Convolutional Neural Network for Track Seed Filtering at the CMS HLT

Adriano Di Florio, Felice Pantaleo, Maurizio Pierini

On Behalf of the CMS Collaboration
The Compact Muon Solenoid (CMS) is a general purpose detector designed for the precision measurement of leptons, photons, and jets, among other physics objects, in proton–proton as well as heavy ion collisions at the CERN LHC.

**ONLY ABOUT 1000 EVENTS/SEC [@ ~ 1 MB/EVENT] CAN BE RECORDED ON DISK**

**L1 TRIGGER**
- 40 MHz in / 100 KHz out
- ~500 KB / event
- Processing time: ~10 μs
- Based on coarse local reconstructions
- FPGAs / Hardware implemented

**HIGH LEVEL TRIGGER (HLT)**
- 100 KHz in / 1 KHz out
- ~500 KB / event
- Processing time: ~30 ms
- Based on simplified global reconstructions
- Software implemented on CPUs
In CMS, the tracking algorithm consists of an iterative procedure, in which tracks are reconstructed according to progressively looser quality criteria starting from hits on the silicon tracker detector.
In CMS, the tracking algorithm consists of an iterative procedure, in which tracks are reconstructed according to progressively looser quality criteria starting from hits on the silicon tracker detector.

**ONLINE RECONSTRUCTION (HLT)**

Practically the same reconstruction procedure as the one run offline. It has to undergo stringent time limits: $O(100)$ ms. It is based on pixel–only reconstruction.
What's next?

LHC / HL-LHC Plan

LHC

Run 1 | Run 2 | Run 3 | HL-LHC

LS1  | 13 TeV | 13.5-14 TeV | LS2  | 14 TeV | LS3  | 14 TeV


experiment beam pipes
splice consolidation
button collimators
R2E project

eYETS

injector upgrade
cryo Point 4
DS collimation
P2–P7(11 T dip.)
Civil Eng. P1-P5

experiment upgrade
phase 1

HL-LHC installation

experiment upgrade phase 2

7 TeV 8 TeV

30 fb⁻¹ 150 fb⁻¹ 300 fb⁻¹ 3000 fb⁻¹

75% nominal luminosity
nominal luminosity
cryolimit interaction regions
radiation damage
2 x nominal luminosity
integrated luminosity

2nd IML Workshop 9–12 April 2018 – CERN – Geneva
What’s next?

INSTANTANEOUS LUMINOSITY & SIMULTANEOUS COLLISIONS (PILE-UP) INCREASE

Ultimate goals

\[ \mathcal{L} = 5 \cdot 10^{34} \text{cm}^2 \text{s}^{-1} \]

\[ <PU> \sim 200 \]

The already complex online and offline track reconstruction has to deal not only with a much more crowded environment but also with data coming from a more complex detector.

INCREASED DETECTOR COMPLEXITY (SINCE MARCH 2017)

10 layers (6 endcap + 4 barrel)
Doublets Generation

**Doublet seeds generation:** bottleneck due to huge combinatorial background.

For a single $t\bar{t}$ at $\sqrt{s} = 13\text{TeV}$ with $\langle PU \rangle = 35$ simulated event: $O(10^5)$ doublets produced with fake ratio $\sim O(100)$ corresponding to $O(1000)$ true doublets.

But doublet selection is based only on geometrical compatibility checks.
There is some more information...
Each doublet is built from a couple of hits on the silicon pixel tracker detector. Each hit is not simply a point on the detector but it is a collection of pixels (2D) on or off. Each pixel is associated with an ADC (16 bit levels) level proportional to the charge deposited by a particle. Each hit is then a 2D pixel image centered on the center of charge. (15x15)

Typical pattern recognition problem (true/fake classification): suitable for a Convolutional Neural Network approach.
Convolutional Neural Networks are a specialized kind of neural network for processing data that has a grid-like structure, such as 2D images. The building block of a CNNs is a layer that uses *discrete convolution* in place of general matrix multiplication.

**Pooling**: its function is to progressively reduce the spatial size of the representation.

**Fully connected**: Neurons in a fully connected layer have full connections to all activations in the previous layer, as seen in regular Neural Networks. Reduce input to a unique score: *softmax*.

