



Contribution ID: 392

Type: poster presentation

## Acceleration of ensemble machine learning methods using many-core devices

Multivariate training and classification methods using machine learning techniques are commonly applied in data analysis at HEP experiments. Despite their success in looking for signatures of new physics beyond the standard model it is known that some of these techniques are computationally bound when input sample size and model complexity are increased. Investigating opportunities for potential performance improvements is therefore of great importance if these techniques are to be used with the much larger data volumes expected from Run 2 operations at the Large Hadron Collider.

It has been previously shown that a large degree of algorithm parallelisation was possible for MLP-based artificial neural networks by the use of many-core devices such as GPUs. Improved scaling was observed in network complexity and qualitative performance gains were attainable through the simultaneous processing of multiple neural networks. Here we investigate how many-core devices can be used to accelerate ensemble machine learning methods that are gaining traction in HEP data analysis.

We present a case study into the acceleration of decision forests using many-core devices in collaboration with Toshiba Medical Visualisation Systems Europe (TMVSE). TMVSE have developed software to process three-dimensional medical imaging data (such as CT or MRI scans), using automatic detection of anatomical landmarks defined on the skeleton, vasculature and major organs. Landmark detection underpins a semantic understanding of the medical data and thus has many diverse applications, for example, it facilitates rapid navigation to a named organ. TMVSE have applied ensemble machine learning methods such as classification by random decision forests to efficiently compute the bounding boxes of organs in processed image data volumes. It is important that their applications using this algorithm runs efficiently and quickly. After data preparation and optimisation the execution time is on average 4.5 seconds per volume with a sub-second processing time being desirable.

Using representative medical image data as input and pre-trained decision trees we will demonstrate how the decision forest classification method maps onto the GPU data processing model. It was found that a GPU-based version of the classification method resulted in over 130 times speed-up over a single-threaded CPU implementation with further improvements possible. We will outline the main optimisation steps undertaken to maximise GPU performance and detail how this was implemented using device profiling to evaluate thread occupancy and execution efficiency.

As this solution was developed to be context independent we will demonstrate how this work can be applied to a suitably formed HEP dataset to determine potential gains in event throughput and classifier discrimination. We will also explore how the advanced analysis techniques applied to automatic landmark detection in medical data can be applied to HEP dataset to achieve further increases in performance.

**Primary author:** WASHBROOK, Andrew John (University of Edinburgh (GB))

**Co-author:** Mr DANIEL, Wyeth (Toshiba Medical Visualisation Systems Europe (TMVSE))

**Presenter:** WASHBROOK, Andrew John (University of Edinburgh (GB))

**Track Classification:** Track8: Performance increase and optimization exploiting hardware features