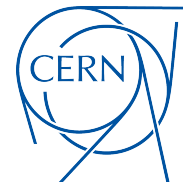


CVN Update

Saul Alonso-Monsalve, Alex Radovic, Leigh Whitehead

LBL Meeting

12 March 2018



Content

- Last CVN state (January 2018)
- Moving forward
- Data generation
- Update
- Conclusion

Last CVN state (January 2018)

- CVN (Convolutional Visual Network):
 - Deep Learning Framework: Caffe (developed by Berkeley AI Research).
 - CVN inspired by the GoogleNet (arXiv:1409.4842) CNN architecture.
 - Trained on the GPUs in the Wilson cluster.
 - Performance on train and test samples similar.
 - ~72% test accuracy.
 - The most time-consuming part of the training is the conversion of inputs into the levelDB format.
 - LevelDB is the format Caffe requires for the network inputs.

Moving forward

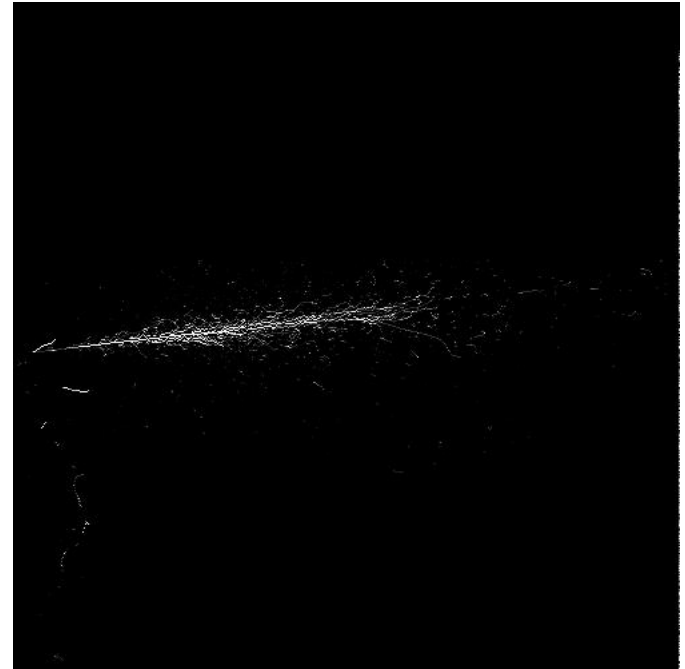
- Avoid the conversion of the inputs to levelDB format.
- Move from Caffe to Tensorflow (developed by Google Brain Team) for ease of design for more complex network topologies.
 - Tensorflow is also included in LArSoft.
 - Moving to Tensorflow also eliminates the levelDB step.
- Try to identify ways to improve both the network architecture and the data generation.

Data generation

- We regenerated all the data without the levelDB step.
 - It was six times faster (we were able to perform the generation in parallel using six nodes) than with the levelDB step (that cannot be performed in parallel).
- We found an overflow error in the generation of the images.
- Pixel range was from -128 to 127 instead of from 0 to 255.
- The brightest (highest charge) pixels (with a brightness value > 127) cannot be appreciated.
- We have regenerated all the images after correcting the overflow error.

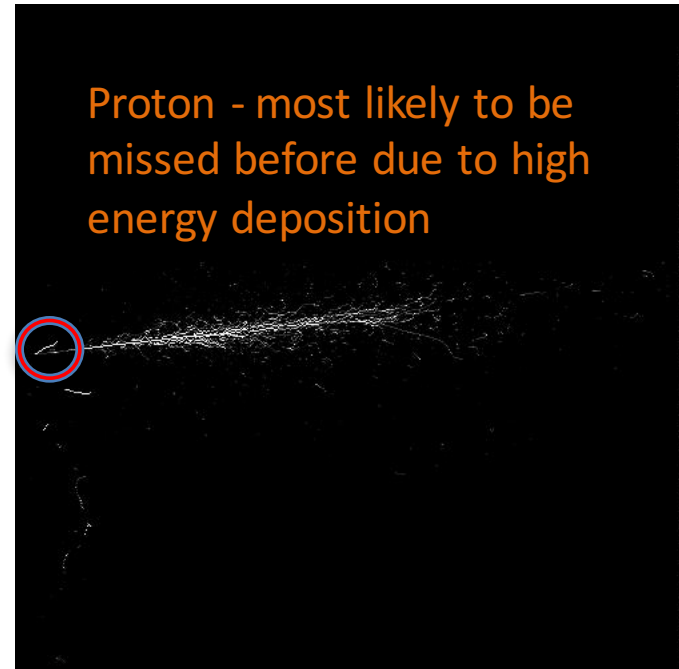
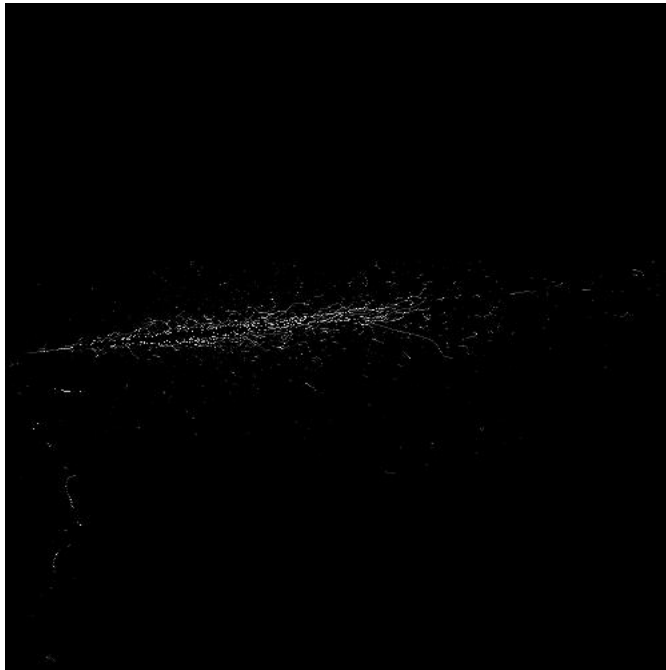
Data generation

