Demo: a Guide to Real-Time Embedded Deep Learning Deployment for Elk Audio OS

Domenico Stefani
Dept. of Information Engineering
and Computer Science
University of Trento
Trento, Italy
domenico.stefani@unitn.it

I. Motivation and Introduction of Demo

Recent advancements in deep learning architectures for music [1] [2] and the availability of powerful embedded computing platforms for low-latency audio processing [3] [4] [5] [6] have opened up exciting possibilities for new digital and Smart Musical Instruments [7]. However, deploying neural networks to various embedded audio platforms remains a challenge. In particular, while Elk Audio OS [3] on the Raspberry Pi 4 proved to be a capable platform for deep learning and audio processing [8] [9] [10], the deployment process has not been documented yet.

In response, this demo proposal accompanies a systematic guide to deploying deep learning models for audio on embedded systems using Elk Audio OS. The proposed demo will cover the entire process, from creating a compatible code project to executing and diagnosing a VST plugin with deep learning inference on a Raspberry Pi. The demo will explore different approaches for real-time execution of deep learning inference on embedded devices and provide solutions for handling larger neural network models. To facilitate implementation and future updates, an online repository is provided[1], offering clean *code templates*, *functional examples*, and *precompiled library binaries* for TensorFlow Lite and ONNX runtime inference engines. The primary aim of this demo is to help bridge the gap between deep learning model development and real-world deployment on embedded systems [11], fostering the creation of self-contained digital musical instruments and audio devices equipped with real-time deep learning capabilities (i.e., embedded AI [12]). By promoting the deployment of neural networks to embedded devices, this demo seeks to contribute to the development of Smart Musical Instruments capable of providing musicians and audiences with unprecedented services. Attendants will gain insights into the process of deploying deep learning models on embedded computers with Elk Audio OS.

II. Details of the Demo

A. Scope and Objectives

The primary goal of this demo is to bridge the gap between deep learning model development and real-world deployment on embedded systems, with a focus on enabling real-time deep learning capabilities in musical instruments and audio devices. Through practical examples, attendants will gain insights into the process of deploying deep learning models on embedded computers using Elk Audio OS. The deployment process is presented in Fig.1

B. Demo Highlights

1) **Creating a Compatible Code Project:** attendants will be guided through the process of setting up a compatible code project for deploying deep learning models on embedded devices. This step involves selecting appropriate libraries, dependencies, and configurations for Elk Audio OS.
2) **Cross-compilation for Elk Audio OS:** The attendants will be exposed to the process of cross-compilation of library dependencies and plugins for Elk Audio OS and the Raspberry Pi 4.
3) **Real-time Execution of Deep Learning Inference:** Attendants will explore different approaches for achieving real-time execution of deep learning inference on embedded devices and different modes of execution. Neural network optimization for low-latency real-time execution will be discussed.
4) **Diagnostic process:** The demo will discuss how to diagnose a VST plugin with real-time inference on a Raspberry Pi 4 running Elk Audio OS, showcasing real-world application scenarios for deep learning in audio processing.

C. Online Repository

An online repository will be provided to facilitate implementation and support future updates. It contains clean code templates, functional examples, and precompiled library binaries for TensorFlow Lite and ONNX runtime inference engines. Attendants can access these resources for further exploration and experimentation.

D. Benefits

By attending this demo, attendants will:

- Gain practical knowledge and skills in deploying deep learning models on embedded systems with Elk Audio OS.

