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

Camila Cendra, Justin Chen , Khaled Jedoui {ccendra, jyc100, thekej}@stanford.edu
Mentored by Ariel Schwatzman, Murtaza Safdari, Lucio Dery

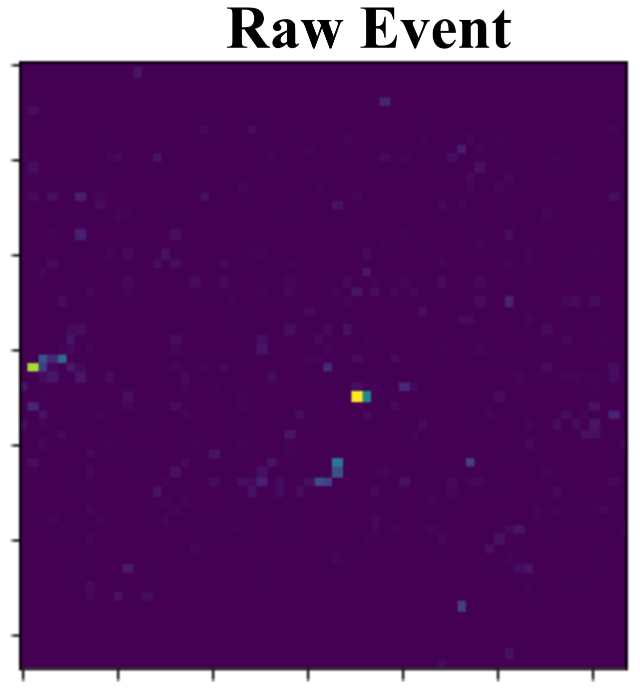
Problem

- 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.

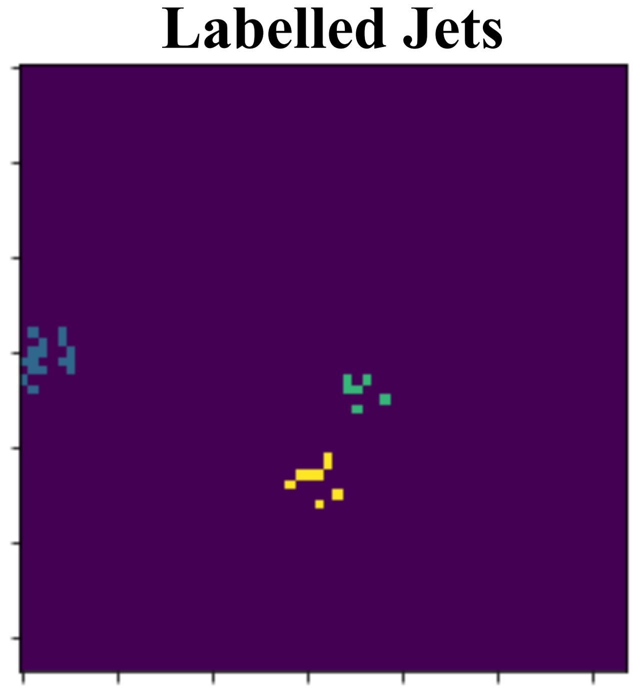
Data

- 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 [1] on the input images.
- Thresholded for jets with $pT > 20\ GeV$.

Raw Event



Labelled Jets



- Bounding boxes are calculated around each labelled jet before training.
- Data is split 99% / 1% between train / dev sets.

