Using Deep Learning to Enhance the Particle Identification Abilities of the Transition Radiation Tracker

Geoff Gaugler | Mark Kruse | Duke University

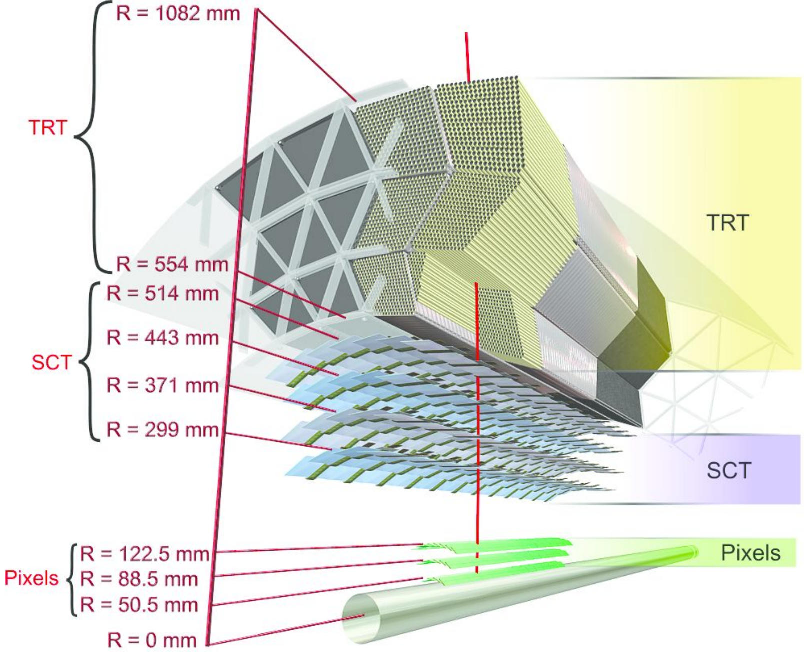
Background - the TRT

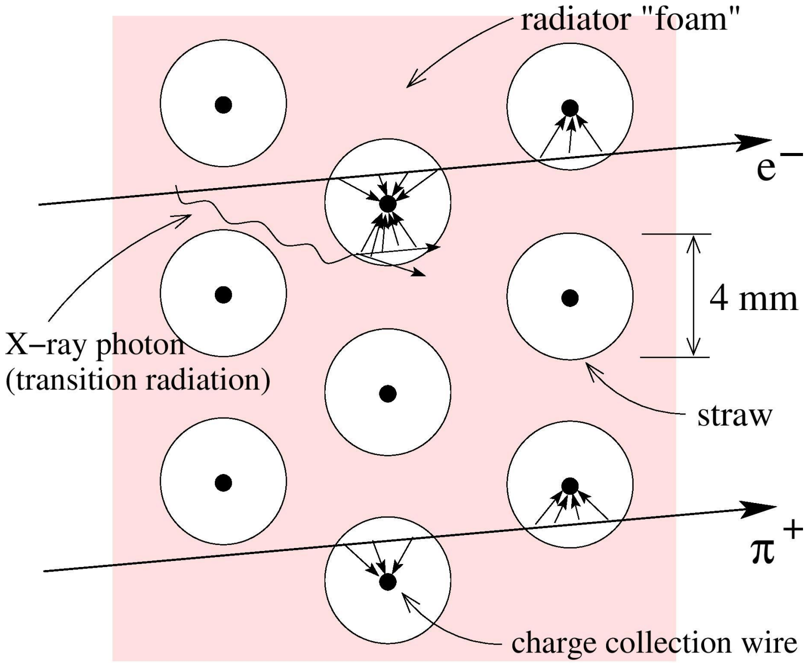
Structure

- 350,000 straw tubes filled with an ionizing gas
    - Run I: all xenon
    - Run II: mixture
    - Run III: likely all argon
- Seated in a radiator foam
- Wire in center of each straw