In our use case the CNN acts as a **binary classifier** (*signal* or *background*) and reduce the whole picture to a single score corresponded to $p_{true}(x)$.
Generation of $t\bar{t}$ at $\sqrt{s} = 13$ TeV with $<PU> = 35$ simulated events (via PYTHIA integrated in CMS reconstruction software, CMSSW): $O(10^5)$ doublets produced with fake ratio $\sim O(100)$ equals to a $O(1000)$ true doublets.

Association RECO - MC
1. get list of all matched reconstructed tracks track hits
2. get list of all doublets produced
3. true doublets = doublets formed by hits from the same sim matched track

To each doublet we the associate
- two $15 \times 15$ images (one for inner and one for outer hit)
- set of local informations ($x, y, z$, charge, ...)
Data preprocessing

Make of use of the layer structure of the detector to extend each single doublet from two pictures to 20 channels, one per each layer (6 barrel + 4 endcap)

**Channels**: a common picture is usually a superimposition of different color levels. E.g. RGB levels.

E.g. A doublet on **barrel2** and **forward endcap3**

<table>
<thead>
<tr>
<th>Inner Hit</th>
<th>BPix1</th>
<th>BPix2</th>
<th>BPix3</th>
<th>BPix4</th>
<th>FPos1</th>
<th>FPos2</th>
<th>FPos3</th>
<th>FNeg1</th>
<th>FNeg2</th>
<th>FNeg3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Outer Hit</th>
<th>BPix1</th>
<th>BPix2</th>
<th>BPix3</th>
<th>BPix4</th>
<th>FPos1</th>
<th>FPos2</th>
<th>FPos3</th>
<th>FNeg1</th>
<th>FNeg2</th>
<th>FNeg3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

20 channel image as input
The model

- A single doublet is a 20 levels image. **Concatenates**:
  - **CNN architecture**: stack of convolutional layers (4) and max pooling (2)
  - **“DENSE” architecture**: dense layers (2) fed with the 1-dim reduced images + doublets infos (inX, inY, inZ, ...)  

- **Dropouts & early stopping** to prevent overfitting

- **Train & val** datasets balanced (0.5)
Results on 2.5 millions doublets training

Accuracy (training on 250k dataset)

\[ ACC = \frac{VP + VN}{P + N} > 0.90 \]

Cross Entropy (training on 250k dataset)

\[ H(p, q) = -\sum_x p(x) \log q(x). \]
Results on 2.5 millions doublets training

\[ ACC = \frac{VP + VN}{P + N} > 0.90 \]

\[ H(p, q) = -\sum_x p(x) \log q(x). \]

Accuracy (training on 250k dataset)

Cross Entropy (training on 250k dataset)

<table>
<thead>
<tr>
<th>Whole Dataset Training</th>
<th>@ Max Acc</th>
<th>Rej @ Eff</th>
<th>Eff @ Rej</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AUC</td>
<td>ACC</td>
<td>Efficiency</td>
</tr>
<tr>
<td>Test</td>
<td>0.968</td>
<td>0.90</td>
<td>0.96</td>
</tr>
<tr>
<td>Val</td>
<td>0.966</td>
<td>0.90</td>
<td>0.95</td>
</tr>
<tr>
<td>Train</td>
<td>0.969</td>
<td>0.91</td>
<td>0.96</td>
</tr>
</tbody>
</table>
Results on 2.5 millions **doublets training**

\[ ACC = \frac{VP + VN}{P + N} > 0.90 \]

\[ H(p, q) = -\sum_x p(x) \log q(x). \]

<table>
<thead>
<tr>
<th>Whole Dataset Training</th>
<th>@ Max Acc</th>
<th>Rej @ Eff</th>
<th>Eff @ Rej</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AUC</td>
<td>ACC</td>
<td>Efficiency</td>
</tr>
<tr>
<td>Test</td>
<td>0.968</td>
<td>0.90</td>
<td>0.96</td>
</tr>
<tr>
<td>Val</td>
<td>0.966</td>
<td>0.90</td>
<td>0.95</td>
</tr>
<tr>
<td>Train</td>
<td>0.969</td>
<td>0.91</td>
<td>0.96</td>
</tr>
</tbody>
</table>
Results - II

