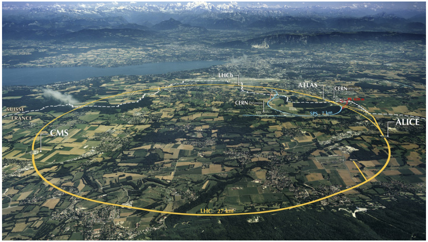
Convolutional Neural Networks for Processing of ATLAS Liquid Argon Calorimeter Signals with FPGAs

Anne-Sophie Berthold

IKTP, TU Dresden

DPG Conference Dresden 2023, March 20

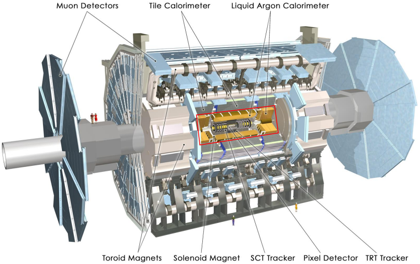
The ATLAS Detector at the LHC



Source: `https://static1.bmbfcluster.de/3/4/3/8_ef6a5eef8f44963/3438meg_22ce2885dae52af.jpg`

The Large Hadron Collider (LHC)

- 27 km circular collider at CERN/Geneva
- Accelerated proton bunches collide with 25 ns spacing (40 MHz)
- 2029: Start of **High Luminosity LHC** (HL-LHC) with 5 to 7.5 x nominal lumi



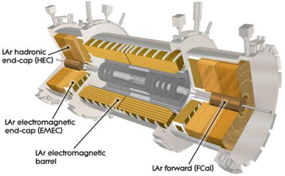
Source: The ATLAS Experiment at the CERN Large Hadron Collider, The ATLAS Collaboration et al, 2008 JINST 3 S08003

The ATLAS Detector

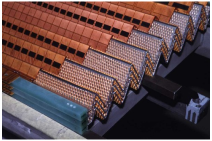
2026-2028: phase-2 upgrade

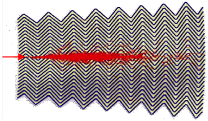
- Has to cope with **140-200 pile-up events** per collision
- Readout electronics of Liquid-Argon (LAr) calorimeters need to be improved

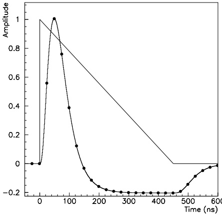
Signal Readout of the ATLAS LAr Calorimeter



Source: https://cds.cern.ch/record/1095928







Source: The ATLAS Experiment at the CERN Large Hadron Collider, The ATLAS Collaboration et al, 2008 JINST 3 S08003

particle shower → readout per cell

Source: http://www.particles.uni-freiburg.de/dateien/vorlesungsdateien/particledetectors/kap8

LAr Calorimeter

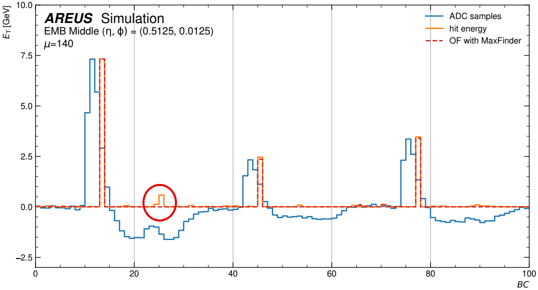
- Absorber material (Pb, Cu, W) and electrodes in accordion geometry
- In between **liquid argon** (→ LAr) as active medium

Signals

- Energy deposits raise triangular pulse
- Shaped by $CR(RC)^2$ analog filter into **bipolar pulse** and digitized
- Amplitude proportional to deposited energy
- Calculation in **realtime** by FPGAs

The Current System

- **Optimal filtering** algorithm (**OF**) applied to calculate deposited energy per cell
- Optimized to suppress noise and to reconstruct peak timing
- Trigger system applies additional **maximum finder** ($\rightarrow$ **OFMax**), is potentially insensitive during undershoot of pulse



$$y_t = \sum_{n=0}^{M-1} x_{t-n} \cdot a_n$$

y_t … OF output for bunch crossing (BC) t

$\vec{x}$ … input ADC samples

a_n … OF coefficients

M … OF filter depth

Convolutional Neural Networks (CNNs)

- Convolutional operation with certain **filter/kernel** size
- **Activation function** gives opportunity to classify, weight, cut

$$y_t = A\left(\sum_{n=0}^{2} x_{t-n} \cdot w_n + b\right)$$

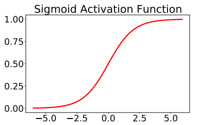
y_t ... CNN output for bunch crossing (BC) t