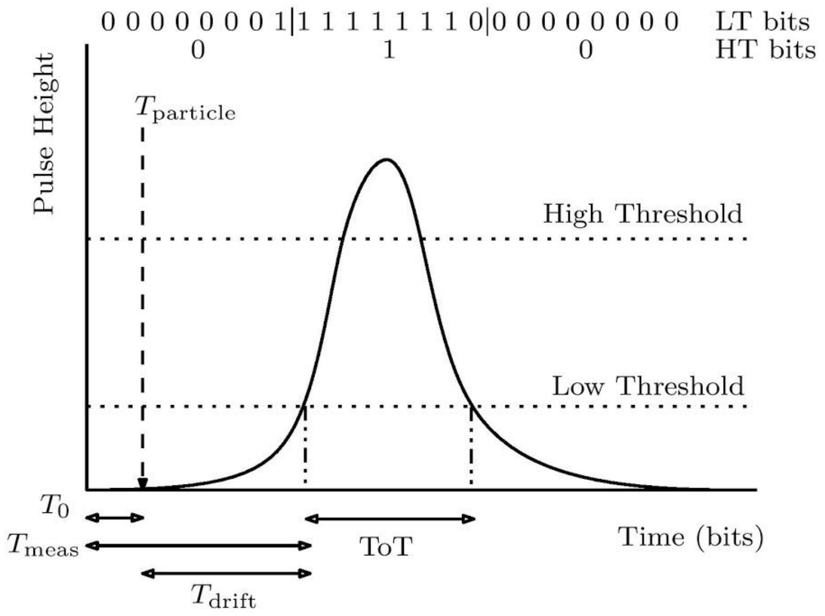
Particle identification (PID)

- Utilize transition radiation
- Mostly to distinguish pions from electrons
- Happens when γ > ~1000
- $E = \gamma mc^2$
- E.g. for electron: m = .5 MeV corresponds to E > .5 GeV
- For pion: m = 140 MeV corresponds to E > 140 GeV





Background - the TRT read-out Bitpattern



How do we use this?

- Precise charge collection measurement not recorded
- Two thresholds: 300 eV (low) and 6 keV (high)
- 75 ns time window per straw
  - 24 low threshold 3.125 ns time bins
  - 3 high threshold 25 ns time bins

The bitpattern

- For each time bin: 1 if signal surpasses threshold, 0 if not
- Result: ‘TRT bitpattern’
- Bit pattern depends on the particle type, gas used, and particle momentum and direction
- Time over threshold = the time a pulse height is above 0 for LT

Background - motivation

Problems created by gas geometry for run III

- Xenon easily ionizes TR photons, giving HT hits for electrons but not for pions
  - ...but expensive!
- Gas leaks -> switch to much cheaper argon
- Argon less sensitive to TR photons -> HT hits less consistent for electrons -> **bad for PID**

Current PID method - eProbHT tool

- Uses HT hits, track occupancy
- Significant performance drop
  - .89 vs. .71 AUC

$$\mathcal{L}^{e,\mu} = \prod_{\text{TRT hits}} \begin{cases} p_{\text{HT}}^{e,\mu} & \text{if HT hit} \\ 1 - p_{\text{HT}}^{e,\mu} & \text{else} \end{cases}$$

What variables not affected by this problem could be useful?

$$\mathcal{P}^{e} = \frac{\mathcal{L}^{e}}{\mathcal{L}^{e} + \mathcal{L}^{\mu}}.$$