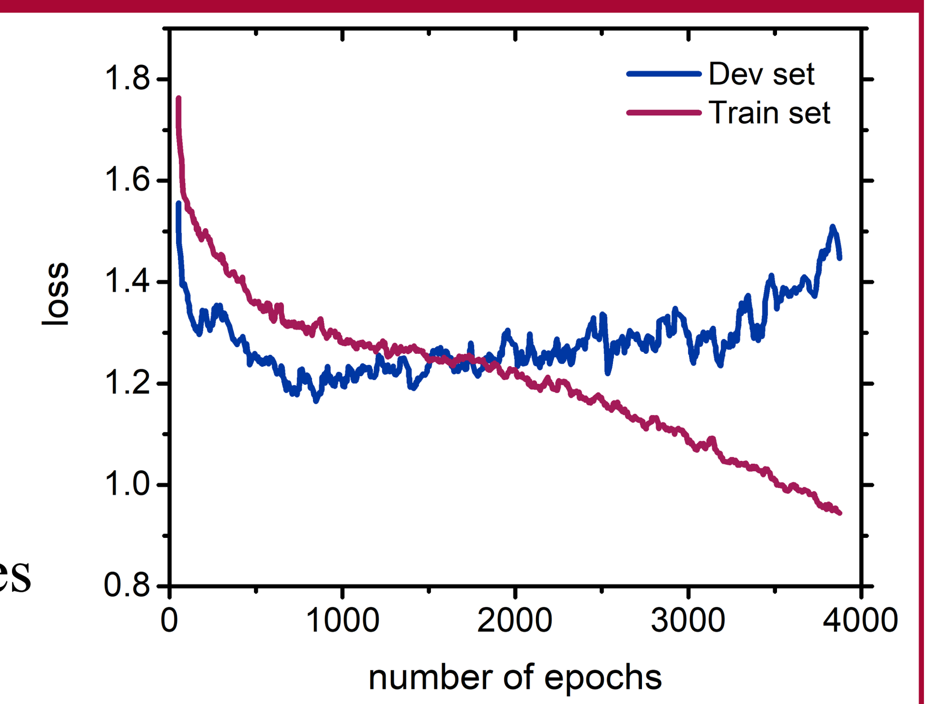
Model

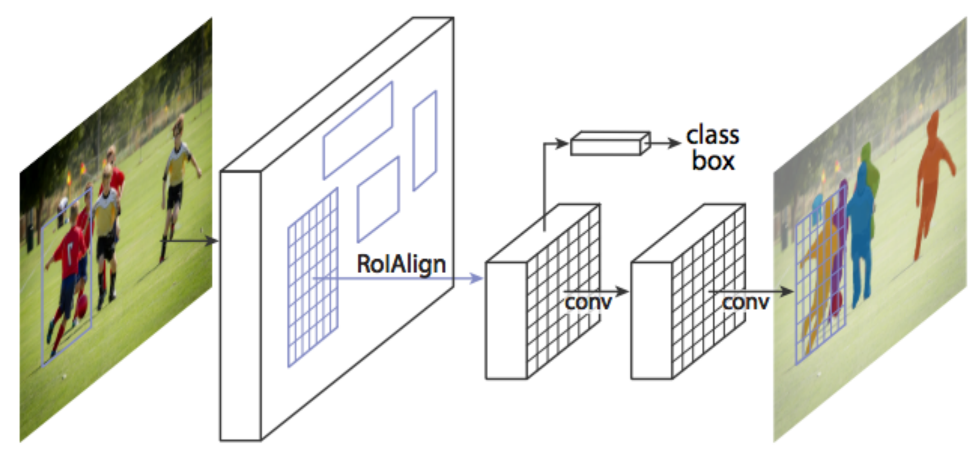
Mask R-CNN [2]

- Region proposal / Detection
- Classification
- Segmentation

Modifications:

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





$$Loss = L_{class} + L_{box} + L_{mask}$$

L_{class}: cross-entropy loss

L_{box}: L_1 loss over box coordinates

L_{mask}: per-pixel cross-entropy loss

Discussion and Future Steps

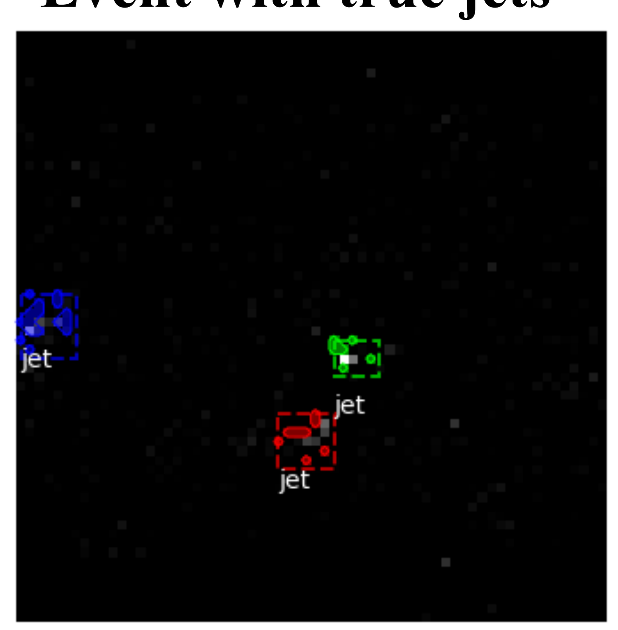
- **Success in identifying high-energy jets.**
- **Worse performance in per-pixel segmentation.**
- Larger training dataset to reduce variance.
- More diverse training set to account for variance in event data between different experiments (more jets, higher pT, etc.).
- Stagger multi-task loss to continue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