Results on 2.5 millions **doublets training**

\[ FVP = \frac{VP}{P} \quad \text{Sensitivity} \]

\[ FFP = \frac{FP}{P} \quad 1\text{-Specificity} \]

**Classifier output score** (training on whole dataset)

- True Positive Rate
- False positives rate

\begin{align*}
\text{Kolmogorov-Smirnov test score} & \quad 0.070 \\
(p\text{-valore} & \quad \sim 0.961)
\end{align*}

**ROC Curve** (training on whole dataset)

| 2nd IML Workshop | 9–12 April 2018 – CERN – Geneva | 12 |
Results - III

Further crosschecks

**ROC Curve (training on 250k sample)**

- **Colorfull** – 0.940
- **B&W** – 0.936

**ROC Curve (training on whole dataset)**

- **TTBar** – 0.966
- **Zee** – 0.952
- **QCD** – 0.950

Setting ADC levels to 0 or 1 only for turned off and on pixels. **Black and white model.**

Comparing with the **colorfull model** with ADC levels.

Testing with other **montecarlo PYTHIA** recipies.
Conclusions and perspectives

- CNN techniques for mitigating combinatorial explosion look very promising
- “Small” use case that shows how CNNs (or DNNs) can help handling tracking @ HEP
- Exploring the integration in the CMS reconstruction Framework
- Verification of the effect on the downstream track reconstruction ongoing
- Exploration of different hardware architecture for fast inference ongoing
- Input variables ranking
- Possible extension to PID @ silicon pixel detector (?)
THANK YOU

"I am putting myself to the fullest possible use, which is all I think that any conscious entity can ever hope to do"

HAL9000
The doublet features

450 ADC pixels [2x15x15 pads]

\[
\text{inPixLab} = ["inPix1","inPix2", \ldots, "inPix224","inPix225"]
\]

\[
\text{outPixLab} = ["outPix1","outPix2", \ldots,"outPix224","outPix225"]
\]

63 features defined for each doublet [true or fake] that may be used as additional features to the pixel pad

\[
\text{headLab} = ["run","evt","detSeqIn","detSeqOut","inX","inY","inZ","outX","outY","outZ","inPhi","inR","outPhi","outR","detCounterIn","detCounterOut","isBarrelIn","isBarrelOut","layerIn","ladderIn","moduleIn","sideIn","diskIn","panelln","bladeln","layerOut","ladderOut","moduleOut","sideOut","diskOut","panelOut","bladeOut","isBigIn","isEdgIn","isBadln","isBigOut","isEdgOut","isBadOut","isFlippedIn","isFlippedOut","iCSize","pixlnX","pixlnY","inClusterADC","iZeroADC","iCSize","iCSizeX","iCSizeY","iOverFlowX","iOverFlowY","oCSize","pixOutX","pixOutY","outClusterADC","oZeroADC","oCSize","oCSizeX","oCSizeY","oOverFlowX","oOverFlowY","diffADC"]
\]

24 labels defined only for MC matched doublets

\[
\text{tailLab} = ["idTrack","px","py","pz","pt","mT","eT","mSqr","rapidity","etaTrack","phi","pdgId","charge","noTrackerHits","noTrackerLayers","dZ","dXY","Xvertex","Yvertex","Zvertex","bunCross","isCosmic","chargeMatch","sigMatch"]
\]