- Example electron antineutrino event:
 - Before (pixel values from -128 to 127).
 - Now (pixel values from 0 to 255).



Data generation

- Example electron antineutrino event:
 - Before (pixel values from -128 to 127).
 - Now (pixel values from 0 to 255).

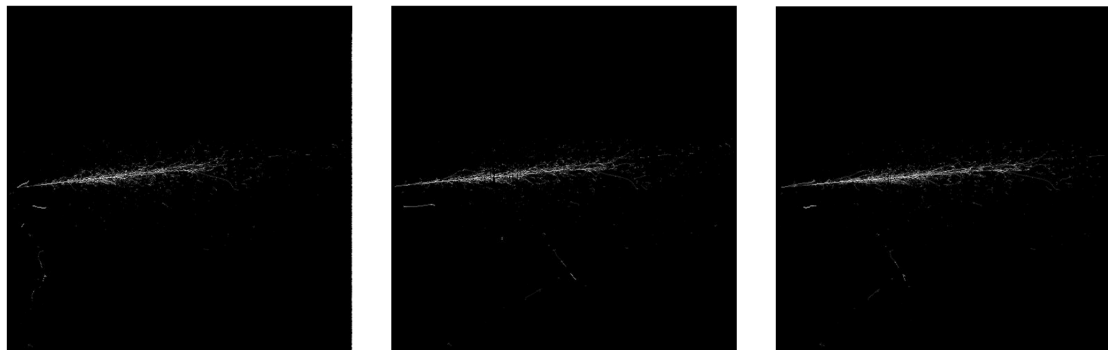


Data generation

- Same event, three views.
 - Before (pixel values from -128 to 127):



- After (pixel values from 0 to 255):

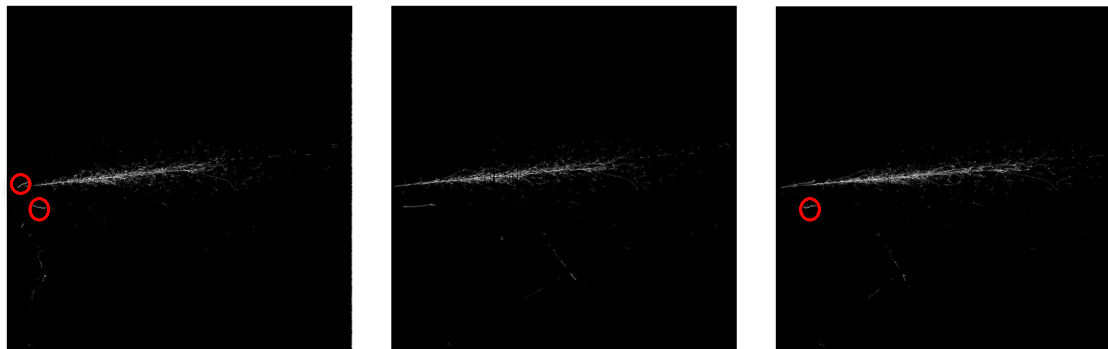


Data generation

- Same event, three views.
 - Before (pixel values from -128 to 127):



- After (pixel values from 0 to 255):



Updates

- This issue might be damaging the CVN performance.
 - The neural network would find difficulties in detecting some relevant features of the images.
- We are designing a new Tensorflow network that will use the regenerated data.
 - We already have a small network prototype that is able to get a 68% accuracy with only 9,000 images! (the entire dataset contains millions of images).
- We are also about to retrain the old Caffe network with the regenerated data.
 - We just need to regenerate all the levelDB files.
 - This will allow for a clear comparison of performance.

Conclusion

- Faster data generation
 - Old data generation method (with levelDB step) takes ~ one week of computation, while the new method (without levelDB step) can perform in parallel (it took only one day of computation using six nodes).
- Overflow error solved
 - We believe that this fact may improve the performance of the CVN.
- Optimistic moving to Tensorflow
 - It is true that we trained our Tensorflow network with a small subset of the whole dataset (~9,000 vs a few million images), but our Tensorflow got almost the same test accuracy as the old CVN (68% vs 72%).
- We expect to have new results in the following days.