Deep Learning and Data Monitoring

KM3NeT Outing, 24.05.2019

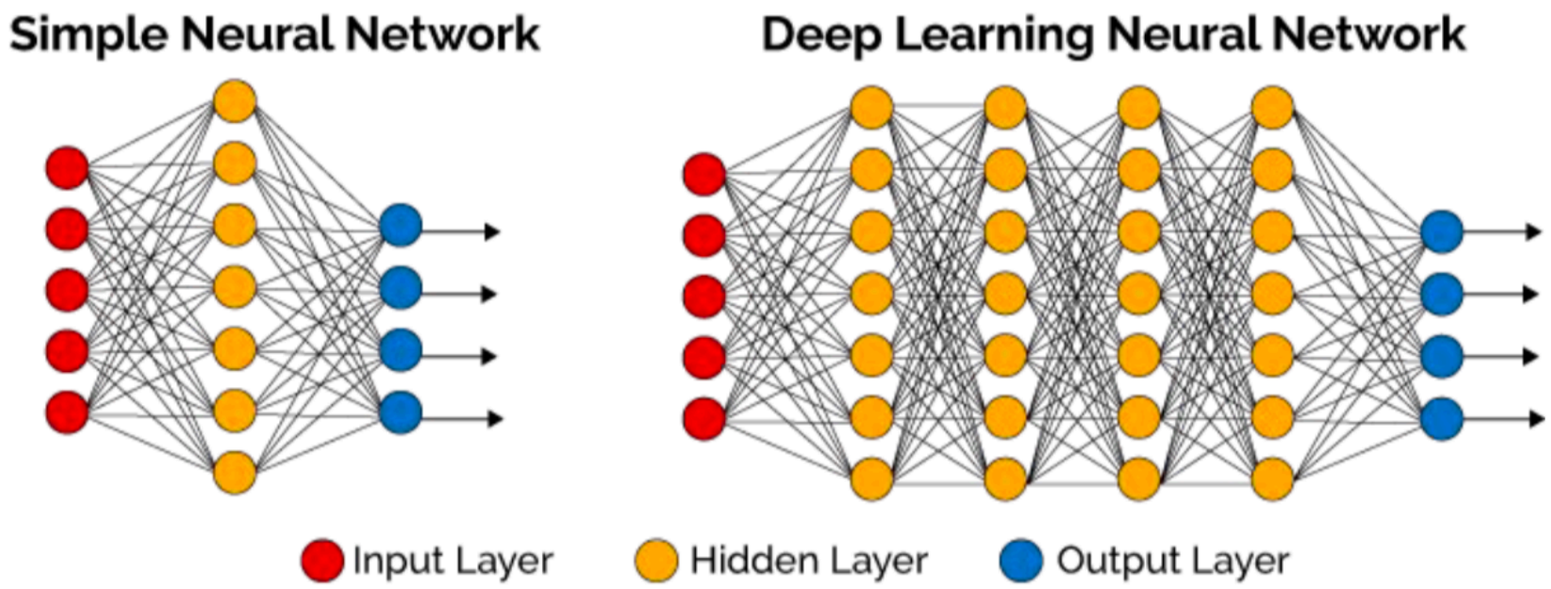
Fatih Bay

Outline

- Deep Learning methods used for Computer Vision
- Studies done and further steps
- DAQ Issues for DUNE
- Road Map

