Improving Multi-Higgs sensitivity in the hadronic final state using ML

Daniel Diaz, Javier Duarte, Sanmay Ganguly, Brian Sheldon

14/07/2023
Di-Higgs production at higher energies

Constraining $\kappa_\lambda$ is a major physics program of LHC and future colliders.

We aim to optimize a generic selection strategy for multi-Higgs final state in all-hadronic channel.

arXiv: 1910.00012
Resonant Higgs-pair: ATLAS

\[ X_{HH} = \sqrt{\left( \frac{m(H_1) - 120 \text{ GeV}}{0.1 \times m(H_1)} \right)^2 + \left( \frac{m(H_2) - 110 \text{ GeV}}{0.1 \times m(H_2)} \right)^2} \]

\[ R_{HH}^{VR} = \sqrt{(m(H_1) - 1.03 \times 120 \text{ GeV})^2 + (m(H_2) - 1.03 \times 110 \text{ GeV})^2} < 30 \text{ GeV} \]

\[ R_{HH}^{CR} = \sqrt{(m(H_1) - 1.05 \times 120 \text{ GeV})^2 + (m(H_2) - 1.05 \times 110 \text{ GeV})^2} < 45 \text{ GeV} \]

SR: \( X_{HH} < 1.6 \)

The pairing is chosen using a BDT trained by LightGBM (gradient BDT).

ATLAS
\( \sqrt{s} = 13 \text{ TeV}, 126 \text{ fb}^{-1} \)
Data (2b)
Resolved channel

Phys. Rev. D 105 (2022) 092002
Y+H interpretations of 4b : CMS

\[ H \rightarrow b\bar{b} \] tagging is done using ParticleNet (GNN network)

\[ \frac{P(R \rightarrow bb)}{P(R \rightarrow bb) + P(QCD)} \]

Same partucleNet tagger was used for H->cc analysis.

Object identification using NN is now the standard prescription.
Event classification using GNN

For boosted di-Higgs production we look for two Ak-08 jets with track subjets

pic credit: https://francis.naukas.com/2014/08/22/el-campo-de-higgs/
Event as a graph

Locate all the R=0.8 PF jets and 0.4 track jets in the \( \eta, \phi \) plane.
Connect the k-NN neighbour through edges.
For each nodes: assign 4-vector + two and one subjettness observables.
Use this graph representation for the events to be fed in GNN.
In a graph, each node can “learn” about the state of neighboring node through message passing operation

\[
(x')_{i}^{l+1} = \max_{j \in \mathcal{N}(i)} \Theta_x(x_j^l - x_i^l) + \Phi_x(x_i^l)
\]

\[
(e')_{i}^{l+1} = \text{mean}_{j \in \mathcal{N}(i)} \Theta_e(e_j^l - e_i^l) + \Phi_e(e_i^l)
\]
The graph network

\[
(x')^l_{i}^{l+1} = \max_{j \in \mathcal{N}(i)} \Theta_x(x_j^l - x_i^l) + \Phi_x(x_i^l)
\]

\[
(e')^l_{i}^{l+1} = \text{mean}_{j \in \mathcal{N}(i)} \Theta_e(e_j^l - e_i^l) + \Phi_e(e_i^l)
\]

Forms a K-NN graph and performs the message passing

Learned "coordinates"

Learned "energy"
The graph network

\[(\vec{x}_1, E_1)\]
\[(\vec{x}_2, E_2)\]
\[(\vec{x}_3, E_3)\]
\[\vdots\]
\[(\vec{x}_N, E_N)\]

\[EC_{L1}(d_{in} = 3, d_{out} = M)\]

Forms a K-NN graph and performs the message passing

\[(z_1^M, E_{11}^L)\]
\[(z_2^M, E_{21}^L)\]
\[(z_3^M, E_{31}^L)\]
\[\vdots\]
\[(z_N^M, E_{N1}^L)\]

After \(p\) message passing layers, the \(q\)-th node has following energy representation:

\[\begin{pmatrix}
E_q, E_{q1}^L, E_{q2}^L, \ldots, E_{qM}^L
\end{pmatrix}\]

\[\xrightarrow{\text{MLP}}\]

\[\left[ p_{\text{sig}}, p_{\text{bkg}} \right]\]
Some distributions of the features

arXiv:2203.07353
Training outputs

The preliminary NN has 3X times background rejection for the same efficiency around cut-based WP.
An attempt with LE GNN

\[
m^l_{ij} = \phi_e \left( h^l_i, h^l_j, \psi(\|x^l_i - x^l_j\|^2), \psi(\langle x^l_i, x^l_j \rangle) \right)
\]

\[
x^{l+1}_i = x^l_i + c \sum_{j \in [N]} \phi_x(m^l_{ij}) \cdot x^l_j
\]

\[
h^{l+1}_i = h^l_i + \phi_h(h^l_i, \sum_{j \in [N]} w_{ij} m^l_{ij})
\]

Lorentz Group Equivariant Block (L Geb)

LorentzNet
Both the MP-networks have similar event classification efficiency.
Pre & post selection distributions

- Large-R jet having two subjets.
- Jet1 $p_T > 400$ GeV, Jet2 $p_T > 300$ GeV
- $|\text{Jet } \eta| < 2.5$

![Signal Histogram](image1)

![Background Histogram](image2)
Pre & post selection distributions

- Large-R jet having two subjets.
- Jet1 $p_T > 400$ GeV, Jet2 $p_T > 300$ GeV
- $|\text{Jet} \ \text{eta}| < 2.5$

**Background Histogram GNN>0.75**

**Signal Histogram GNN>0.75**
Pre & post selection distributions

- Large-R jet having two subjets.
- Jet1 $p_T > 400$ GeV, Jet2 $p_T > 300$ GeV
- $|\text{Jet \eta}| < 2.5$

Larger stat fluctuation on QCD tail due to rejected events.

![Background Histogram GNN$>0.98$](image1)

![Signal Histogram GNN$>0.98$](image2)
Pre & post selection distributions

![Image of a plot showing signal significance vs. $kappa_{\lambda}$ with different selection methods indicated: Cut-based, GNN Pre-Selection, and GNN WP=95. The plot includes data points and curves for each method.]
Take away

- For multi-b final states event level classifiers are capable of increasing sensitivity.
- These preliminary studies (for FCC-hh) are generalizable across hh or hhh searches.
- Compared two different GNN models: probably a general GNN will do the required job.
- The individual tagging score and pairing scores should improve the sensitivity.

THANK YOU!!