Artificial Intelligence Execution Accelerator Platform
We execute artificial intelligence faster

What is SoyNet :

Artificial Intelligence Execution Accelerator

Company: SOYNET Co., Ltd
Website: soynet.io
Country: Republic of Korea
Location: Hi Tech Campus, Daewangpangyo-ro 815, Sujeong-gu, Seongnam-si, Gyeonggi-do
Market: With data and domain knowledge, the company that operates artificial intelligence to resolve its own issue, or companies that want to enter into business, or who want to secure the performance of artificial intelligence services in Cloud

Profile
Started development of SoyNet from 2016 and established in 2018. Currently providing artificial intelligence execution accelerator platform

Various pilot projects using MLP, CNN, RNN models and the latest artificial intelligence models have been porting to SoyNet

Solution
SoyNet is an artificial intelligence execution accelerator platform developed by SOYNET Co., Ltd., which solves the performance issues in artificial intelligence service

Benefits
- Relentless AI execution under no internet service
- Low memory consumption, High speed performance guaranteed
- Increased DevOps productivity (No more obfuscated AI codes vs. simpler API to use SoyNet)

"We are introducing the latest AI based image analysis models to the fields and are conducting a project to classify surface defects of products through image analysis. However, when 100 images are input, only 10 images are processed, and 90 images are not processed due to its execution latency. So, while considering the introduction of additional high-specification GPU servers, fortunately, We have found SoyNet and solved performance issues. "

A Voice of Customer

What is AI execution Accelerator?

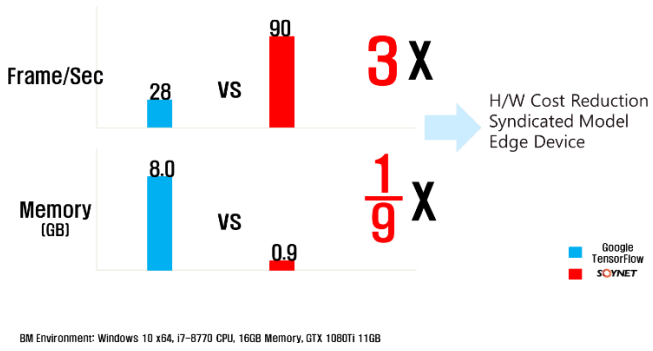
Recently, the introduction and utilization of artificial intelligence has been taking place in various industries very positively with the advent of the 4th Industrial Revolution. However, the performance is not supported at the actual application stage, and it has also been known as conundrum in edge device.

In general, global companies such as Google and NVIDIA provide an API repository called Model Zoo through TensorFlow, Caffe, and Pytorch. Artificial intelligence developers design and train artificial intelligence models using these APIs. As artificial intelligence is running on an integrated environment with training plus execution, modules of training engines that are not meaningful at runtime become heavy in memory, and custom

layers which not tuned up for GPU affect performance.

SoyNet takes the trained engines, such as model and weight parameters out of a general AI platform and can run them to get the same results 2 times or 50 times faster depending on models. Any model trained in general AI platform such as TensorFlow, Caffe, Pytorch can be configured in SoyNet. In addition, we applied proprietary technologies (Patented) to optimize the AI environment such as memory, CPU, and GPU and accelerated performance by converting all modules into CUDA, CUDNN, tensorRT and OpenCL.

SoyNet Performance



BM Environment: Windows 10 x64, i7-8770 CPU, 16GB Memory, GTX 1080Ti 11GB

In order to extract benchmarking data in the same environment, performance was measured with a CNN-based Object Detection Model called YOLOv3 in common.

BM Environment: OS(Windows 10 x64), CPU(i7-7700K), Mem(16GB), GPU(GTX1080TI 11GB)

Applied Model : YOLOv3-DarkNet53 (CNN Object Detection), Model Size : 416x416, Float32, TensorFlow 1.12 (TensorRT 4.0 Version)

One of the differentiating parts of SoyNet is the significantly lower memory usage than traditional AI platforms. (Noted that memory: GPU (0.9 GB vs. 8.0 GB). This gives the edge devices the opportunity to utilize artificial intelligence

In addition, it can be used as a general-purpose while FPGA type of accelerator can be used to a specific area with integrated H/W circuits limitation, thus SoyNet facilitates implementation convenience and various type of scale requirements, thereby ensuring higher ROI.

And while the normal artificial intelligence loads the unnecessary training engine in the execution environment, and takes a lot of boot time, while SoyNet booting time is very short and the service latency is very low. Additionally, some custom layers developed by modellers can utilize GPU thanks to the optimized forwarding layers of SoyNet. This makes much difference in terms of high-seed performance. As such, SoyNet has developed various custom layers to fully support GPUs.

Available AI Models with SoyNet

- AlexNet
- BNInception
- CaffeResNet101
- DenseNet121
- DenseNet161
- DenseNet169
- DenseNet201
- DenseNet201
- DualPathNet68
- DualPathNet92
- DualPathNet98
- DualPathNet107
- DualPathNet113
- FBResNet152
- FCN
- InceptionResNetV2
- InceptionV3
- InceptionV4
- MASK-R-CNN
- NASNet-A-Large
- NASNet-A-Mobile
- PNASNet-5-Large
- PolyNet
- ResNeXt101_32x4d
- ResNeXt101_64x4d
- ResNet101
- ResNet152
- ResNet18
- ResNet34
- ResNet50
- SENet
- SE-ResNet50
- SE-ResNet101
- SE-ResNet152
- SE-ResNeXt50_32x4d
- SE-ResNeXt101_32x4d
- SqueezeNet
- VGG11
- VGG13
- VGG16
- VGG19
- VGG11_BN
- VGG13_BN
- VGG16_BN
- VGG19_BN
- Xception
- YOLOV3

Benefits

Works when the Internet is disconnected: The trained artificial intelligence plays a role even when the Internet is disconnected, but cloud-based artificial intelligence services are impossible. SoyNet has been developed to solve these problems by distributing artificial intelligence on edge devices.

Reduce high-cost GPU server: Even when running a simple AI model with TensorFlow or a caffe, server-class high-end equipment is required. SoyNet is an artificial intelligence execution accelerator platform that can execute various artificial intelligence models at low cost servers and IOT devices at high speed.

Easy at the application stage of AI: Usually, artificial Intelligence Developers need to know the obfuscated AI API to develop. However, if you use SoyNet, you can develop applications that utilize artificial intelligence in easier way. This is because

SoyNet provides the convenience of working with a simple API.

Price Policy

[Run Time License]

- Edge : ARM based Mali GPU and NVIDIA Jetson device based licensing
- Workstation : GPU based licensing
- Server: NVIDIA Tesla, Titan series of high-end GPU equipped , GPU based licensing

[Developer License]

- Developer PC/Notebook equipped with GPU (Development Only)

[Model License]

- AI Model configuration License

As to each AI model, the price is different. We claim one-time model fee for SoyNet to each legal entity. Please contact sales@soynet.io for more details. .

SOYNET SOYNET Co., Ltd. has started its own artificial intelligence execution accelerator platform since 2016 and established it in 2018.