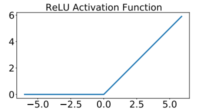
$\vec{x}$ ... input ADC samples

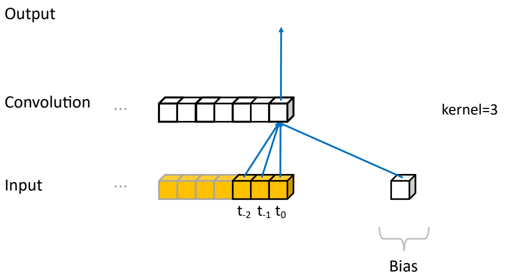
w_t ... weights of kernel with size 3

b ... bias term

A ... activation function

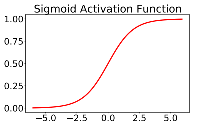


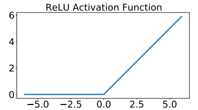


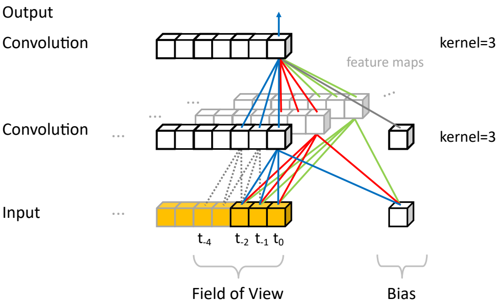


Convolutional Neural Networks (CNNs)

- Convolutional operation with certain **filter/kernel** size
- **Activation function** gives opportunity to classify, weight, cut
- **Feature maps** focus on different properties
- **Training** minimizes difference between output and target

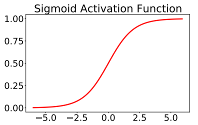


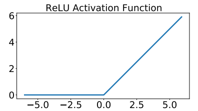


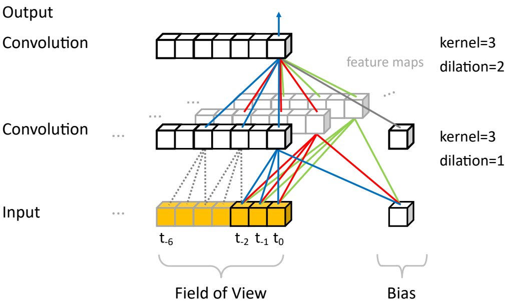


Convolutional Neural Networks (CNNs)

- Convolutional operation with certain **filter/kernel** size
- **Activation function** gives opportunity to classify, weight, cut
- **Feature maps** focus on different properties
- **Training** minimizes difference between output and target
- **Dilation** varies field of view (FOV) without increasing parameters
- Keep parameters low ($\approx$ 100) and FOV realistic ($\approx$ 24) since FPGA implementation planned



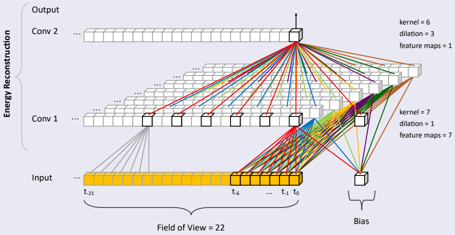




Energy Reconstruction by CNNs

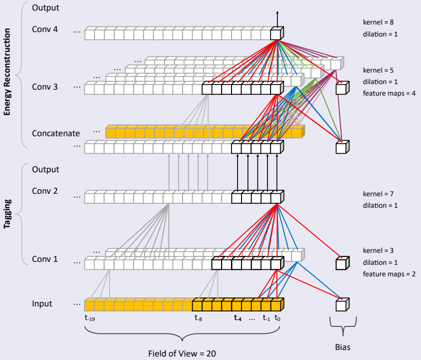
Plain 2-layered CNN (2CNN)

- Input: digitized signals
- 2 layers ReLU activation function
- Uses dilation
- Output: reconstructed energy



4-layered CNN with Tagging (4TCNN)

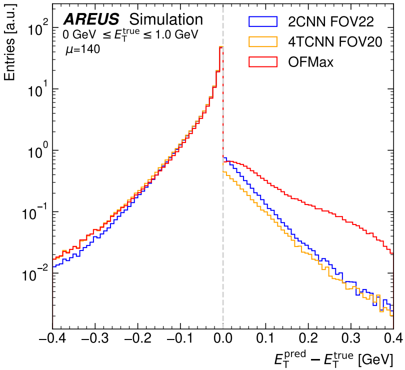
- Input: digitized signals
- 2 layers sigmoid activation function
- Intermediate output tags signals
- 2 layers ReLU activation function
- Output: reconstructed energy



