Enabling Artificial Intelligence On-Board for Image Processing Applications

Pablo Ghiglino, Mandar Harshe

Klepsydra Technologies AG

On-Board Image Processing Using AI

A novel approach to on-board artificial intelligence (AI) is presented here that is based on state-of-the-art academic research of the well known technique of data pipeline. Algorithm pipelining has seen a resurgence in the high performance computing work due its low power use and high throughput capabilities. The approach presented here provides a very sophisticated threading model combination of pipeline and parallelization techniques applied to deep neural networks (DNN), making these type of AI applications much more efficient and reliable. This new approach has been validated with several DNN models developed for Space application (including asteroid landing, cloud detection and coronal mass ejection detection) and two different computer architectures. The results show that the data processing rate and power saving of the applications increase substantially with respect to standard AI solutions, enabling real AI on space.

Use Case: Earth Observation

Earth observation images are not useful when there is cloud cover obstructing the view.

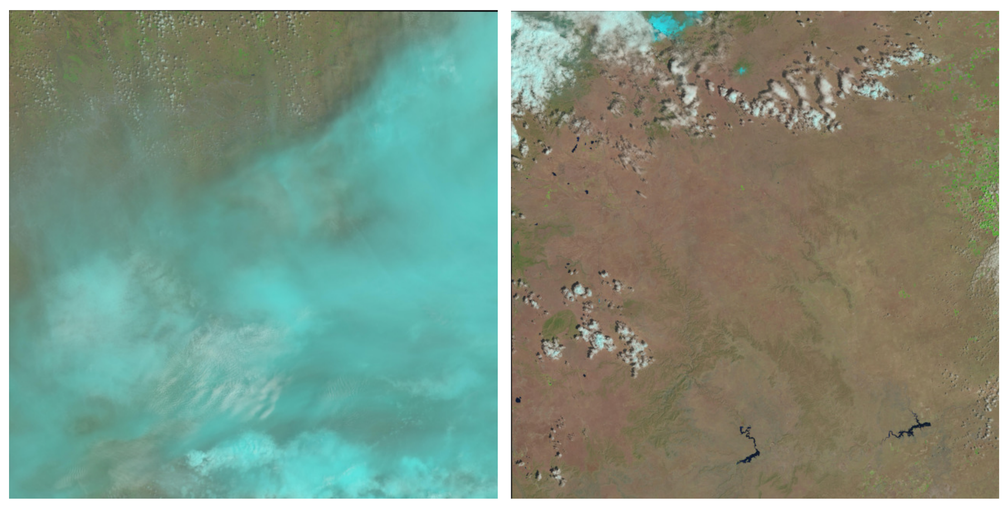


Fig. 1: Cloud cover makes images unusable. Image on the left conveys very little information compared to image on the right

Image Segmentation allows us to segment the image into visible areas & areas covered by clouds.

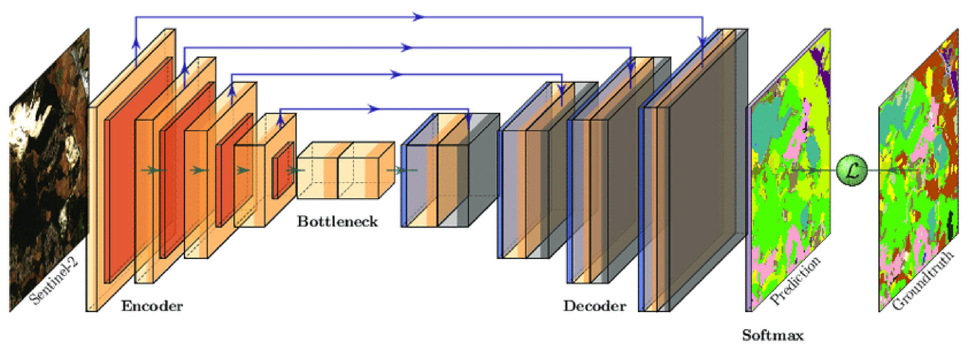


Fig. 2: A U-Net based DNN could identify cloud cover on satellite itself

Segmentation tasks bring in multiple challenges:

- Convolutions & Deconvolutions are slow on CPUs
- Large number of parameters & intermediate data
- Skip connections - Intermediate data has to be stored

This problem is not unique to segmentation tasks. Modern DNNs have multiple skip connections, large number of layers and expect GPU availability for fast predictions.