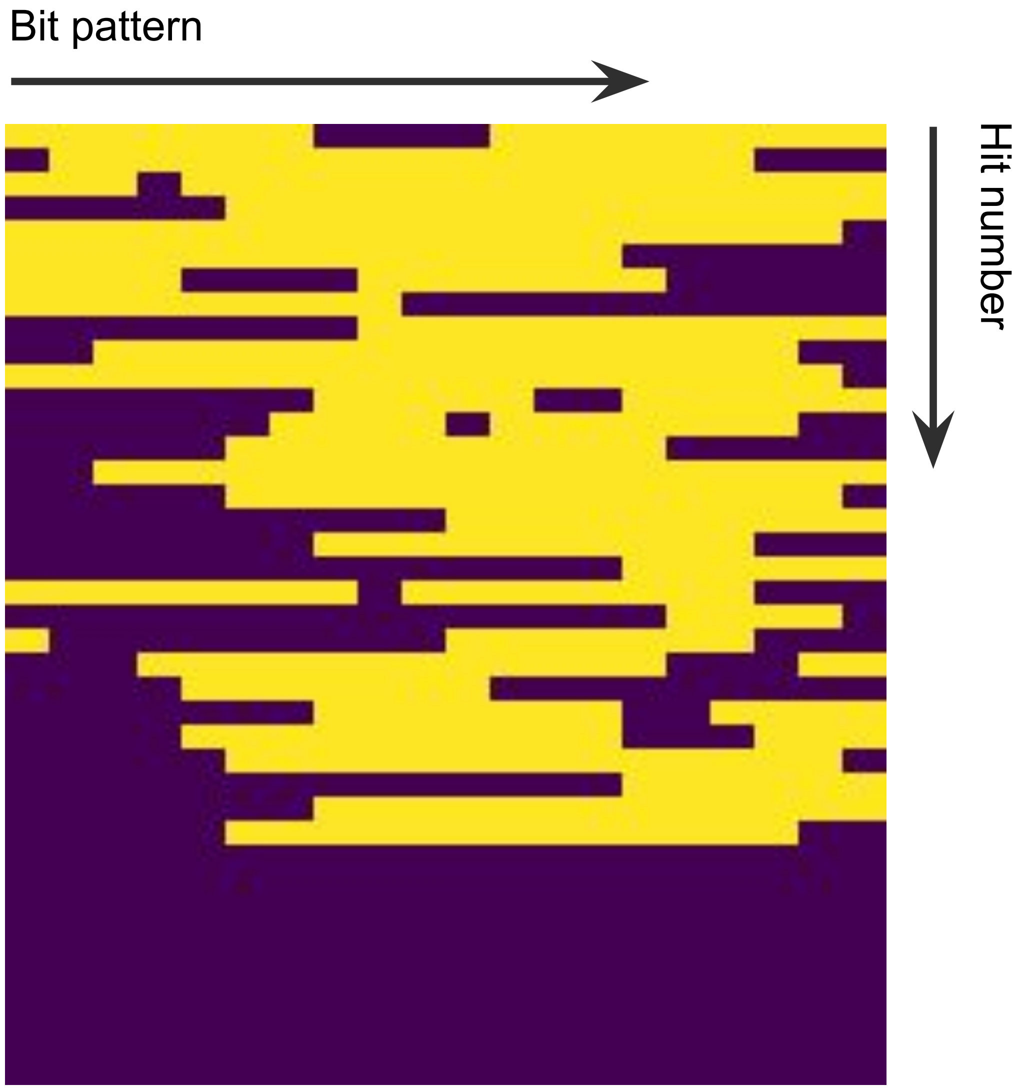
Approach

Many properties of hit extracted from low threshold bit pattern

- Using just these variables -> lose other info contained in entire BP
- Any discriminating power in LT BP?
  - E.g. ToT length, ToT ‘islands’, correlations of patterns between straws
- We don’t know what we don’t know!

Treat LT bit patterns as images

- 2D array - 20 bits in each bit pattern, ~40 hits per track
- Computer vision techniques on LT BP - CNNs
- Feed hit-level variables into LSTMs (gas type, R of hit, etc.)
- Feed track-level ones into FCNs (track occupancy, track η, etc.)
- Using simulated $Z \rightarrow ee$ and $Z \rightarrow \mu\mu$ data