Performance Evaluation: Energy Resolution

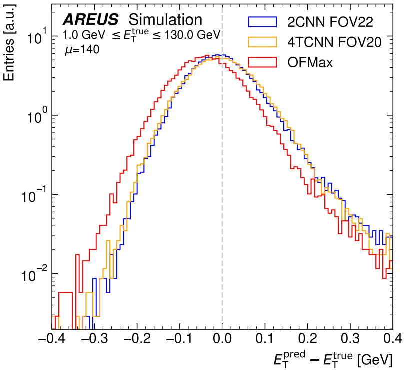
Pile-up Energy Region

0 GeV - 1.0 GeV



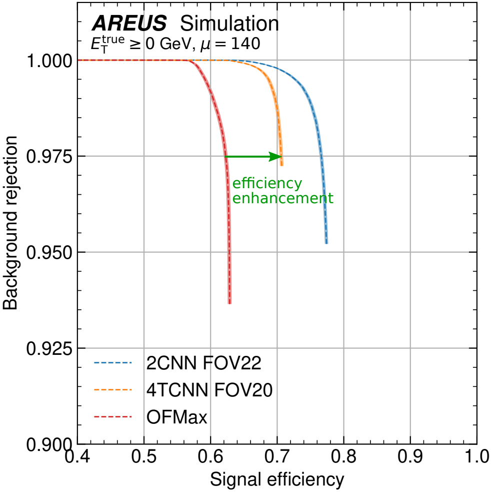
Signal Energy Region

1.0 GeV - 130.0 GeV



- CNNs improve energy resolution especially in low energy region
- CNNs show smaller energy bias especially in high energy region

Performance Evaluation: Signal Efficiency

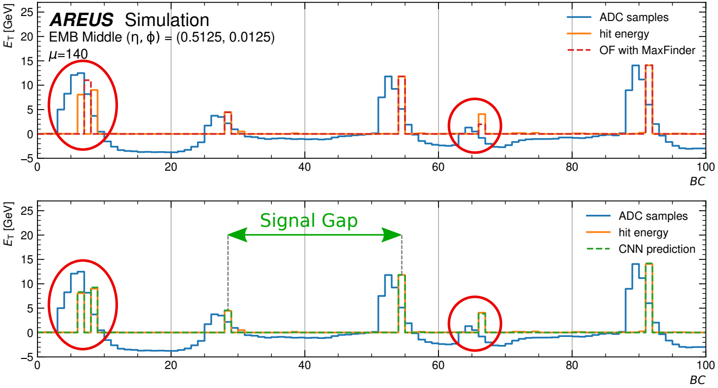


Receiver Operating Characteristic (ROC) Curves

- Indicate detection performance
- Signal efficiency $= \frac{\text{true positives}}{\text{true positives} + \text{false negatives}}$
- Background rejection $= \frac{\text{true negatives}}{\text{true negatives} + \text{false positives}}$
- Dependent on threshold

CNNs reach **higher signal efficiencies** at same background rejection level compared to OFMax

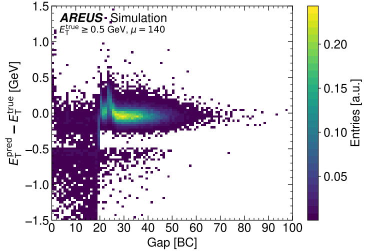
Performance Evaluation: Influence of Overlapping Signals



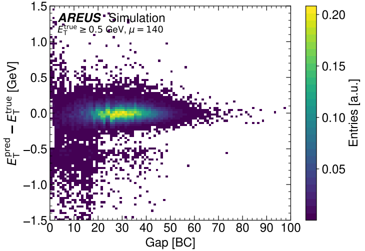
- New trigger accept scheme → overlapping signals possible
- CNN optimized to reconstruct overlapping signals
- Performs well even with very close overlaps (BC6-8), when OF cannot resolve it

Performance Evaluation: Performance on Subsequent Hits

Optimal Filter with Maximum Finder



CNN with 2 Convolutional layers



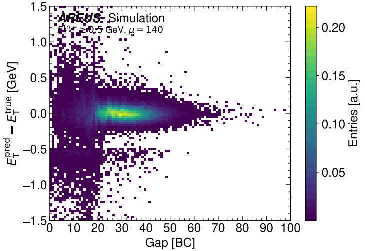
OF

- Not able to resolve overlapping signals < 20 BCs
- Shows artifact at BCs 20-25, at rising edge of pulse

