeq (2.903 \pm 0.005) \times 10^{-13}$ s, which for a τ lepton of 30 GeV corresponds to a decay length of:

$$\lambda_\tau = c * \mathcal{T}_\tau * \gamma\beta \simeq (3 \times 10^{11} \text{ mm/s}) * (2.9 \times 10^{-13} \text{ s}) * (30 \text{ GeV}/1.78 \text{ GeV}) \simeq 1.5 \text{ mm}$$

with c speed of light and $\beta\gamma = p/m$, with momentum (p) and rest mass (m) written in natural units. The innermost layer of the silicon tracker in CMS is at a distance of ~3 cm from the beamline. This means that the fraction of τ leptons produced close to the beam line (also referred to as *prompt*) that decay after reaching the innermost layer of the detector is negligible. Tau leptons are therefore reconstructed based on their decay products. Fig. 1 summarizes the main tau decay channels and their corresponding measured branching fractions.
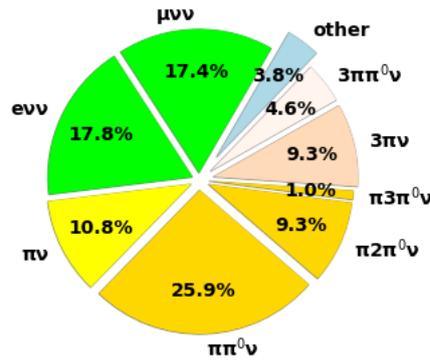

Figure 1: Pie chart representing the tau decay channels with relative branching fractions.

Tau decays always involve at least one charged particle, traditionally referred to as *prong*. In particular, leptonic decays are all classified as 1 prong decays, while hadronic decays are mostly mediated by mesonic resonances leading to final states with 1 or 3 prongs. 5 prong decays are also allowed and represent 0.1% of the total decays.

3. Hadronically decaying tau lepton reconstruction

Hadronically decaying tau leptons (τ_h) appear in the detector as well collimated jets. Their decay products are identified using the PF algorithm [7], and assigned to a tau decay by the Hadron-plus-strip (HPS) algorithm [8]. A track in the inner tracker matching a deposit in the hadronic calorimeter is assigned to a *hadron* by the PF and HPS algorithms, while electrons, positrons and photons found in a region narrow in pseudorapidity (η) and elongated in the azymuthal direction (ϕ), i.e. matching a π^0 decay signature, are assigned to a *strip*. If these objects are sufficiently close to each other and their invariant mass matches the one corresponding to a tau decay channel, they are assigned to a tau decay. Four decay channels are reconstructed by the HPS algorithm and their signatures are shown in Fig. 2.
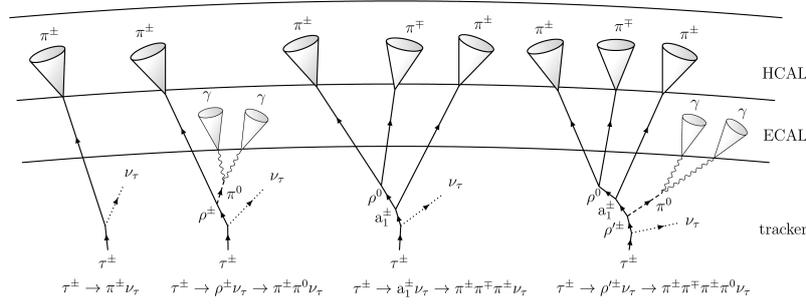
Figure 2: Hadronic decay channels reconstructed by the HPS algorithm. Moving left to right the decay channels are: 1 prong, 1 prong + π^0s, 3 prong and 3 prong + π^0.

Several objects can be misidentified as hadronically decaying tau leptons by the HPS algorithm:

- jets: a highly collimated quark or gluon jet can be mistaken for any tau decay,
- muons: can produce a signature similar to a 1 prong tau,
- electrons: can emit photons via bremsstrahlung radiation and mimic the ρ decay.

The misidentification rate by jets and light leptons is reduced by the use of a neural network based algorithm: DeepTau [9].

4. DeepTau Identification

The DeepTau identification is based on a multiclass convolutional NN used to reduce the misidentification of jets, muons and electrons as hadronically decaying taus. It takes as inputs low and high level features associated to the τ_h candidate and produces 4 output scores, one for each class: genuine taus (y_τ), jets (y_j), muons (y_μ), and electrons (y_e). The basic structure of the NN is shown in Fig. 4.