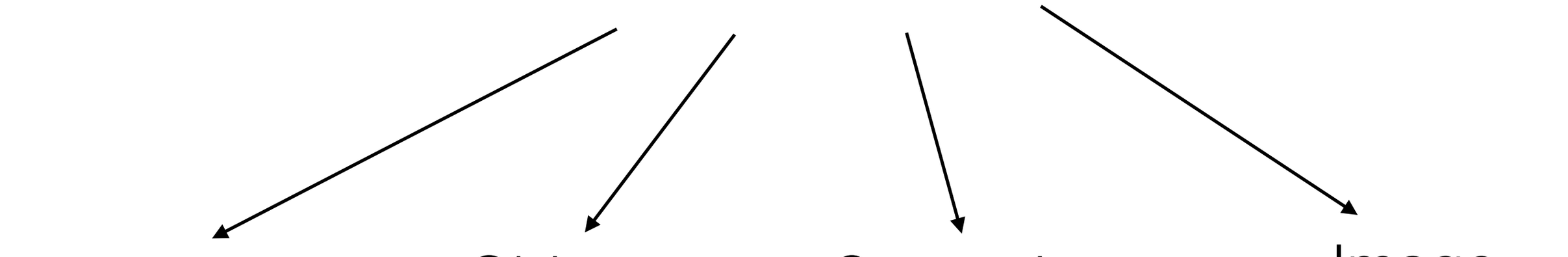
Deep Learning (DL)

- DL is a subfield of Machine Learning.
- The algorithms are inspired by the structure and function of the brain.
- The structure is called artificial neural network.



By Antonella Massini - MIT

DL employed for Computer Vision



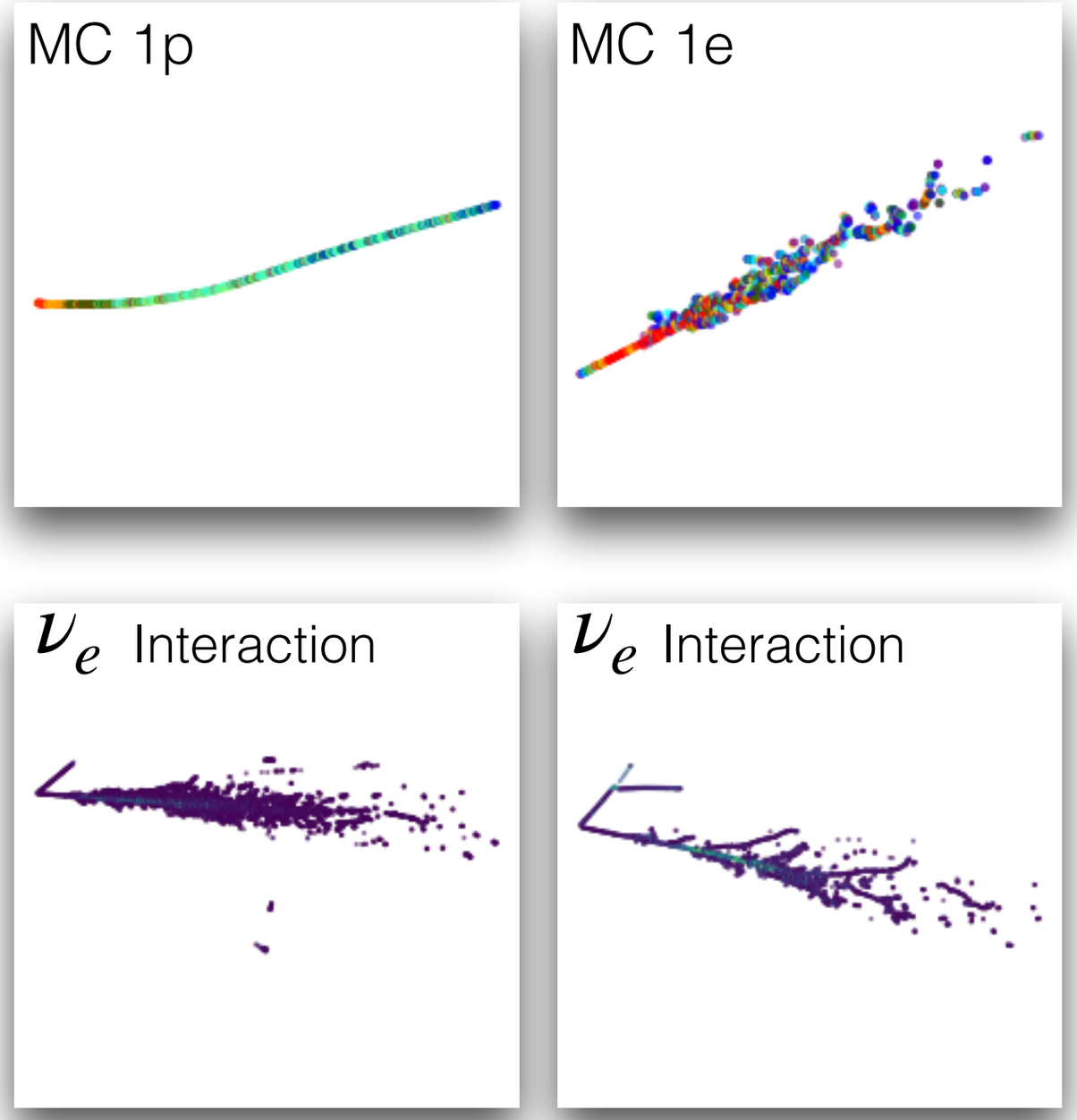
For state-of-the-art studies, models, papers, etc:
https://paperswithcode.com/sota