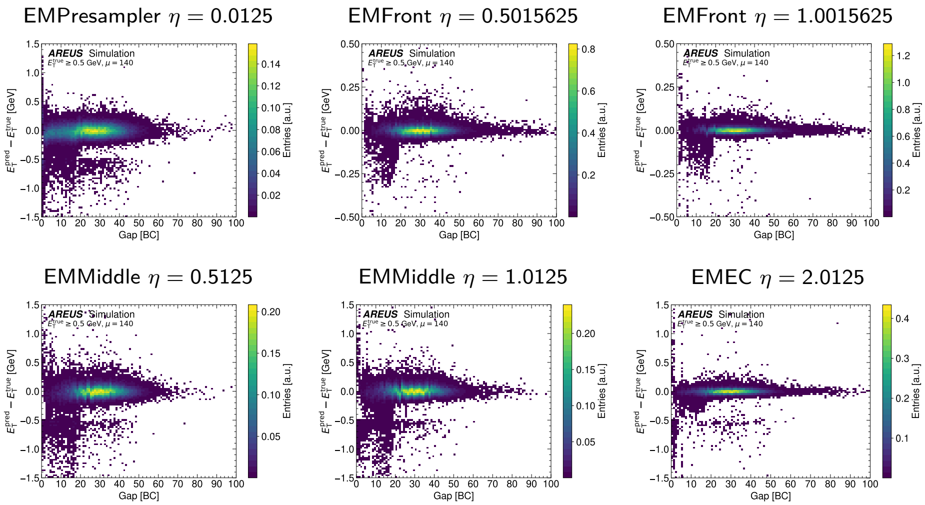
CNNs

- Much better overall performance
- Small offsets

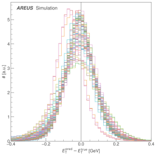
TCNN with 4 Convolutional layers



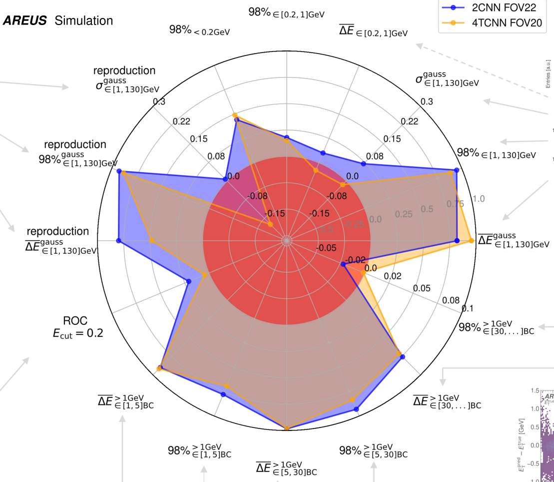
Performance Evaluation: Different Detector Cells

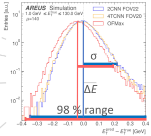


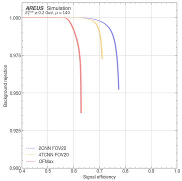
Performance evaluation on different, representative detector cells show similar results

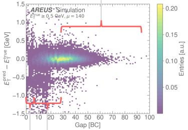


- Score indicates performance
- Approaching outer circle (=1) is best
- Red circle indicates OFMax yield
- Crossing red circle: performance worse than OFMax









→ **CNNs outperform OF for most parameters**

Summary

Convolutional Neural Networks ...

- Proved to be able to replace the Optimal Filtering algorithm
- Show better performance than the OF regarding energy resolution and signal efficiency, special benefit lies in detection and reconstruction of overlapping signals
- Application on different cells is consistent
- Evaluation with spider diagram enables to compare most important parameters
- Have small number of parameters ($\approx$ 80-100), so that FPGA resource limitations can be satisfied (see also talks by Johann Voigt and Alexander Lettau)

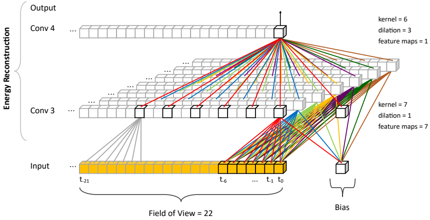
Thanks for your Attention!