Klepsydra AI Building blocks

In order to address performance issues faced by tasks like Segmentation, Object Detection & Classification, and Regression, Klepsydra uses an innovative approach to Parallelization. Klepsydra AI incorporates a 2 dimensional threading model along with event loops.

- Lock-free EventLoops to connect successive layers.
- FPU vectorisation to accelerate the matrix multiplications.
- One EventLoop per thread - each layer operation in one thread, with optional parallelisation.

The two dimensions of the threading model are :

Pipelining

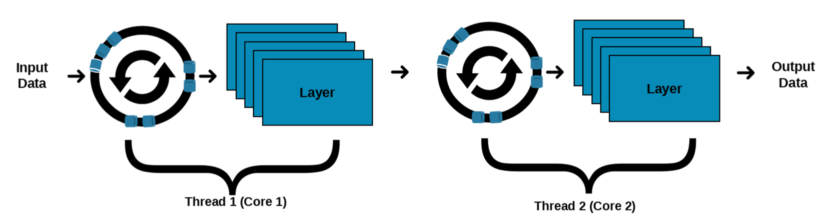


Fig. 3: Pipelining applied to layers of Deep Neural Network

Parallelization

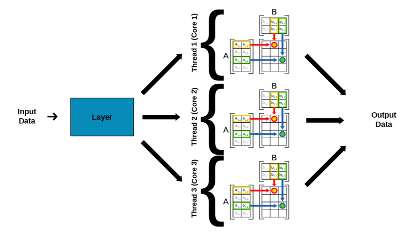


Fig. 4: Second-dimension: Parallelization in each Event Loop

The 2D threading allows configuration of various parameters:

- Internal Memory Buffer size
- Number of event loops
- CPU core distribution
- Layer grouping

Performance "Spectrum"

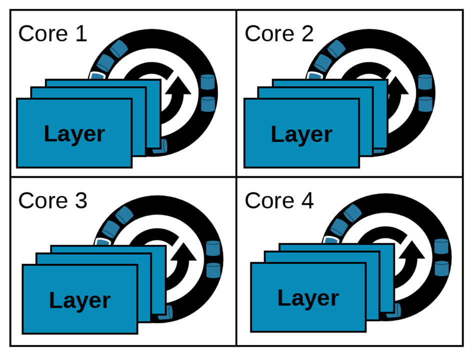


Fig. 5: Full Parallelization

- Low CPU use
- High throughput
- High latency

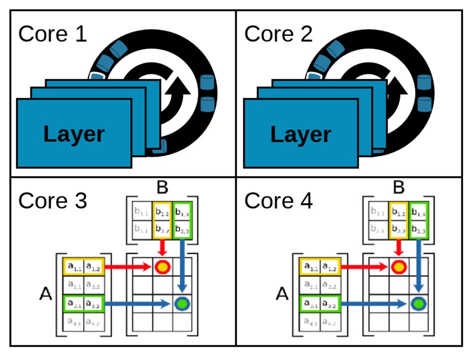


Fig. 6: Balanced

- Mid CPU use
- Mid throughput
- Mid latency

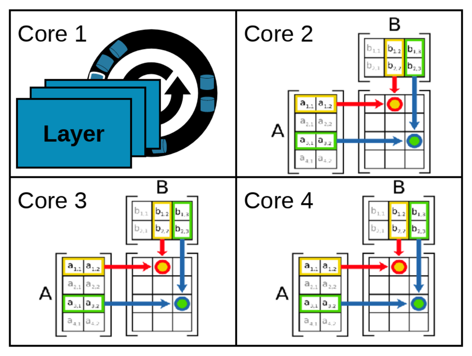


Fig. 7: Full Pipelining

- High CPU use
- Mid throughput
- Low latency

API & Configuration

```
auto dnnFactory = createDNNFactory(modelFile, configFile);
auto dnn = dnnFactory->getDNN();
dnn->setCallback(processPredictionResults);
dnn->predict(NewImage);
```

