

End-to-end Deep Learning Reconstruction for CMS Experiment

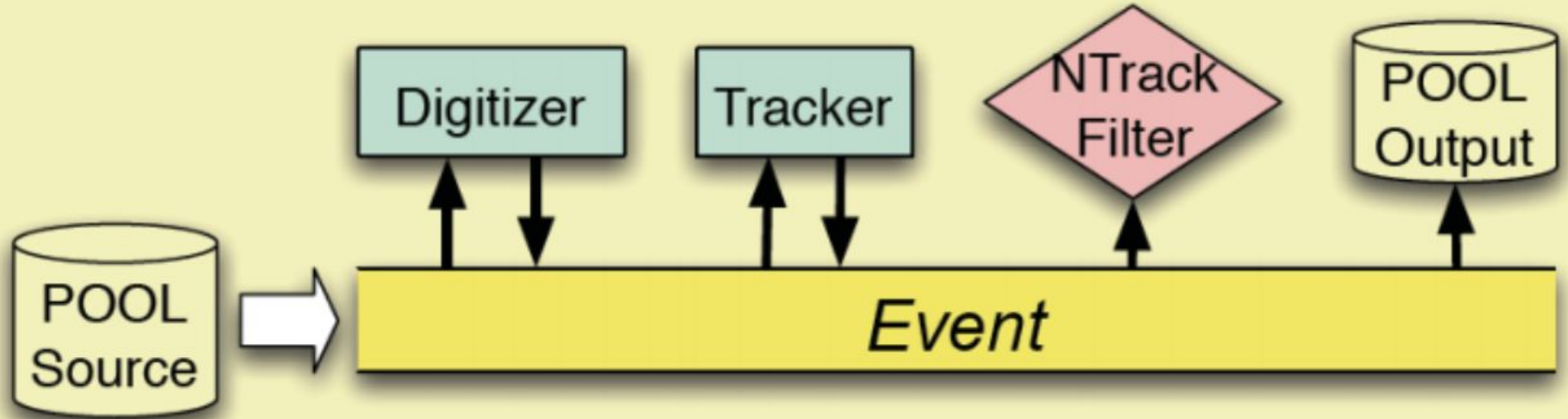
A Google Summer of Code 2020 Project

Mentors:

Sergei Gleyzer
Michael Andrews
Emanuele Usai

- Shravan Chaudhari (shravan25.chaudhari@gmail.com)

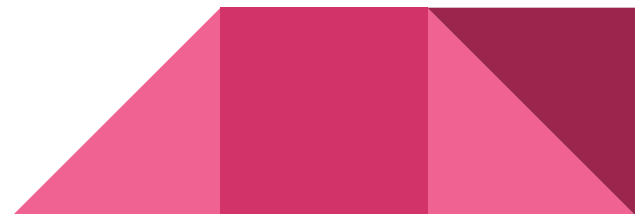
General Data flow in CMSSW:



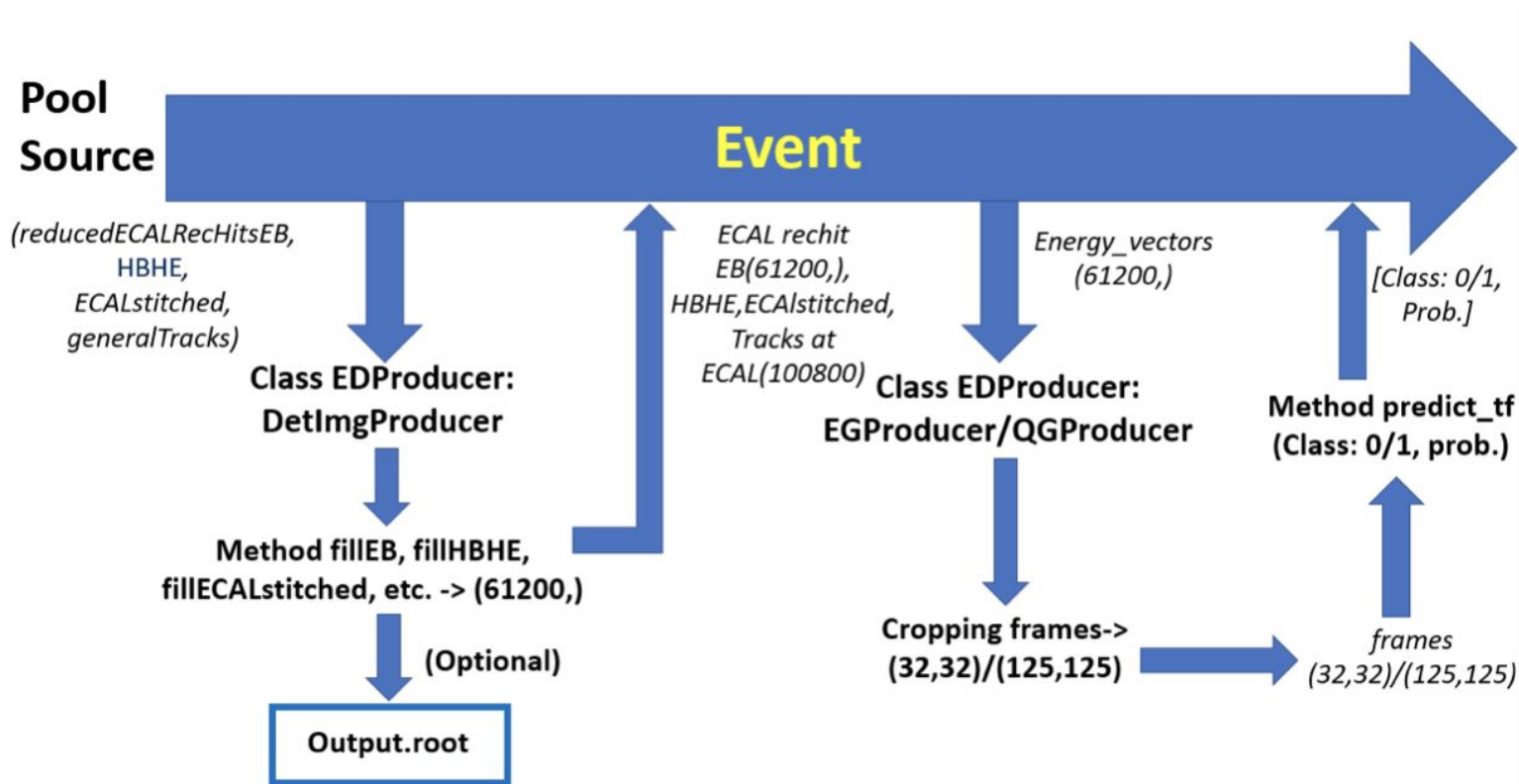
Objectives:

- Reading Detector Images and storing the detected Image frames in the production file (edm root file).
- Extracting Jet seeds Coordinates
- Preparing the frames for inference based on the seeds
- Running the inference and storing the predictions.
- Creating the appropriate configuration files.