Normalization with incident angle
**QCD processes recipe**

---
**PYTHIA Process Initialization**

We collide p+ with p+ at a CM energy of 1.300e+04 GeV

<table>
<thead>
<tr>
<th>Subprocess</th>
<th>Code</th>
<th>Estimated max (mb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>g g → g g</td>
<td>111</td>
<td>2.865e-07</td>
</tr>
<tr>
<td>g g → q qbar (uds)</td>
<td>112</td>
<td>7.838e-09</td>
</tr>
<tr>
<td>q qbar → g g</td>
<td>113</td>
<td>1.819e-06</td>
</tr>
<tr>
<td>q (bar)' → q q (bar)'</td>
<td>114</td>
<td>3.727e-07</td>
</tr>
<tr>
<td>q qbar → g g</td>
<td>115</td>
<td>3.415e-09</td>
</tr>
<tr>
<td>q qbar → q' qbar' (uds)</td>
<td>116</td>
<td>2.381e-09</td>
</tr>
<tr>
<td>g g → c cbar</td>
<td>121</td>
<td>2.361e-09</td>
</tr>
<tr>
<td>q qbar → c cbar</td>
<td>122</td>
<td>7.937e-10</td>
</tr>
<tr>
<td>g g → b bbar</td>
<td>123</td>
<td>2.361e-09</td>
</tr>
<tr>
<td>q qbar → b bbar</td>
<td>124</td>
<td>7.936e-10</td>
</tr>
</tbody>
</table>
---
**End PYTHIA Process Initialization**

---
**PYTHIA Event Listing (hard process)**

<table>
<thead>
<tr>
<th>no</th>
<th>id</th>
<th>name (system)</th>
<th>status</th>
<th>mothers</th>
<th>daughters</th>
<th>colours</th>
<th>p_x</th>
<th>p_y</th>
<th>p_z</th>
<th>e</th>
<th>m</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>90</td>
<td>(system)</td>
<td>-11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>13000.000</td>
<td>13000.000</td>
</tr>
<tr>
<td>1</td>
<td>2112</td>
<td>(p+)</td>
<td>-12</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>6500.000</td>
<td>6500.000</td>
</tr>
<tr>
<td>2</td>
<td>2112</td>
<td>(p+)</td>
<td>-12</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-6500.000</td>
<td>6500.000</td>
</tr>
<tr>
<td>3</td>
<td>21</td>
<td>(g)</td>
<td>-21</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>6</td>
<td>101</td>
<td>102</td>
<td>0.000</td>
<td>1811.756</td>
</tr>
<tr>
<td>4</td>
<td>21</td>
<td>(g)</td>
<td>-21</td>
<td>2</td>
<td>0</td>
<td>5</td>
<td>6</td>
<td>103</td>
<td>101</td>
<td>0.000</td>
<td>-125.216</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>(t)</td>
<td>-22</td>
<td>3</td>
<td>4</td>
<td>7</td>
<td>8</td>
<td>103</td>
<td>0</td>
<td>-89.558</td>
<td>-23.812</td>
</tr>
<tr>
<td>6</td>
<td>-6</td>
<td>(tbar)</td>
<td>-22</td>
<td>3</td>
<td>4</td>
<td>9</td>
<td>10</td>
<td>0</td>
<td>102</td>
<td>89.558</td>
<td>23.812</td>
</tr>
<tr>
<td>7</td>
<td>24</td>
<td>(W)</td>
<td>-22</td>
<td>5</td>
<td>0</td>
<td>11</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>3.982</td>
<td>38.668</td>
</tr>
<tr>
<td>8</td>
<td>5</td>
<td>b</td>
<td>23</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>103</td>
<td>0</td>
<td>-93.481</td>
<td>-54.477</td>
</tr>
<tr>
<td>9</td>
<td>-24</td>
<td>(W-)</td>
<td>-22</td>
<td>6</td>
<td>0</td>
<td>13</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td>125.765</td>
<td>29.471</td>
</tr>
<tr>
<td>10</td>
<td>-5</td>
<td>bbar</td>
<td>23</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>102</td>
<td>0</td>
<td>-36.215</td>
<td>-5.659</td>
</tr>
<tr>
<td>11</td>
<td>-1</td>
<td>dbar</td>
<td>23</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>184</td>
<td>0</td>
<td>11.092</td>
<td>45.836</td>
</tr>
<tr>
<td>12</td>
<td>2</td>
<td>u</td>
<td>23</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>104</td>
<td>0</td>
<td>-7.160</td>
<td>-15.171</td>
</tr>
<tr>
<td>13</td>
<td>3</td>
<td>s</td>
<td>23</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>105</td>
<td>0</td>
<td>27.651</td>
<td>-28.416</td>
</tr>
<tr>
<td>14</td>
<td>-4</td>
<td>cbar</td>
<td>23</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>105</td>
<td>0</td>
<td>98.114</td>
<td>49.887</td>
</tr>
</tbody>
</table>