[1] https://github.com/CIMIL/elk-audio-AI-tutorial

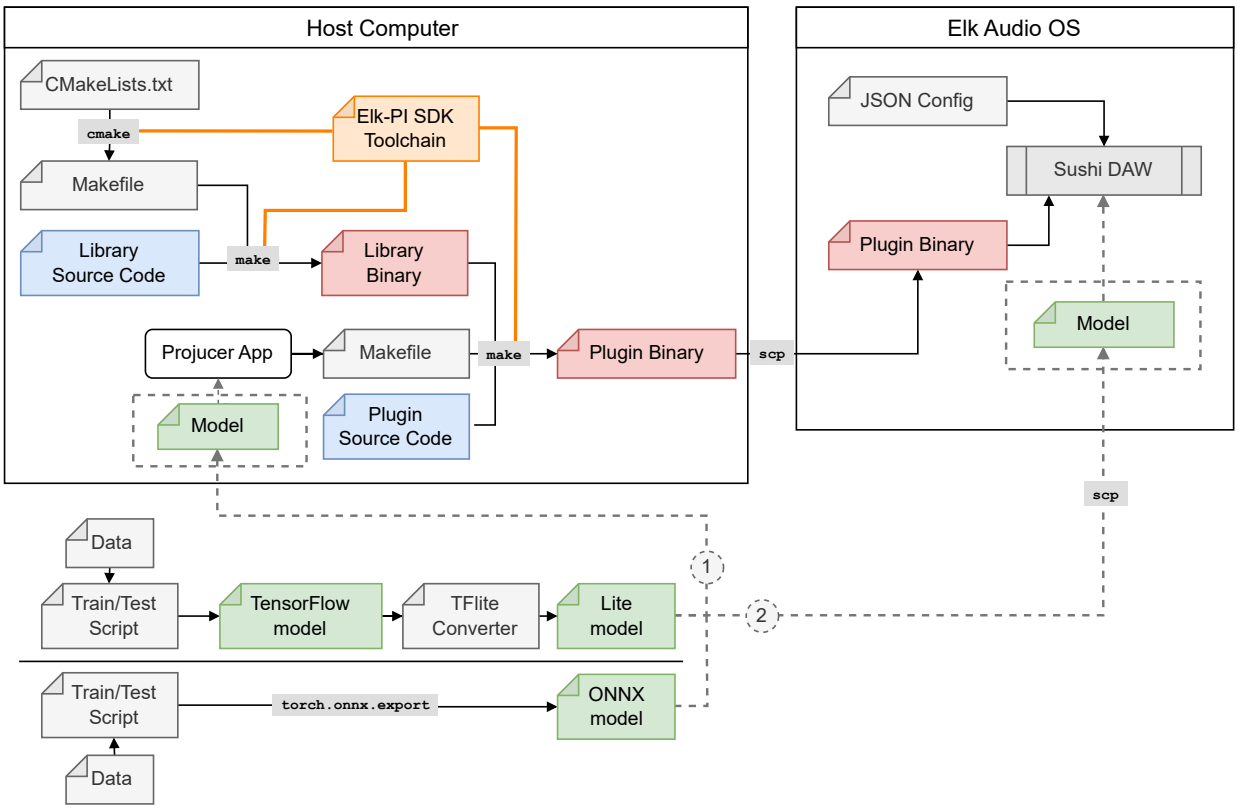

Fig. 1. Diagram depicting the deployment process for a deep learning model to an embedded device running Elk Audio OS. Plugin and dependency compilation happens on a Host computer (top left), where the Elk-PI toolchain allows to cross-compile the source code. Library binaries are linked during plugin compilation, which produces the binary file for a VST plugin that can execute on the target device (right). The compiled plugin has to be moved to the embedded computer running Elk Audio OS, where Elk's DAW Sushi can be configured to load it into a new track and process audio in real-time. The lower part represents the training, testing, and model export phases required for either TensorFlow or frameworks that export to the ONNX format, such as Pytorch. Dashed arrows labeled as 1 and 2 represent two distinct options for model integration: in option 1 the deep learning model gets integrated as JUCE BinaryData into the plugin's binary, while for option 2 the model can be simply copied to the device. For the latter, the plugin code must load the model from a path relative to the target's folder structure.

- Understand the challenges and solutions for achieving real-time execution of neural networks in musical instruments and audio devices.
- Explore real-world use cases and applications of deep learning in audio processing.

References

[1] J. Engel, L. H. Hantrakul, C. Gu, and A. Roberts, "DDSP: Differentiable digital signal processing," in *International Conference on Learning Representations*, 2020. [Online]. Available: https://openreview.net/forum?id=B1x1ma4tDr

[2] A. Caillon and P. Esling, "RAVE: A variational autoencoder for fast and high-quality neural audio synthesis," *CoRR*, vol. abs/2111.05011, 2021. [Online]. Available: https://arxiv.org/abs/2111.05011

[3] L. Turchet and C. Fischione, "Elk Audio OS: an open source operating system for the Internet of Musical Things," *ACM Transactions on the Internet of Things*, vol. 2, no. 2, pp. 1–18, 2021.

[4] A. McPherson and V. Zappi, "An environment for submillisecond-latency audio and sensor processing on beaglebone black," in *Audio Engineering Society Convention 138*. Audio Engineering Society, 2015.

[5] E. Meneses, J. Wang, S. Freire, and M. M. Wanderley, "A comparison of open-source linux frameworks for an augmented musical instrument implementation," in *Proceedings of the Conference on New Interfaces for Musical Expression*, 2019, pp. 222–227.

[6] L. Vignati, S. Zambon, and L. Turchet, "A comparison of real-time Linux-based architectures for embedded musical applications," *Journal of the Audio Engineering Society*, vol. 70, no. 1/2, pp. 83–93, 2022.

[7] L. Turchet, "Smart Musical Instruments: vision, design principles, and future directions," *IEEE Access*, vol. 7, pp. 8944–8963, 2019.

[8] D. Stefani and L. Turchet, "Demo of the timbreid-vst plugin for embedded real-time classification of individual musical instruments timbres," in *Proc. 27th Conf. of Open Innovations Association (FRUCT)*, vol. 2, 2020, pp. 412–413.

[9] ——, "On the Challenges of Embedded Real-Time Music Information Retrieval," in *Proceedings of the 25-th Int. Conf. on Digital Audio Effects (DAFx20in22)*, vol. 3, Sept. 2022, pp. 177–184.

[10] D. Stefani, S. Peroni, and L. Turchet, "A comparison of deep learning inference engines for embedded real-time audio classification," in *Proceedings of the Digital Audio Effects Conference*, 2022.

[11] T. Pelinski, R. Diaz, A. L. B. Temprano, and A. McPherson, "Pipeline for recording datasets and running neural networks on the bela embedded hardware platform," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, 2023.

[12] T. Pelinski, V. Shepardson, S. Symons, F. S. Caspe, A. L. Benito Temprano, J. Armitage, C. Kiefer, R. Fiebrink, T. Magnusson, and A. McPherson, "Embedded AI for NIME: Challenges and Opportunities," *International Conference on New Interfaces for Musical Expression*, jun 22 2022. [Online]. Available: https://nime.pubpub.org/pub/rwr2c3zs