4.1 Input features

The *low level* features are the tracks and energy deposits of the tau decay products and other PF candidates in the isolation cone, i.e. in the vicinity of the τ_h candidate. The CMS detector is divided in cells of $\eta \times \phi$, and the properties of the leading PF candidate found in each cell are
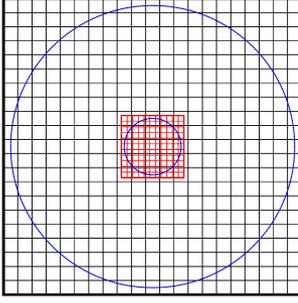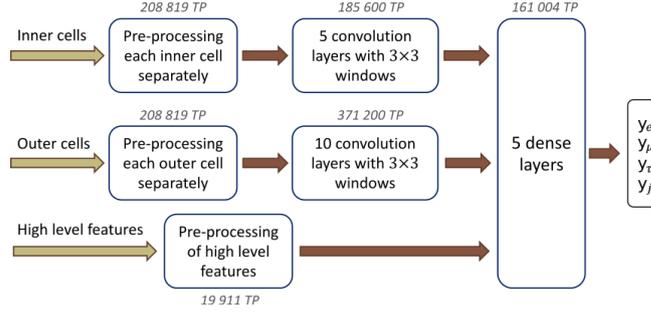
Figure 3: NN input grid.

Figure 4: Structure of the DeepTau NN.

taken as inputs for the NN. A set of 21×21 cells arranged as a square around the leading tau decay product direction of flight are defined with a granularity of $\eta \times \phi = 0.05 \times 0.05$, and overlap with a set of 11×11 cells with a granularity of $\eta \times \phi = 0.02 \times 0.02$ as shown in Fig. 3. The cells with higher granularity are named *inner cells*, while the others are named *outer cells*. The two granularities allow to study more precisely the topology of the event in the region closer to the leading tau decay product, allowing for good discrimination between tau leptons and misidentified leptons and jets, while keeping the number of features manageable from the computational point of view. Other input features are referred to as *high level* as they are related to either the tau candidate, like its transverse momenta, HPS-decay mode, η, ϕ and other properties, or the recorded event in its entirety like the average energy of the event. This results in a number of input features of the order of 100 thousands ($O(100k)$).

4.2 Convolutional layers and training

Each group of input features[1] is pre-processed by 3 convolutional layers of window size 1×1. The low level features are then further processed by convolutional layers of window size 3×3 as shown in Fig. 4. All features are then concatenated and processed via 5 fully connected dense layers leading to the four output nodes. The training is performed using the NAdam algorithm [10] with $O(1.5M)$ trainable parameters. It is a supervised training which uses samples of simulated Drell-Yan (DY) or high mass Z interactions for genuine hadronically decaying tau leptons and misidentified objects, together with simulated processes like $t\bar{t}$, W+Jets, and QCD multijet production to account for different jet topologies.
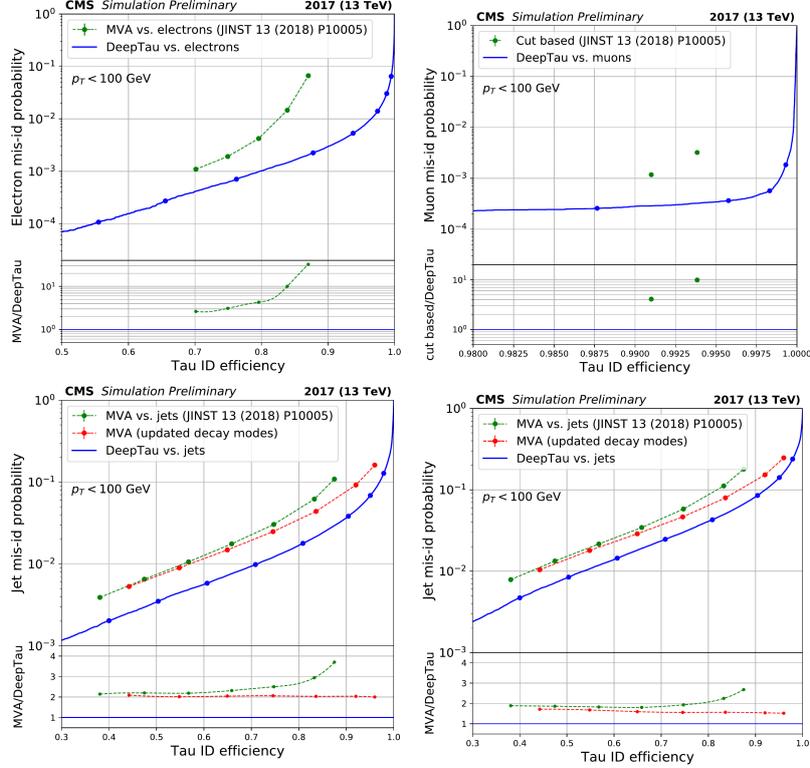
4.3 Classifiers

From the output scores, three different classifiers can be derived:

$$P_{\tau \text{ vs } obj.} = \frac{y_\tau}{y_\tau + y_{obj.}}$$

with $obj. = j, e, \mu$. Fig. 5 shows the misidentification rate as a function of the efficiency for the selection of genuine hadronically decaying tau leptons for the three classifiers. The DeepTau classifiers are compared to the older MVA-based classifiers used in CMS before the deployment of DeepTau. For the same efficiency for the selection of genuine hadronically decaying tau leptons,

---

[1]Low level from the inner and outer cells separately, and high level features.

