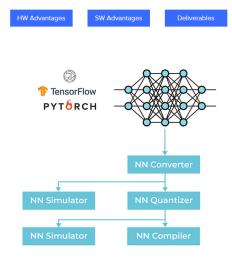


$\mathbf{ENLIGHT}^{\mathsf{TM}}$

Performs various operations of deep neural networks such as convolution, pooling, and non-linear activation functions for the edge computing environments.



It has a solid competitive edge in delivering unparallel compute density with energy efficiency (power, performance, and area)

Hardware Key Advantages

Mixed Precision (4-/8-bit) Computation

 $\bullet\,$ Higher efficiency in PPAs (power, performance, and area), DRAM bandwidth

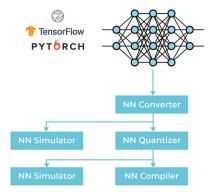
DNN-optimized Vector Engine

• Better adaptation to future DNN changes

Scale-out with Multi-core

Document title: NPU | OPENEDGES Technology Capture URL: https://www.openedges.com/npu

Capture timestamp (UTC): Tue, 01 Mar 2022 22:06:44 GMT



It has a solid competitive edge in delivering unparallel compute density with energy efficiency (power, performance, and area)

Hardware Key Advantages

Mixed Precision (4-/8-bit) Computation

• Higher efficiency in PPAs (power, performance, and area), DRAM bandwidth

DNN-optimized Vector Engine

• Better adaptation to future DNN changes

Scale-out with Multi-core

• Even higher performance by parallel processing of DNN layers

Modern DNN Algorithm Support

• Depth-wise convolution, feature pyramid network (FPN), swish/mish activation, etc.

Software Key Advantages

High-level Inter-layer Optimization

 Grouped layer partitioning and scheduling for reducing DRAM traffic from intermediate data

DNN-layers Parallelization

- Efficiently utilize multi-core resources for higher performance
- Optimize data movements among cores

Aggressive Quantization

• Maximize use of 4-bit computation capability

Deliverables

ENLIGHT Toolkit is available to all eligible companies with the following items:

- RTL design for synthesis
- User guide
- Integration guide

Contact Products Investors News Careers Location Privacy & Legal

© 2022 OPENEDGES Technology, Inc. . All Rights Reserved.

Document title: NPU | OPENEDGES Technology
Capture URL: https://www.openedges.com/npu
Capture timestamp (UTC): Tue, 01 Mar 2022 22:06:44 GMT