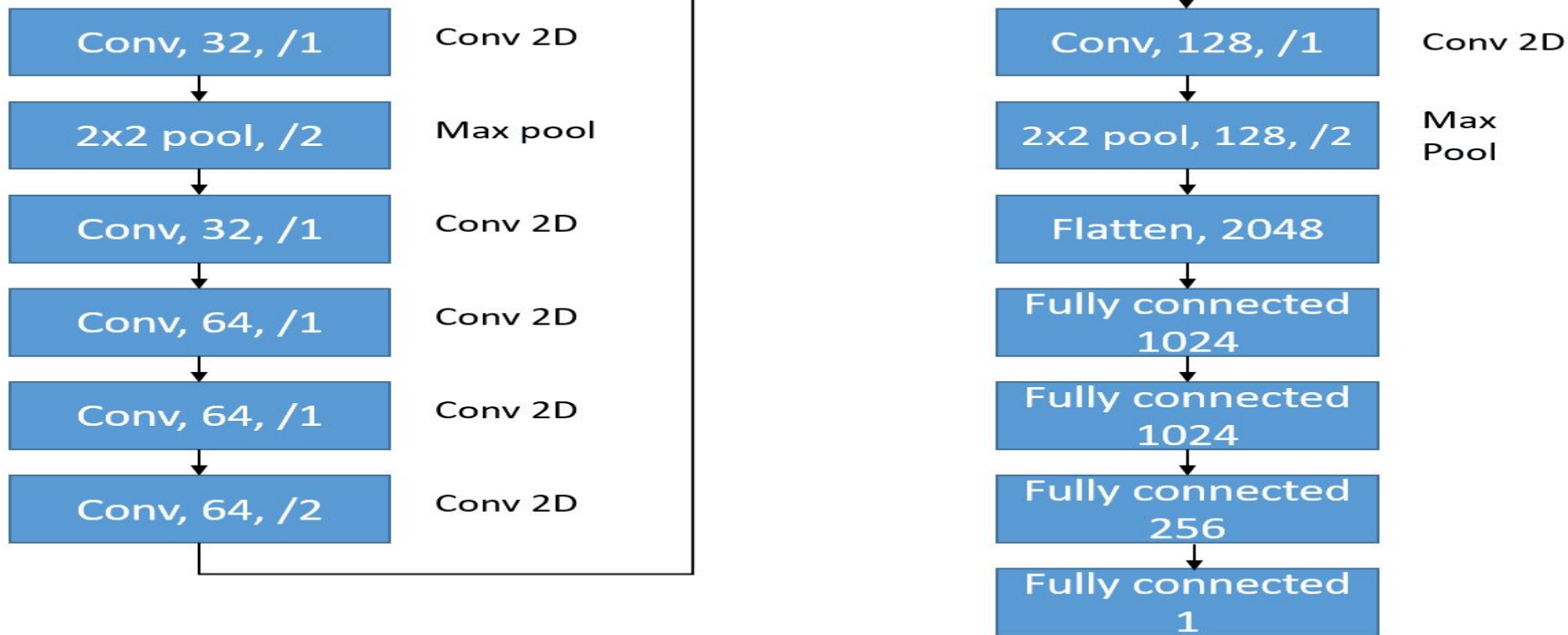
(Similar approach for all the three tasks: Electron-Photon, Quark-Gluon, Top Quarks)