References

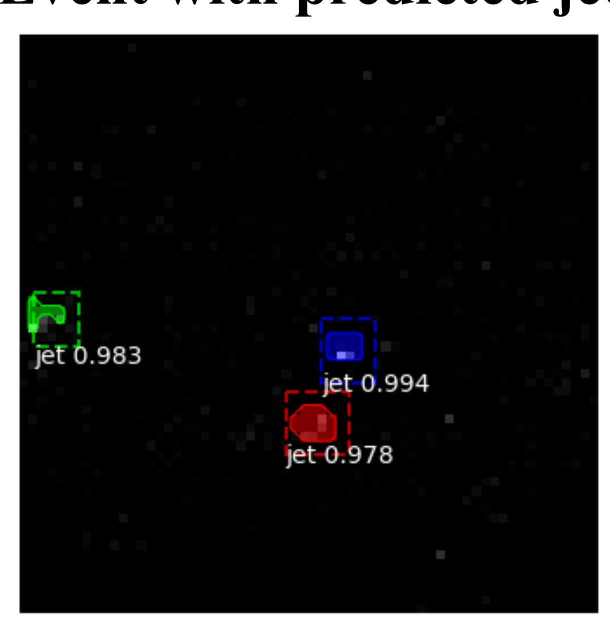
[1] M. Cacciari, G.P. Salam and G. Soyez, Eur.Phys.J. C72 (2012) 1896 [arXiv:1111.6097]
[2] K. He, G. Gkioxari, P. Doll´ar, and R. Girshick. Mask R-CNN. arXiv:1703.06870, 2017
[3] https://github.com/matterport/Mask_RCNN

Results

Event with true jets



Event with predicted jets

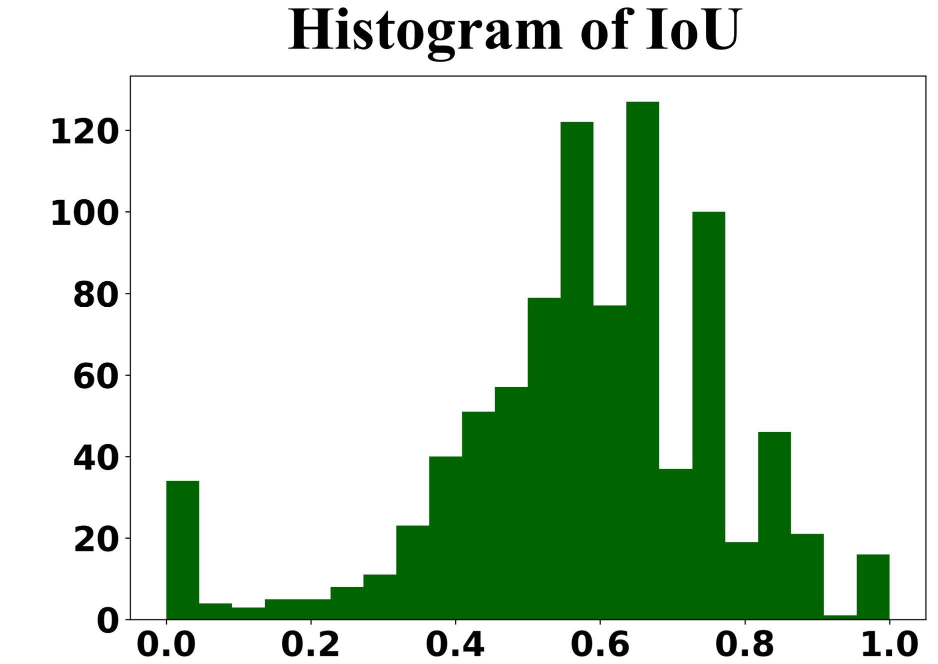


- Implemented early stopping because of overfitting object detection on training data.
- Post-processing of predictions through filtering by energy of jets and confidence score.
- Used F1 score to evaluate performance on dev set.
- True positives defined as predicted bounding boxes with Intersection over Union (IoU) $\geq \beta$ 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.

Post-processing filters *	precision	recall	F1 score
None	0.2459	0.3005	0.2705
pT > 20	0.6433	0.4245	0.5115
pT > 20 or confidence > 0.98	0.5086	0.4135	0.4562
pT > 20 and confidence > 0.98	**0.8592**	**0.5516**	**0.6719**

* Model performance based on IoU threshold: $\beta = 0.3$.

Histogram of IoU



# of jets ($pT > 20$) per event