Charge sum: 0.000
Momentum sum: 0.000 886.540 1136.972 711.865
---
**End PYTHIA Event Listing**
### MC Recipes - II

#### TTbar processes recipe

*--------- PYTHIA Process Initialization -------------------------*

We collide p+ with p+ at a CM energy of 1.300e+04 GeV

<table>
<thead>
<tr>
<th>Subprocess</th>
<th>Code</th>
<th>Estimated max (mb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>g g → t t̄bar</td>
<td>601</td>
<td>7.216e-06</td>
</tr>
<tr>
<td>q q̄bar → t t̄bar</td>
<td>602</td>
<td>9.331e-07</td>
</tr>
</tbody>
</table>

*--------- End PYTHIA Process Initialization -------------------------*

*--------- PYTHIA Event Listing (hard process) -------------------------*

<table>
<thead>
<tr>
<th>no</th>
<th>id</th>
<th>name</th>
<th>status</th>
<th>mothers</th>
<th>daughters</th>
<th>colours</th>
<th>p_x</th>
<th>p_y</th>
<th>p_z</th>
<th>e</th>
<th>m</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>90</td>
<td>(system)</td>
<td>-11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>1</td>
<td>2212</td>
<td>(p+)</td>
<td>-12</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>2</td>
<td>2212</td>
<td>(p̄)</td>
<td>-12</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>3</td>
<td>21</td>
<td>(g)</td>
<td>-21</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>6</td>
<td>101</td>
<td>102</td>
</tr>
<tr>
<td>4</td>
<td>21</td>
<td>(g)</td>
<td>-21</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>6</td>
<td>103</td>
<td>104</td>
</tr>
<tr>
<td>5</td>
<td>21</td>
<td>g</td>
<td>23</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>101</td>
<td>104</td>
<td>187.257</td>
</tr>
<tr>
<td>6</td>
<td>21</td>
<td>g</td>
<td>23</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>103</td>
<td>102</td>
<td>-187.257</td>
</tr>
</tbody>
</table>

Charge sum: 0.000
Momentum sum: 0.000 0.000 1098.312 2174.776 1877.062

*--------- End PYTHIA Event Listing -------------------------*
### ZEE processes recipe

*------- PYTHIA Process Initialization -------*

We collide p+ with p+ at a CM energy of 1.300e+04 GeV

<table>
<thead>
<tr>
<th>Subprocess</th>
<th>Code</th>
<th>Estimated max (mb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>f fbar -&gt; gamma*/Z0</td>
<td>221</td>
<td>3.845e-06</td>
</tr>
</tbody>
</table>

*------- End PYTHIA Process Initialization -------*

<table>
<thead>
<tr>
<th>no</th>
<th>id</th>
<th>name</th>
<th>status</th>
<th>mothers</th>
<th>daughters</th>
<th>colours</th>
<th>p_X</th>
<th>p_Y</th>
<th>p_Z</th>
<th>e</th>
<th>m</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>90</td>
<td>(system)</td>
<td>-11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1300.000</td>
</tr>
<tr>
<td>1</td>
<td>2212</td>
<td>(p+)</td>
<td>-12</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6500.000</td>
</tr>
<tr>
<td>2</td>
<td>2212</td>
<td>(p+)</td>
<td>-12</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6500.000</td>
</tr>
<tr>
<td>3</td>
<td>-2</td>
<td>(ubar)</td>
<td>-21</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>101</td>
<td>0</td>
<td>1.842</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>(u)</td>
<td>-21</td>
<td>2</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>101</td>
<td>0</td>
<td>1.842</td>
</tr>
<tr>
<td>5</td>
<td>23</td>
<td>(Z0)</td>
<td>-22</td>
<td>3</td>
<td>4</td>
<td>6</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1105.446</td>
</tr>
<tr>
<td>6</td>
<td>11</td>
<td>e^-</td>
<td>23</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-10.016</td>
<td>3.637</td>
<td>-13.819</td>
</tr>
<tr>
<td>7</td>
<td>-11</td>
<td>e+</td>
<td>23</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10.016</td>
<td>-3.637</td>
<td>1089.837</td>
</tr>
</tbody>
</table>

Charge sum: 0.000  
Momentum sum: 0.000  

*------- End PYTHIA Event Listing -------*