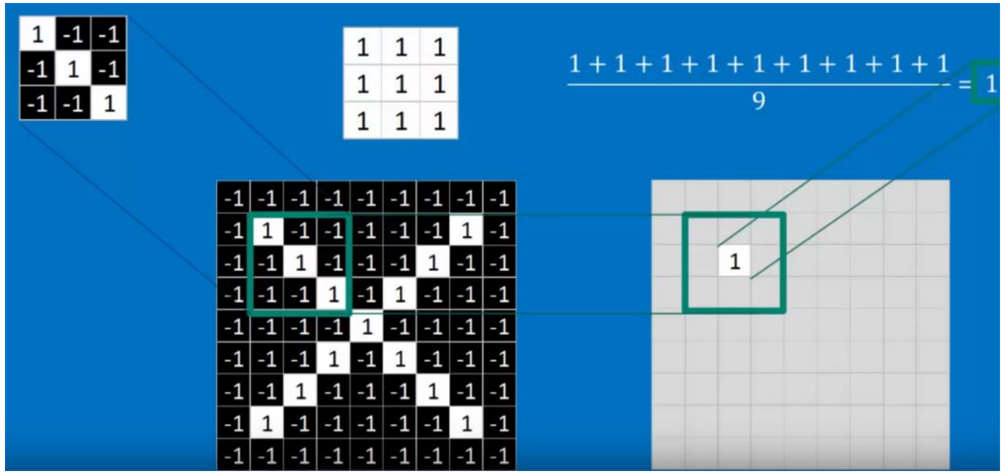
Convolutional Neural Networks

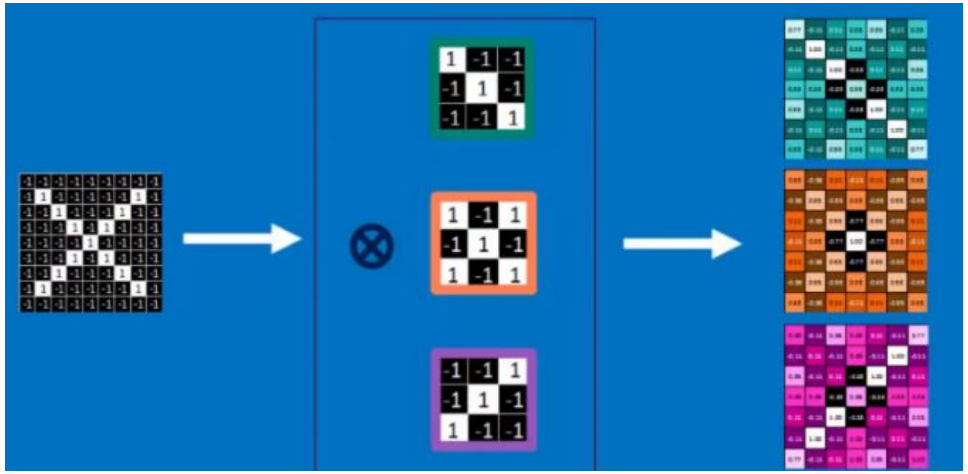
Convolution Layer

- Filtering across every possible image patch
- Doing this for multiple features: single image -> stack of filtered images

Idea

- Learn local patterns in feature input space
- Learn spatial hierarchies of patterns