Image Classification

- Liquid Argon Time Projection Chamber (LAr-TPC) is an imaging calorimeter (mm resolution)

- **Single Images** can be represented as a matrix of pixel values (Input)

- **Particle Identification** and **Neutrino Event Classification** can be realised by using DL Image Classification Method

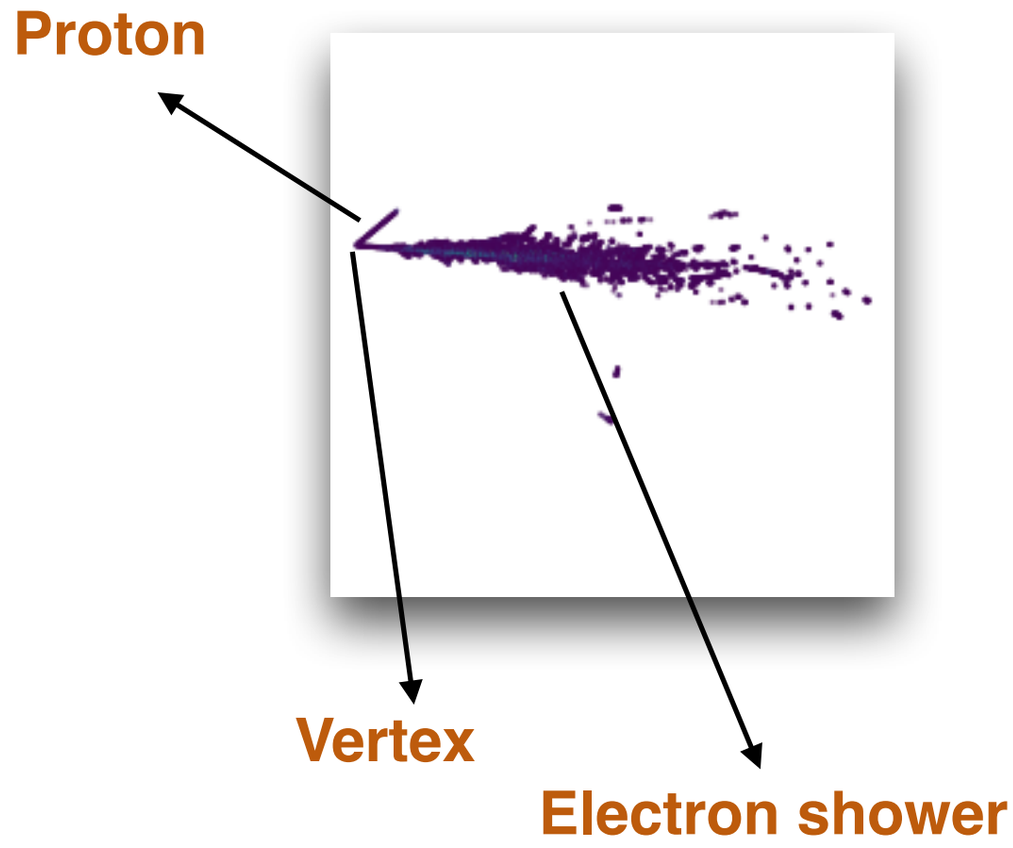


Object Detection

- Liquid Argon Time Projection Chamber (LAr-TPC) is an imaging calorimeter (mm resolution)

