Nizhny Novgorod State University
Institute of Information Technologies, Mathematics and Mechanics
Department of Computer Software and Supercomputer Technologies

Educational course «Modern methods and technologies of deep learning in computer vision»

Lecture №3
Overview of the Intel Distribution of OpenVINO toolkit

Supported by Intel

Vasiliev E.P.

Nizhny Novgorod
2020

Content

1 Abstract

The goal of this lecture is to study the features of the Intel Distribution of OpenVINO Toolkit for deep learning inference [1, 2]. First, an overview of the main components of the OpenVINO toolkit is given. Further, we describe features of the components which provide deep learning inference, their Python API, and give general outlines of deep learning inference implemented by these components.

Intel Distribution of OpenVINOToolkit is a toolkit for solving computer vision and deep learning tasks, it is developing by Intel. The development goal is to ensure the ease of use of various computer vision and deep learning algorithms, as well as to increase the performance of applications based on deep learning and executed on a variety of Intel platforms (Intel CPUs, Intel Processor Graphics, Intel FPGAs).

The Intel Distribution of OpenVINO Toolkit includes the following components:

- ***Deep Learning for Computer Vision***
  - Intel Deep Learning Deployment Toolkit (DLDT)
    - Model Optimizer [3] is a tool for converting deep models from various formats to the intermediate representation (IR).
    - Inference Engine [6] is a component for high-performance inference of deep neural networks.
  - Open Model Zoo [4] is a public repository of trained models for solving various problems, samples and demos. It contains hundreds of trained deep models in various formats (public models and models trained by Intel engineers). Also, there are many examples and demo applications that demonstrate the use of the available models.
  - Deep Learning Workbench is a tool for calibrating models, measuring accuracy and benchmarking of models.
- ***Traditional Computer Vision***
  - OpenCV [5] is a well-known and widely used computer vision library.
- ***Tools & Packages***
  - Set of tools for improving performance of processing graphics and video (Intel Media SDK, OpenCL Drivers & Runtimes, etc).

We describe in detail the components of DLDT (Model Optimizer and Inference Engine), Open Model Zoo, and OpenCV. These components are responsible for deep learning inference. OpenVINO supports inference for models trained using various frameworks (TensorFlow, Caffe, MXNet, Kaldi, Keras, PyTorch) or converted to the ONNX format. We download models from the Open Model Zoo repository using Model Downloader, convert them to the intermediate representation using Model Optimizer, and infer using the Inference Engine component. Deep learning inference can be implemented using the Deep Neural Network (DNN) module of the OpenCV library. Also, OpenCV provides functions for preprocessing of input data (images and video) and processing network output. We provide a description of programming interface and step-by-step tutorials for implementing deep learning inference in Python using the Inference Engine component [6] and the DNN module of the OpenCV library [7].

2 Literature

2.1 References

1. OpenVINO documentation website [https://docs.openvinotoolkit.org].
2. Intel Distribution of OpenVINO Toolkit [https://software.intel.com/en-us/openvino-toolkit].
3. Model Optimizer [https://docs.openvinotoolkit.org/latest/_docs_MO_DG_Deep_Learning_Model_Optimizer_DevGuide.html].
4. Open Model Zoo [https://github.com/opencv/open_model_zoo].
5. OpenCV [https://opencv.org].

6. Inference Engine Python API Overview [https://docs.openvinotoolkit.org/latest/_inference_engine_ie_bridges_python_docs_api_overview.html].
7. OpenCV. Deep Neural Network module [https://docs.opencv.org/master/d6/d0f/group__dnn.html].