# Export

The trained model were exported using ONNX (Open Neural Network Exchange), a standard format for representing machine learning models that enables interoperability between various machine learning frameworks. The end goal is to run the models on TensorFlow Lite, a lightweight machine learning framework designed to run on mobile and embedded devices. A huge benefit of TensorFlow Lite is that it supports NNAPI (Neural Network Application Programming Interface) which allows for easy implementation on the DSP or GPU on the target device. One common problem that emerges when exporting a model from PyTorch to TensorFlow Lite is their mismatch in input sizes. While PyTorch uses the NCHW (Number of Batches, Channels, Height, Width) format, TensorFlow and TensorFlow Lite use the NHWC. To enable a seamless conversion without the need of additional transpose layers another framework called OpenVino is resorted to. The final export flow looks as follows:

PyTorch -> ONNX -> OpenVino -> TensorFlow Lite