- **Objects in an Image** can be localised and classified

- **Background rejection, Vertex finding, Online Event Display** can be realised by using DL Object detection Method

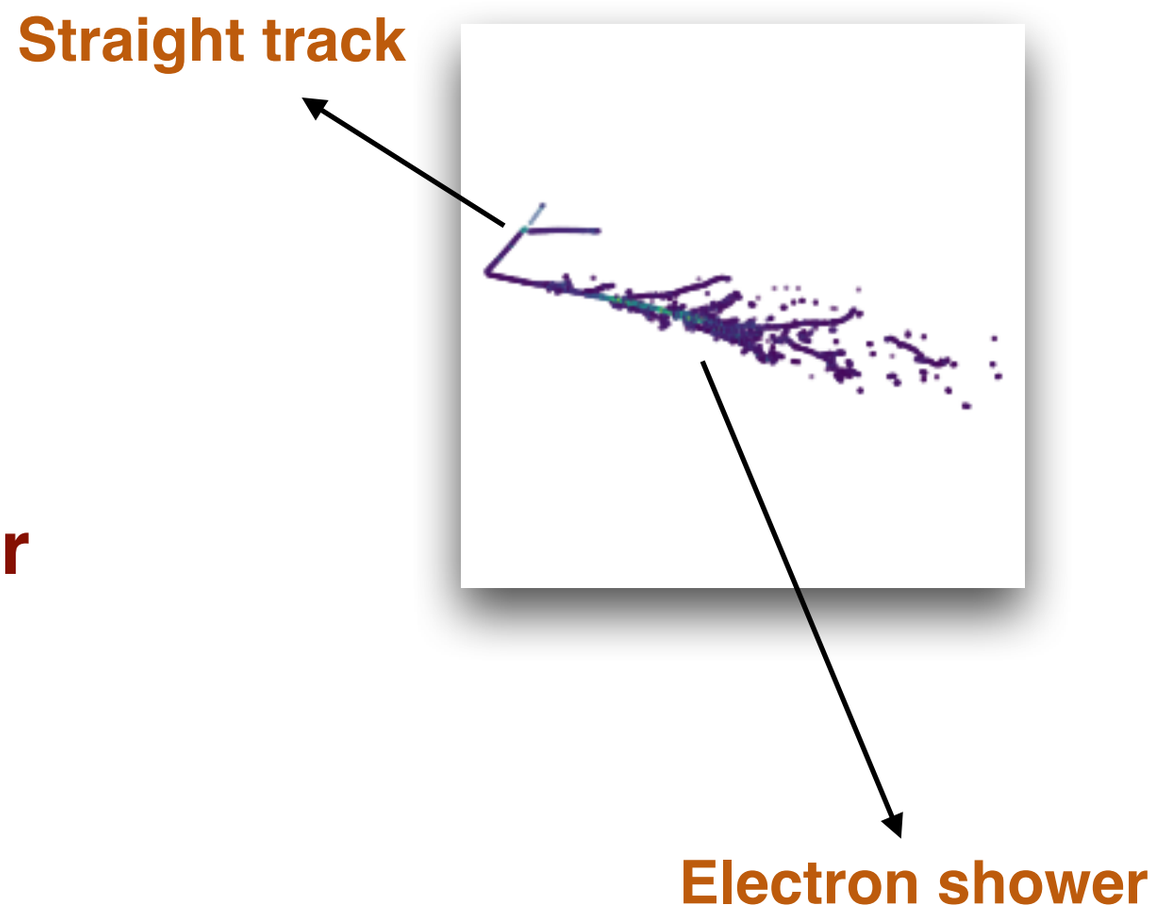


Semantic Segmentation

- Liquid Argon Time Projection Chamber (LAr-TPC) is an imaging calorimeter (mm resolution)

