Release Eagle (V0.5.0)

https://github.com/Adlik/Adlik/releases/tag/v0.5.0

Release Eagle

Feature List

Model Optimizer

Support quantization and distillation of YOLOv5s models, which achieves nearly 2.5 times inference performance improvement with OpenVINO runtime

Model Zoo

· A new repository that stores Adlik optimized and compiled models, including ResNet series models and YOLOv5 series models.

Compiler:

- Support compilation path from oneflow to onnx
- OpenVINO upgraded to version 2022.1.0

Inference Engine

- Support Torch runtime
- OpenVINO upgraded to version 2022.1.0

Benchmark Test

• Benchmark test for BERT model on Intel 8260 CPU, including throughput and other performance indicators.

Fixed issues

- Tensorrt int8 quantization function cannot be used
- OpenVINO runtime doesn't support multiple inputs model
- Can't activate model by grpc
- Segmentation fault when loading more than 16 models in manual mode
- Failed to compile PaddlePaddle model file to OpenVINO model