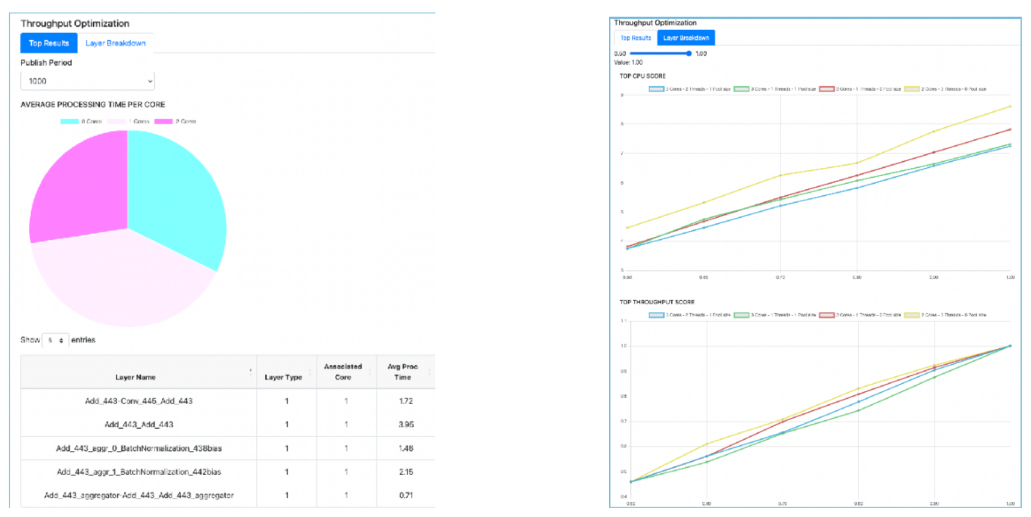


Fig. 8: Optimizer tool to get tuned configuration

KATESU: Testing Performance on Space Qualified Boards

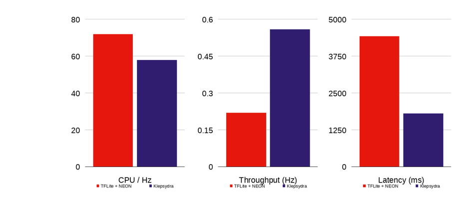


Fig. 9: Performance of a UNet on a LS1046

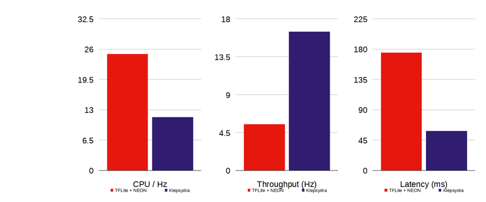


Fig. 10: Performance of convolution net On LS1046

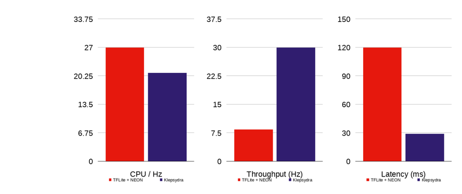


Fig. 11: Performance of Quantized convolution net on LS1046

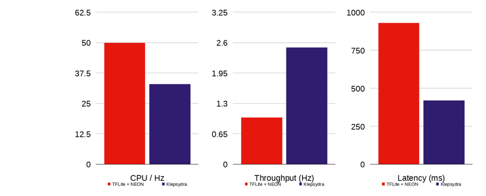
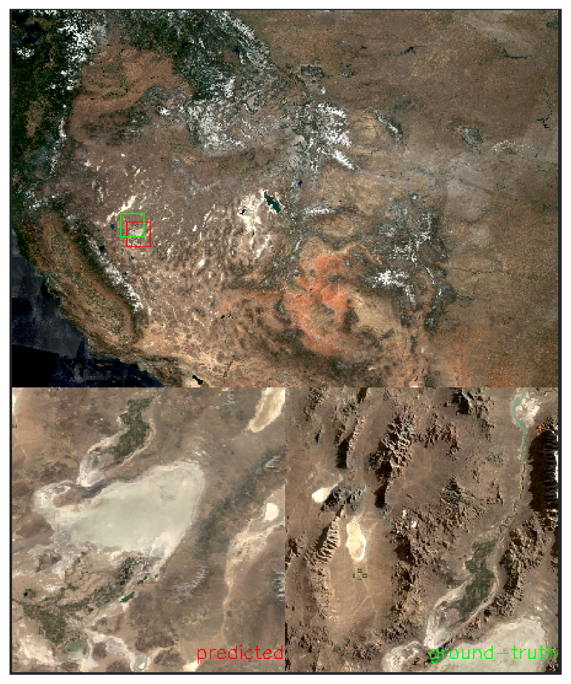
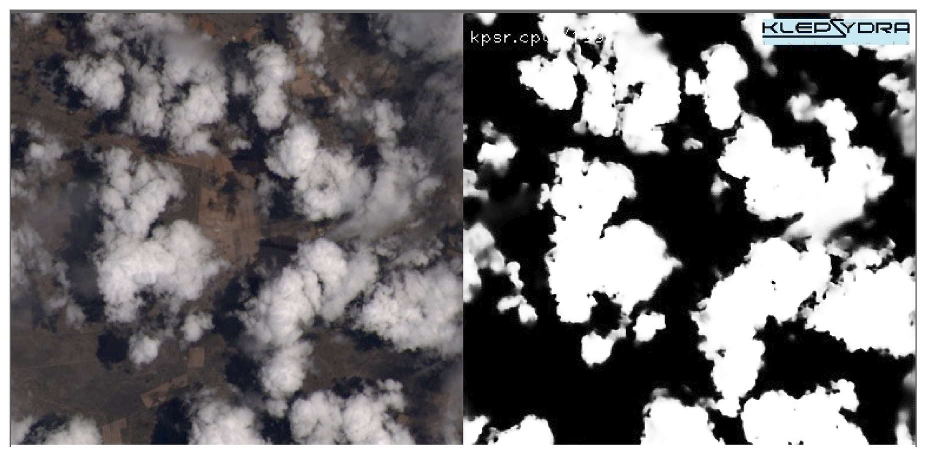


Fig. 12: Performance of Quantized convolution net on Zedboard

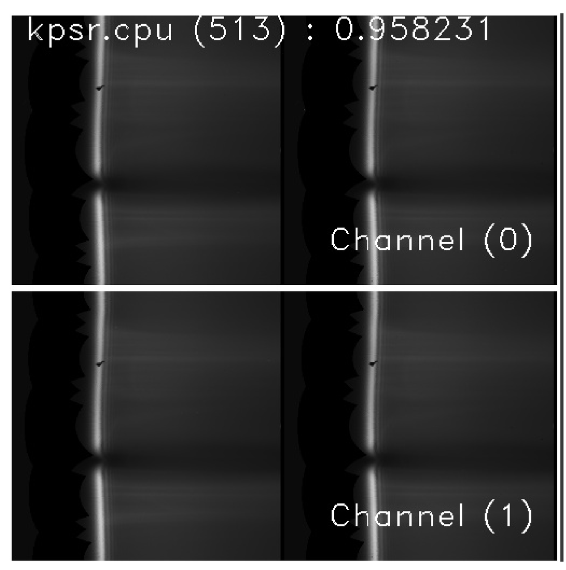
Other On Board AI Applications



GPS sensors can be augmented with image based localization, allowing faster and more accurate geo localization of satellites. GPS can provide coarse localization information to narrow the field of search for image based localization.



On board cloud detection allows quickly detecting the masked and visible regions. A masked image could be discarded, thus saving valuable storage space and reducing downlinking costs.



Solar Coronal Mass Ejections (CME) can cause geomagnetic storms and severely damage electronic equipment and power grids. Early detection of CMEs is therefore crucial for damage containment.

Live demos

See Klepsydra AI being used for the above applications:
www.klepsydra.com/demos