- Pixel level classification

- **Pixels belongings to shower and tracks** can be classified (clustering)



Generative Adversarial Network (GAN)

- Liquid Argon Time Projection Chamber (LAr-TPC) is an imaging calorimeter (mm resolution)

- Generative method instead of discriminative

- **Recovering events (fill dead channels)** and **Particle simulation** can be realised by using GAN

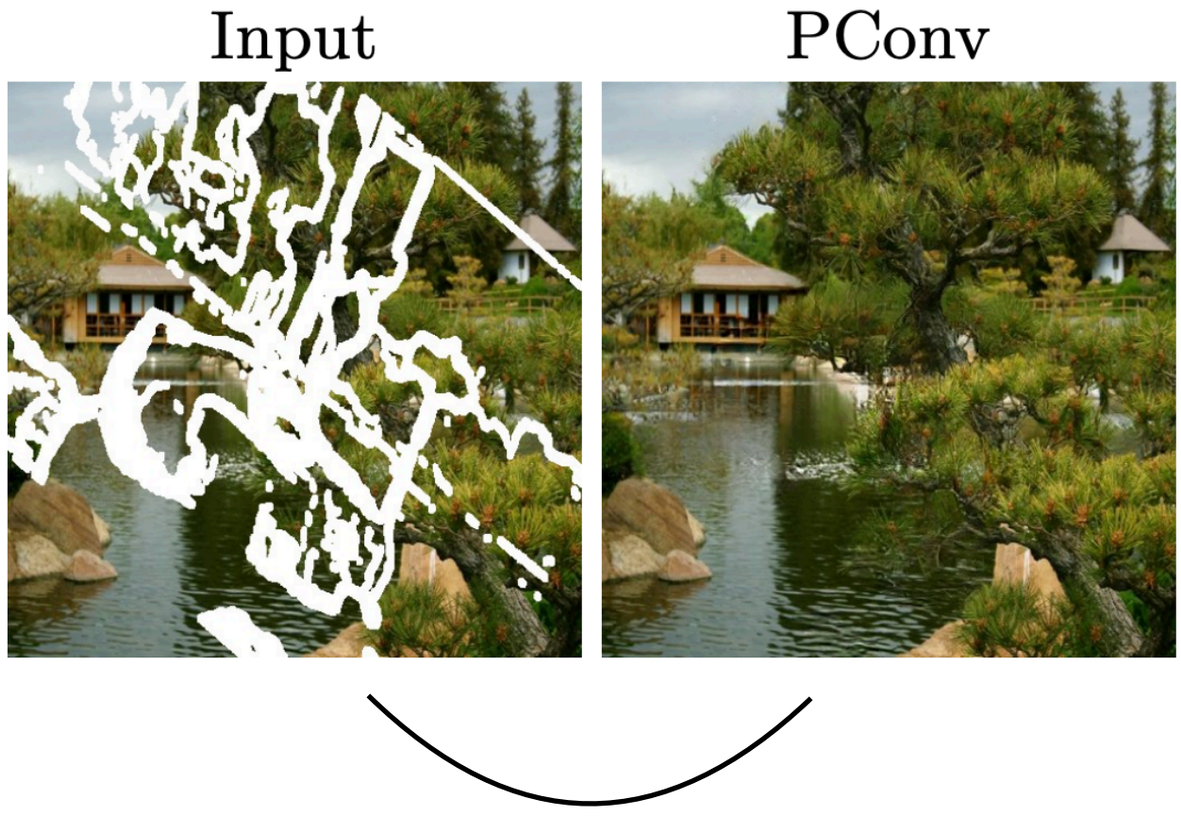


Image recovered by GAN (NVIDIA)

https://arxiv.org/pdf/1804.07723.pdf

Studies Done

- Image Classification method through a Deep Learning algorithm (ResNet50) has been applied to particle identification using DUNE MC events
- Image Optimisation has been completed
- **0.74 test accuracy** has been obtained with 7K images (5 particles classification)
    - NOvA: 0.69 with 4.7M images
    - MicroBooNE: 0.80 with 140K images
- The dataset will be increased and the study will be finalised
- I have started writing the studies done through overleaf as a technical note