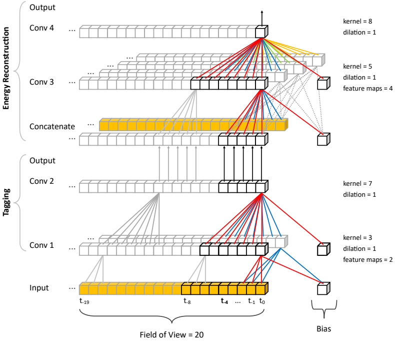
Backup

Architectures

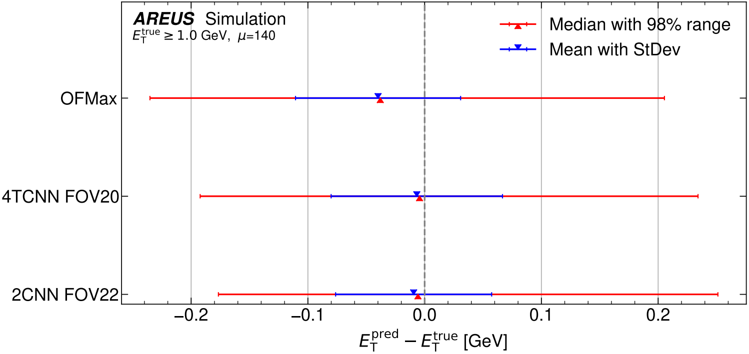
2CNN	
kernel	[7,6]
feature maps	[7,1]
dilation	[1,3]
activation function	[ReLU, ReLU]

4TCNN	
kernel	[3,7,5,8]
feature maps	[2,1,4,1]
dilation	[1,1,1,1]
activation function	[sig, sig, ReLU, ReLU]

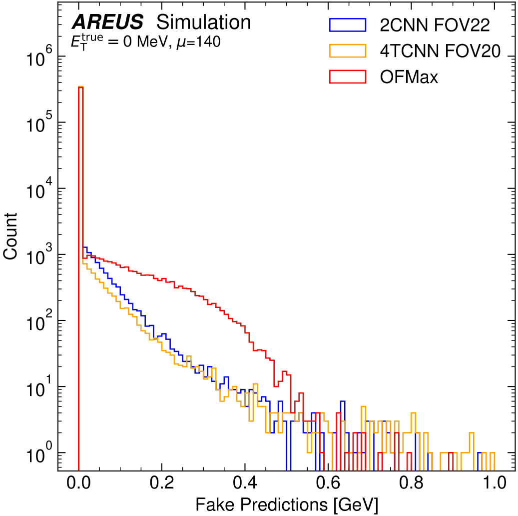




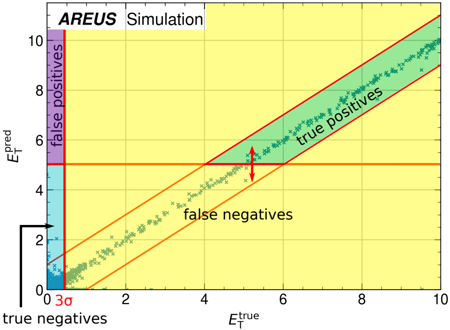
Performance Evaluation: OF vs CNN - Energy Resolution

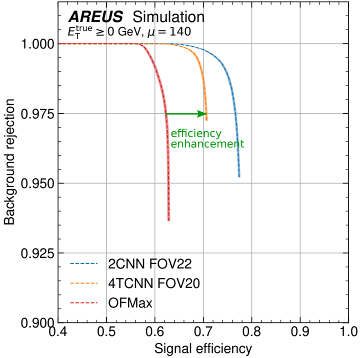


Fake Predictions



ROC curves for CNNs

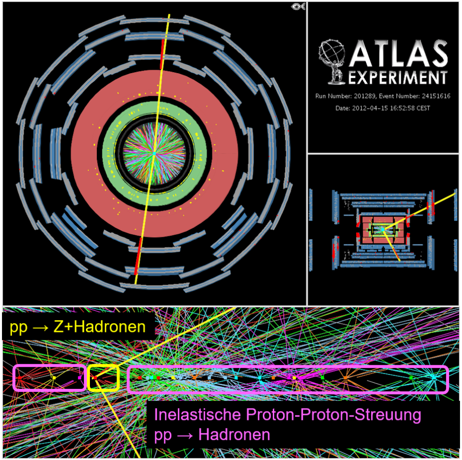




- declare everything above 3σ electronic noise as signal
- exclude predictions with $1\,\text{GeV} > E_{true} - E_{pred}$ to avoid counting bad energy reconstructions as true positives
- use largest rectangle (instead of integral) under the curve as score:

$$score = \frac{ROC_{\square}^{\text{ANN}} - ROC_{\square}^{\text{OFMax}}}{1 - ROC_{\square}^{\text{OFMax}}}$$

Event Display with Pileup



https://iopscience.iop.org/article/10.1088/1742-6596/523/1/012018/pdf