Deep learning jet clustering algorithm for analysis of particle collisions at the Large Hadron Collider

Camila Cendra, Justin Chen, Khaled Jedoui, {ccendra, jye100, thekej}@stanford.edu

Mentored by Ariel Schwatzman, Murtaza Safdari, Lucio Dery

Problem

- ATLAS is a physics detector on the Large Hadron Collider (LHC) looking at proton-proton collisions.
- The LHC makes proton-proton collisions 40 million times per second. However, only 1,000 events per second can be saved.
- Decisions to save events are made based on the presence of sprays of particles, called jets.
- Goal: We train a deep neural network that learns the rules of a theory-based jet clustering algorithm to identify jets.

Data

- The dataset consists of 100k events.
- 64 x 64 pixels
- pT energy
- (η, φ) coordinates
- Ground truth jets are defined by their constituent pixels by running the FastJet anti-kt algorithm on the input images.
- Thresholded for jets with pT > 20 GeV.

Model

- **Mask R-CNN [2]**
  - Region proposal / Detection
  - Classification
  - Segmentation

  Modifications:
  - Reduced output space (binary classification)
  - Reduced model depth
  - Adjusted region proposal parameters for smaller images

Model Loss: $\mathcal{L} = \mathcal{L}_\text{cross-entropy} + \mathcal{L}_\text{mask}$

Results

- **Event with true jets**
- Implemented early stopping because of overfitting object detection on training data.
- Post-processing of predictions through filtering by energy of jets and confidence score.
- Used F1 score to evaluate performance on dev set.
- True positives defined as predicted bounding boxes with Intersection over Union (IoU) ≥ β with ground truth.
- Correlation of 0.647 between energy of maximum pT jet per image in predictions and ground truths.
- Forward propagation in ~64ms per event on Nvidia GTX 1060-6GB.

<table>
<thead>
<tr>
<th>Post-processing filters *</th>
<th>precision</th>
<th>recall</th>
<th>F1 score</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>0.2459</td>
<td>0.3005</td>
<td>0.2705</td>
</tr>
<tr>
<td>pT &gt; 20</td>
<td>0.6433</td>
<td>0.4245</td>
<td>0.5115</td>
</tr>
<tr>
<td>pT &gt; 20 or confidence &gt; 0.98</td>
<td>0.5086</td>
<td>0.4135</td>
<td>0.4562</td>
</tr>
<tr>
<td>pT &gt; 20 and confidence &gt; 0.98</td>
<td>0.8592</td>
<td>0.5516</td>
<td>0.6719</td>
</tr>
</tbody>
</table>

* Model performance based on IoU threshold: β = 0.3.

Discussion and Future Steps

- Success in identifying high-energy jets.
- Worse performance in per-pixel segmentation.
- Larger training dataset to reduce variance.
- More diverse training set to account for variance in event data between different experiments (more jets, higher pT, etc.).
- Stagger multi-task loss to continue training on segmentation without overfitting detection.

References

  arXiv:1703.06870, 2017