Further Steps for Deep Learning

- Do Event Interaction classification by using Image Classification Method

- Make a research plan for the application of the other CNN methods

- Topical Lectures at Nikhef “Machine Learning” (4-6 April 2018) https://indico.nikhef.nl/event/1122/

- Search for collaboration opportunity with a close lab for the development of a CNN like (GAN) for Neutrino Field

AMLAB
Amsterdam
Machine Learning Lab


Max Welling

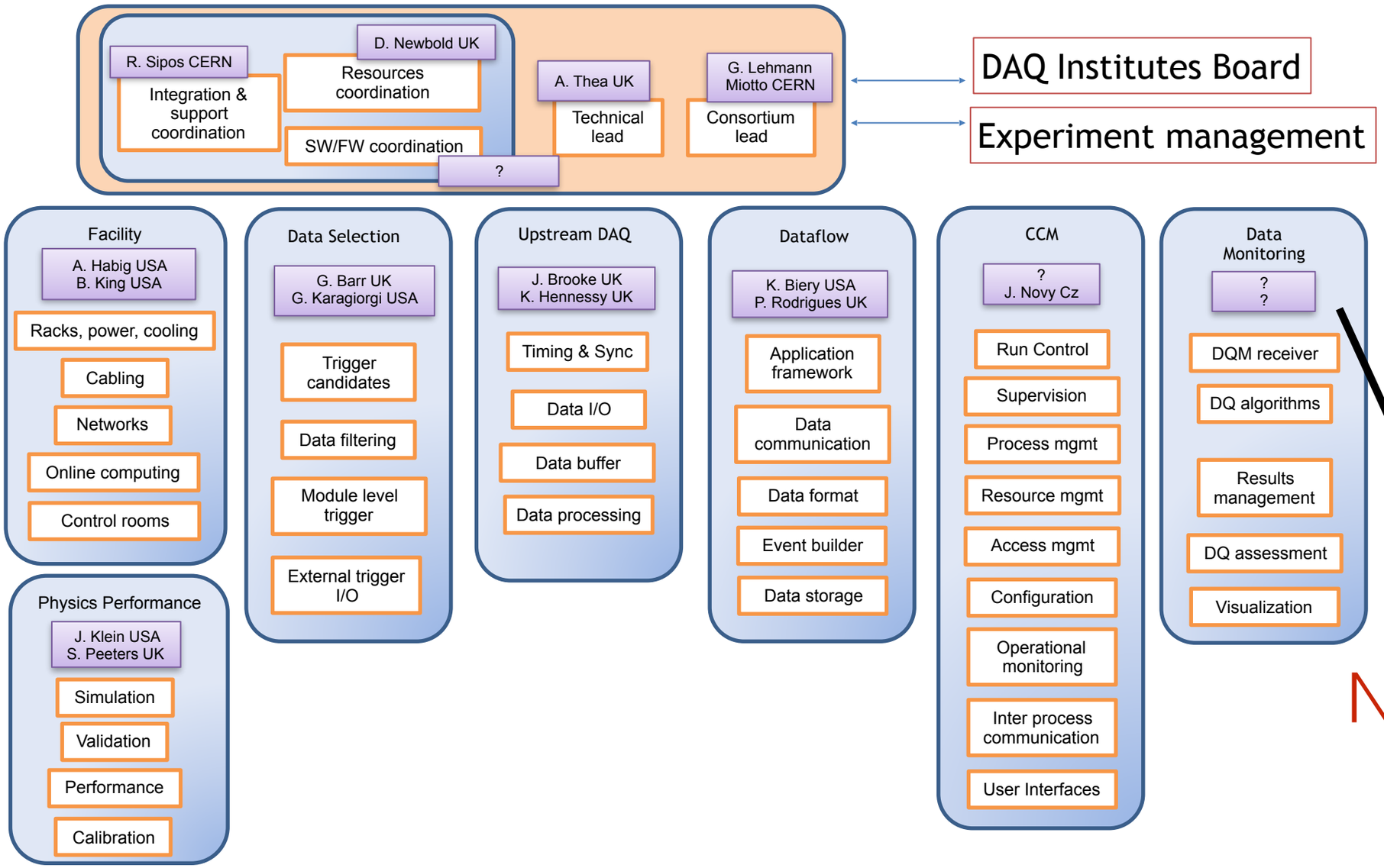

Joris Mooij


Zeynep Akata


Herke van Hoof

DAQ Issues

DAQ Consortium Organization



Giovanna's Talk - CERN

DQ: Data Quality
DQM: Data Quality Monitor

DUNE CDR Review - ProtoDUNE-SP Lessons Learned

- Monet (OM webviever) was too slow for LArTPC datasets. (Not enough for DUNE scales)

- OM needs to be more than “data viewing” - it should know about slow control and run control states, and beam info

- Some users found the OM based on LarSoft very tricky to maintain and inefficient - some fresh ideas are needed here.

- Best effort doesn’t work. It needs official status and groups responsible.

Karol’s talk - Liverpool

What can Nikhef do?

- Take the responsibility of Data Monitoring WG.
- First step is to study lessons learned by ProtoDUNE, MicroBooNE, etc by making a survey? and/or by discussing with related people face to face, email exchange, etc.
- Take comments of the experts from the Collaborations (DUNE, MicroBooNE, ProtoDUNE).
- Specify what we need as a Data Monitoring and make a research plan including motivation, objectives, work packages, milestones, team, etc. Take the feedback for the final draft.
- Find one or two highly skilled software engineers dedicated to this study and coordinate the work (50 Institutes in the DUNE collaboration).
- Find the most modern way to monitor DUNE data by the end users.
- Hold weekly meetings.