Obstacles

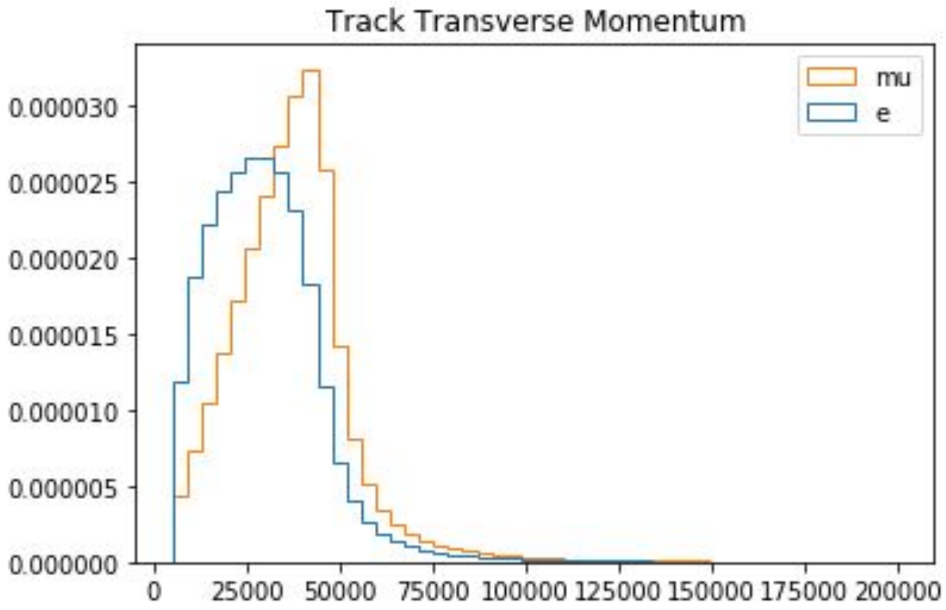
Data

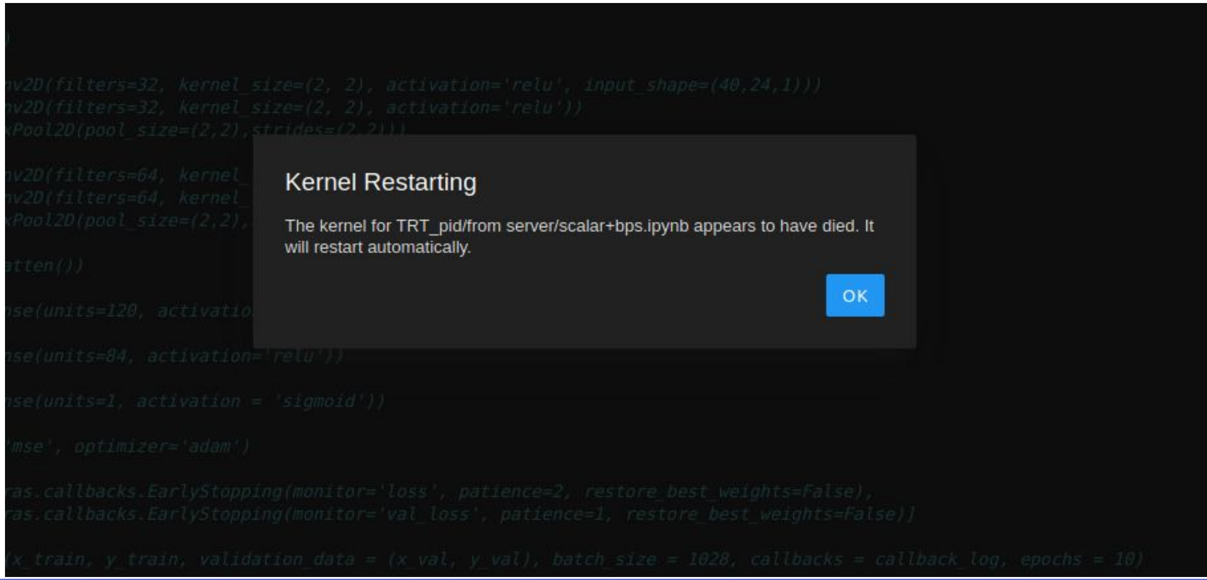
- Amount
- Availability

Computationally intensive

- Can't use personal machine
- Frequent server downtime
- Long training times

Constantly learning - double edged sword!





Current Performance and Future Work

Outperforming current PID tool

- Currently achieving .86 AUC
  - 20% increase over current tool!

Will continue this fall

- Making model shift-invariant for timing corrections
- Explore autoencoders
- Experiment with ensembling methods

Great potential for use in an all-Argon TRT gas geometry in Run III

  - Continue to work with and report to the TRT software group to continue its development

Acknowledgments

US ATLAS SUPER Program

Mark Kruse, Doug Davis

TRT SW group members

- Fred Luehring
- Arif Bayirli