MICHIGAN ENGINEERING
UNIVERSITY OF MICHIGAN

Machine Learning and Instrument Recognition

Team Bassoon: Aiden Wenzel, Tony Valencia, Emre Yavuz, Adarsh Bharathwaj

3/15/2024

Figure 1: Machine Learning
Source: Adapted from [1]

Figure 2: Bassoon Player
Source: Adapted from [2]

MICHIGAN ENGINEERING
UNIVERSITY OF MICHIGAN

Overview

Table of Contents

Context

- Traditional approach to instrument recognition
  - Use original audio signal and manually filter out noise
  - Requires the use of data analysis and Fast-Fourier Transforms (FFTs)

- Fast-Fourier Transform (FFT)
  - Uses an input signal and calculates underlying coefficients and frequencies to be used for calculation

- Machine Learning (ML)
  - Allows for prediction given an input signal [1]

Problem

- Researchers cannot efficiently analyze noisy signals
  - Filtering signals loses data and decreases accuracy of predictions
  - Time-consuming process

- Need to facilitate research process of audio recognition
  - Eliminate the need for signal filtration
  - Understand the viability of other methods rather than using an FFT

User Group & Needs

- User Group
  - Signal Processing Researchers
  - Audio / software engineers who specialize in developing music recommendation and recognition algorithms

- User Needs
  - Human-comparable identification accuracy of instrument/instruments within given audio sample
  - Resilience to noise
  - Computationally efficient model (Small size, low memory use, quick performance)

Functional Requirements

- Functional and usable
  - Takes in audio inputs and returns ideally return list of correct outputs
  - Can accurately identify at least 4 unique instruments
- Test accuracy ≥ 90%
  - Provide incentive to use ML model
  - Show that ML is reliable and accurate
- Simple to use
  - Can abstract away signal processing theory
  - Has a GUI as an interface for input and output

Proposal

- Create a FNN with input audio signal, output instrument type
  - Use a Machine Learning (ML) framework from either Python or Julia
  - Compile audio samples from online databases to train the model
  - Candidate model types include CNN, RNN-LSTM, and MLP [6] [7]

- Comparison study between the accuracy of using FFT approach versus Machine Learning model
  - Determine the viability of Machine Learning in various signal processing applications

Choosing a Language

- Python
  - Well documented + expansive libraries such as TensorFlow [3]
  - Data analysis (audio) may be more difficult than using Julia

- Julia
  - Flux.jl integrated package in Julia that contains most ML models [4]
  - Julia combines speed and accuracy of data analysis with ML
  - Better suited for heavy batch jobs

Machine Learning

- Using a FNN [5]
  - Train the model using single-instrument audio-signals
  - Single- and multi-instrument signals dataset [8]
  - Create a graphical user interface (GUI) for the user to input their own signal for prediction

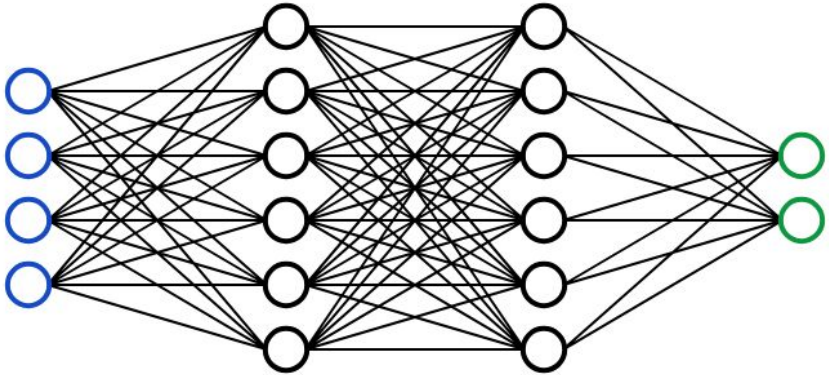

Figure 3: Neural Network
Source: Adapted from [5]

Flowchart

- Multi-step process
  - Can be done in parallel
  - Utilize team member strengths

Conclusion

- Goal:
  - Streamline the process of identifying instruments in noisy audio samples
  - Compare the effectiveness of machine learning to traditional signal processing methods when identifying instruments
- Methods:
  - Use an open source machine learning library to build a neural network
  - Train the network using audio data from open-source datasets
  - Compare the rate of success in identifying instruments using machine learning vs traditional signal processing

Figure 4: Flux.jl Logo
Source: Adapted from [4]

Figure 5: TensorFlow Logo
Source: Adapted from [3]

References (IEEE)

[1] V. Kanade, “What is machine learning? understanding types & applications,” Spiceworks, https://www.spiceworks.com/tech/artificial-intelligence/articles/what-is-ml/ (accessed Mar. 14, 2024).

[2] J. Syme, “Sci Fi,” Pinterest, https://www.pinterest.com/pin/star-wars-bassoon-3-3-3--224546731392673978/ (accessed Mar. 14, 2024).

[3] “Module: tf,” TensorFlow, https://www.tensorflow.org/api_docs/python/tf (accessed Mar. 14, 2024).

[4] “Flux: The julia machine learning library,” Fluxml, https://fluxml.ai/Flux.jl/stable/ (accessed Mar. 13, 2024).

[5] V. Zhou, “Neural Networks from scratch,” Neural Networks From Scratch, https://victorzhou.com/series/neural-networks-from-scratch/ (accessed Mar. 11, 2024).

[6] S. M. Elghamrawy and S. Edin Ibrahim, "Audio Signal Processing and Musical Instrument Detection using Deep Learning Techniques," 2021 9th International Japan-Africa Conference on Electronics, Communications, and Computations (JAC-ECC), Alexandria, Egypt, 2021, pp. 146-149, doi: 10.1109/JAC-ECC54461.2021.9691427 (accessed Mar. 12, 2024)

[7] Blaszke, M., & Kostek, B. (2022). Musical Instrument Identification Using Deep Learning Approach. Sensors (Basel, Switzerland), 22(8), 3033. https://doi.org/10.3390/s22083033 (accessed Mar. 12, 2024)

[7] E. Humphrey, S. Durand and B. McFee, "OpenMIC2018: An Open Data-set for Multiple Instrument Recognition", ISMIR, pp. 438-444, September 2018. (accessed Mar. 12, 2024)

Team Contact Information

Aiden Wenzel: aidenwen@umich.edu

Tony Valencia: avalen@umich.edu

Emre Yavuz: eyavuz@umiche.du

Adarsh Bharathwaj: adarshb@umich.edu

MICHIGAN ENGINEERING
UNIVERSITY OF MICHIGAN

Questions & Answers