- Complete work packages on time. Do permanence checks. Prepare easily understandable manuals at the end for the safety or DAQ shifters, etc.

Road Map

- Complete Particle Classification and its technical note

- Present it in the collaboration soon

- Start neutrino interaction classification

- Search for other CNN methods and any collaboration opportunity to develop our own CNN model

- Continue DAQ studies on behalf of Nikhef

BACKUP

DL Studies in Neutrino Field

Experiment	Aim	CNN Method	DL Model (s)	Dataset	DL Software Framework	Results
NOvA [2]	Neutrino Event Classification	Image Classification	Convolutional Visual Network (CVN) inspired by GoogLeNet	2D Images: 4.7 events (80% Train and 20% Test)	-Caffe for training - Training on NVIDIA Tesla K40s	Test accuracy: ~0.69
MicroBooNE [3]	Single particle and neutrino event classification	Image Classification and Object Detection	AlexNet, GoogLeNet, Faster-RCNN, Inception-ResNet-v2, ResNet	2D Images: 22,000 (train) and 3,800 (test) events per particle	Caffe for CNN training and Analysis	Test accuracy: ~0.8
MicroBooNE [4]	track/shower separation at the pixel level	Image classification and semantic segmentation via Transfer Learning	U-ResNet, a hybrid of U-Net and ResNet	2D and 3D images: 140,000 events in total (100,000 train, 20,000 test and 20,000 validation)	-Caffe -NVIDIA TitanX	Incorrectly Classified Particle Fraction (ICPF) = 1.9% avarage
Our analysis in DUNE	π^0 separation through Single particle and neutrino event classification	Image Classification via Transfer Learning	Resnet50 NasNet more??	2D Images with Dataset??	Google Colab for training and analysis (GPU based training)	???

[2] JINST 11 (2016) no.09, P09001 arXiv:1604.01444 [hep-ex] FERMILAB-PUB-16-082-ND

[3] JINST 12 (2017) no.03, P03011 FERMILAB-PUB-16-538-ND DOI:10.1088/1748-0221/12/03/P03011 e-Print: arXiv:1611.05531

[4] FERMILAB-PUB-18-231-ND e-Print: arXiv:1808.07269