DELTATEC

On-board Deep-Learning for semantic imaging data reduction

MOTIVATIONS

Requirements:

- More precise, hyperspectral observations
- Needs for increased re-visit rates
- Cost reduction

Proposed solution:

- Fleets of small satellites (cubesats)

Barriers:

- Small downlink bandwidth available
- Reduced on-board power availability

ASTERIA space telescope (source NASA)

PURPOSE

Downlink bandwidth reduction techniques with huge compression rates (i.e. factor 100)

- Generate noise in data
  - De-noising is not recovering data
- Is not content aware
  - Bandwidth consumption for application non relevant data

On-board AI can enable

- Transmission of semantic data only
  - Maps
  - Event or« object » detection
- Content aware compression control
  - Adapt compression scheme to content and its application relevance
- Content aware transmission control
  - Transmit data only for relevant areas

Image par Steve Buissinne (Pixabay)

DEEP-LEARNING FOR ON-BOARD IMAGE PROCESSING

NNs are now the best solution on ground for image segmentation, classification and object detection.

NICE !

- CNN are scalable and have a data flow structure
  - Well suited for progressive image acquisition
- NNs are predictable in time
  - NN execution does not vary as a function of the content
  - Nice property for RT system design

NOT NICE !

- NNs are computing intensive processes…
  - Not compatible with current generation on-board computing devices

AI ON-BOARD BECOME FEASIBLE

New generations of SoCs in embedded world: VCU – TPU – FPGA

- Low power
- Huge computing capabilities per Watt and cm^3

Qualification / hardnessing of some SoCs is on-going:

- Xilinx Zynq 7000 on-board computers in production phase
- Intel-Movidius Myriad 2 radiation tests by ESA and NASA

Significant on-board computing capabilities should become available in short future for small platforms

« Significant » does not mean « huge » !

Image: Intel news room

CHALLENGE 1: DESIGN METHODOLOGY

Resources are sparse and data flow huge:

- We need to maximize throughput and computing efficiency

Setup a design methodology for light weight network architectures

- Deep-learning looks often « magic » but is not
- Avoid reusing general purpose models directly
- Key is understanding designed networks behavior

Develop network reduction techniques

- Network pruning – Transfer learning
- Network quantization
  - Myriad 2 = int 8
  - FPGA => strong quantization welcome (i.e. binary networks)

Image Pexels (Pixabay)

CHALLENGE 2: DATA AVAILABILITY

Data is the key for the design of any deep-learning system

- Availability of **<u>annotated</u>** data is to be questioned
- Even for simple network, having very large database is the key for success

Techniques for data augmentation

- Building modified or assembled pictures
- Imply characterization of what is to be detected

Building synthetic training data sets

- Full synthetic pictures for generic tasks like compression control
    - Need characterization of features
- Augmentation of original pictures by synthesis (for instance adding clouds or picture defects)
- Style transfer networks and GANs for realistic database generation

Image source NASA

CONCLUSIONS

Space and on-ground embedded applications share the same technical challenges

- Use of dedicated low power devices
- Light network design and design reduction techniques
- Database augmentation
- Synthetic database generation

Deep Learning is a privileged domain for ground to in-flight technology transfer

Push on AI dedicated devices in the embedded world will push on-board technologies in near future

Image Gerd Altmann (Pixabay)

CONCLUSIONS (CONTINUED)

Deep-Learning is only a part of the solution

- SoCs like Zinq or Myriad 2 are not purely dedicated to Deep-Learning
- They are powerful devices for any « scientific » computing
- Significant processing will move from ground to on-board to optimize resource usage.

We are entering a ground to on-board transition period

Image Piro4D (Pixabay)