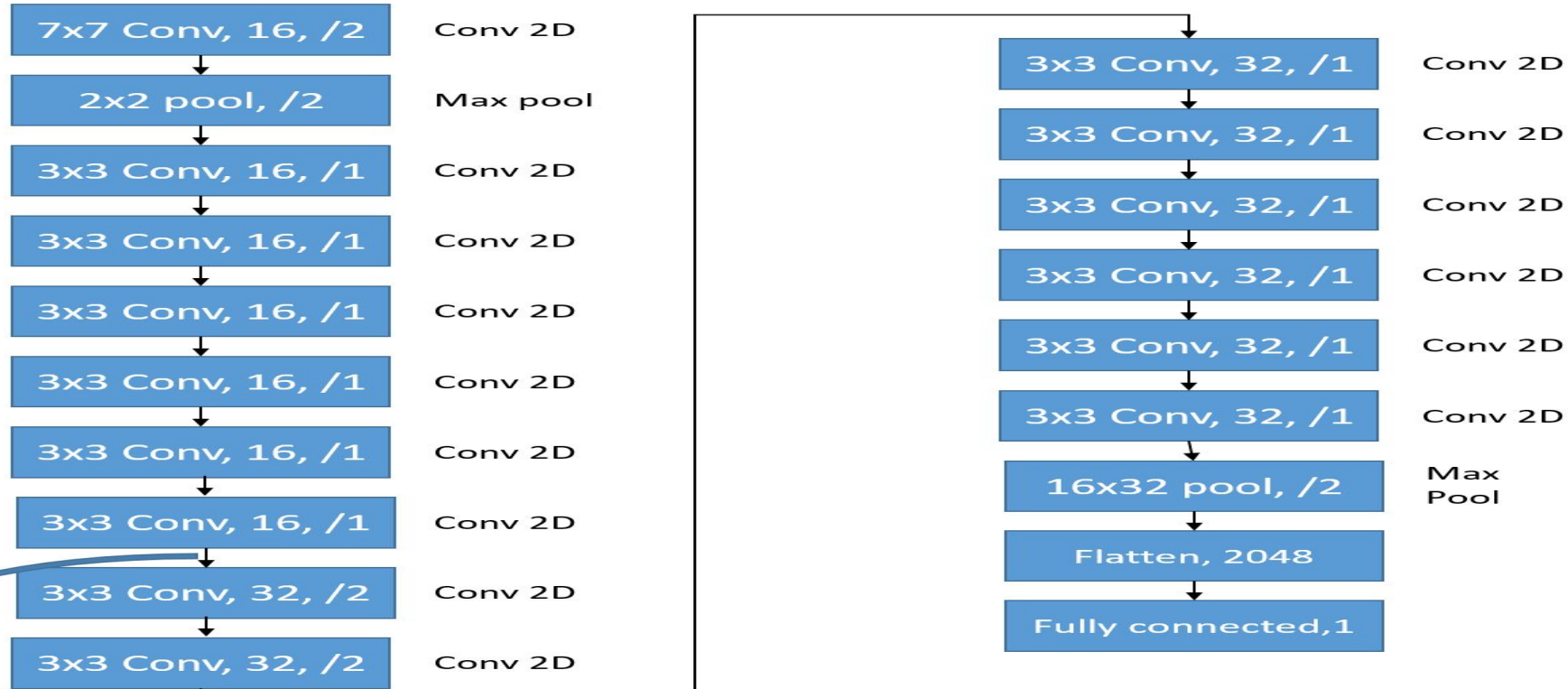
E2e Workflow:



Architecture: Electron Vs Photon (variant of VGG)



Architecture: Top Quarks and Quark vs Gluon (inspired from Resnet)



Executing the code:

```
$ git clone https://github.com/Shra1-25/E2eDL.git
```

```
$ scram b
```

```
$ cmsRun E2eDL/E2eDLrec/python/E2eDLrec_cfg.py \  
> inputFiles=file:[file_location] \  
> maxEvents=100 \  
> EGModelName=[e_vs_ph_model_name.pb] \  
> QGModelName=[q_vs_g_model_name.pb] \  
> TopQuarksModelName=[topq_model_name.pb]
```

* Note: Before running cmsRun command please ensure that all the tensorflow model pb files (pref: 1.6.0) are present inside the E2eDL folder.

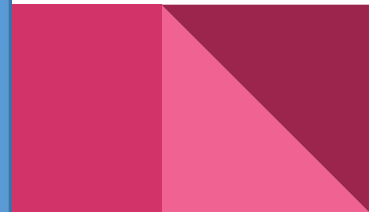


Inference Timing results(on a CPU):

No. of Events	Minimum time for an event (+/- 1sec)	Maximum time for an event (+/- 1sec)	Average time (sec)	Total job time (sec)	Event throughput (event/sec)
5	0.3	6.5	2.9	39.8	0.3
10	0.1	5.5	1.4	36.8	0.7
100	0.1	5.3	0.6	86.4	1.6
300	0.1	3.5	0.5	160.2	1.8

Inference Memory Results:

No. of Events	Peak virtual size (Mb)	Total job/ record (Mb)
5	2897	23.2
10	2897	24.6
100	6555	78.4
300	12199	163.6



Thank You!!