4

the DeepTau classifiers present a noticeable reduction of the misidentification rate. For the against-electron and muon classifiers, the misidentification rate is reduced by at least a factor 3, while the against-jet classifier achieves a reduction of $\sim 20\%$.

Figure 5: Misidentification rate as a function of the efficiency for the selection of hadronically decaying tau leptons. The top row shows the against-electron (left) and against-muon (right) classifiers validated on $Z \rightarrow ee$ and $Z \rightarrow \mu\mu$ samples, while the bottom row shows the against-jet classifiers validated on two different jet topologies: $t\bar{t}$ (left) and W+Jets (right) [9].
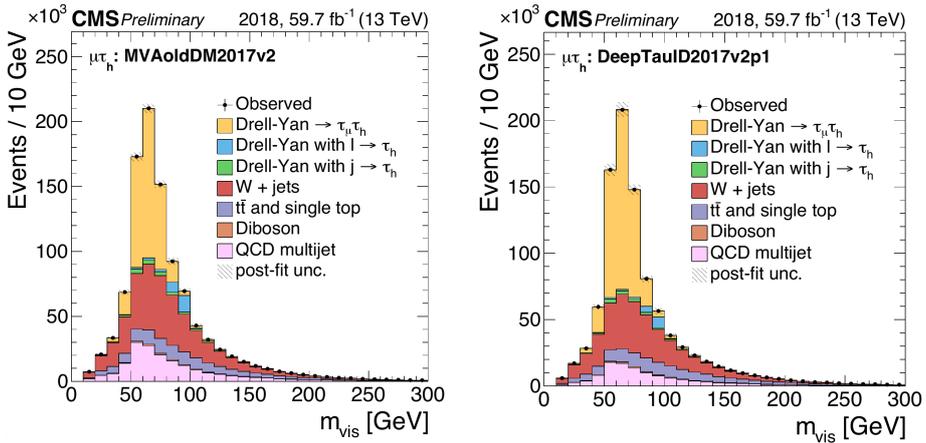
Figure 6: Visible mass distribution for $\tau_\mu \tau_h$ channel using on the left the old MVA-based classifiers and on the right the DeepTau ones [9].

5. Conclusions

The DeepTau NN based identification algorithm improves the efficiency in identifying genuine hadronically decaying tau leptons by $\sim 20\%$ while reducing the misidentification rate from light leptons and jets by $\sim 23\%$. This already allowed several analyses to improve their sensitivities. As an example, Fig. 6 shows the improvement in the reconstruction of di-tau events where a tau lepton decays to a muon (τ_μ) and the other decays hadronically. The visible mass distribution of the $\tau_\mu \tau_h$ pair obtained with the DeepTau classifiers shows a greater contribution from the DY$\to \tau_\mu \tau_h$ process and reduced contributions from other processes, compared to the distribution obtained with the old MVA-based classifiers. Simulation agreement with data is maintained through the use of correction scale factors of the order of $\sim 10\%$.

References

[1] CMS Collaboration, *Measurement of Higgs boson production in the decay channel with a pair of τ leptons*, CMS-PAS-HIG-19-010, cds.cern.ch/record/2725590.

[2] CMS Collaboration, *Analysis of the CP structure of the Yukawa coupling between the Higgs boson and τ leptons in proton-proton collisions at $\sqrt{s}$ = 13 TeV*, CMS-PAS-HIG-20-006, cds.cern.ch/record/2725571.

[3] CMS Collaboration, *Search for additional neutral MSSM Higgs bosons in the $\tau\tau$ final state in proton-proton collisions at $\sqrt{s}$ = 13 TeV*, JHEP **09** (2018), 007, doi:10.1007/JHEP09(2018)007.

[4] CMS Collaboration, *The CMS Experiment at the CERN LHC*, JINST **3** (2008), S08004 doi:10.1088/1748-0221/3/08/S08004

[5] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008), S08001 doi:10.1088/1748-0221/3/08/S08001

[6] R. M. Barnett *et al.*, Particle Data Group, *Review of particle physics. Particle Data Group*, Phys. Rev. D **54** (1996), 1-720 doi:10.1103/PhysRevD.54.1

[7] CMS Collaboration, *Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET*, CMS-PAS-PFT-09-001, cds.cern.ch/record/1194487.

[8] CMS Collaboration, *Performance of reconstruction and identification of tau leptons in their decays to hadrons and tau neutrino in LHC Run-2*, CMS-PAS-TAU-16-002, cds.cern.ch/record/2196972.

[9] CMS Collaboration, *Performance of the DeepTau algorithm for the discrimination of taus against jets, electron, and muons*, CMS-DP-2019-033, cds.cern.ch/record/2694158.

[10] T. Dozat, *Incorporating Nesterov Momentum into Adam*, in Proceedings of *4th International Conference on Learning Representations*, Workshop Track, 2016.