

GESTURE-BASED MIDI SYNTHESIS

MACHINE LEARNING PROJECT

ASEEL IBRAHIM
SARIGA NAIR



INTRODUCTION

Our project, titled “**Gesture Based MIDI Synthesis of D Chords**,” is a simple web-based application that lets users play piano sounds using their computer keyboard. It does not require a real piano or any musical instrument. Instead, it uses the **Web MIDI API**, which allows web browsers to produce musical sounds.

In this project, we have focused only on the D chords — specifically:

- **D Major**
- **D Minor**
- **D7**



WHAT IS MIDI?

MIDI stands for Musical Instrument Digital Interface, a widely used communication protocol that enables electronic musical instruments, computers, and other related devices to interact. MIDI doesn't transmit audio but rather digital signals that indicate which notes to play, their velocity, duration, and other control changes. These messages are interpreted by synthesizers or software instruments to generate sounds. This project uses MIDI because it offers a flexible and standardized way to trigger musical notes and chords, making it compatible with numerous digital audio workstations and virtual instruments.

WHY GESTURE BASED MUSIC?

Using hand gestures as an interface for playing music provides a natural and intuitive mode of interaction. Unlike pressing keys or buttons, gestures allow performers to control music dynamically and expressively without physical contact. This method is particularly valuable in live performances where musicians can blend movement and sound seamlessly. Gesture control is also beneficial for people with physical disabilities or those seeking alternative musical expression forms. Furthermore, it introduces a novel creative tool for education and entertainment, making music-making more engaging and accessible.

OBJECTIVES

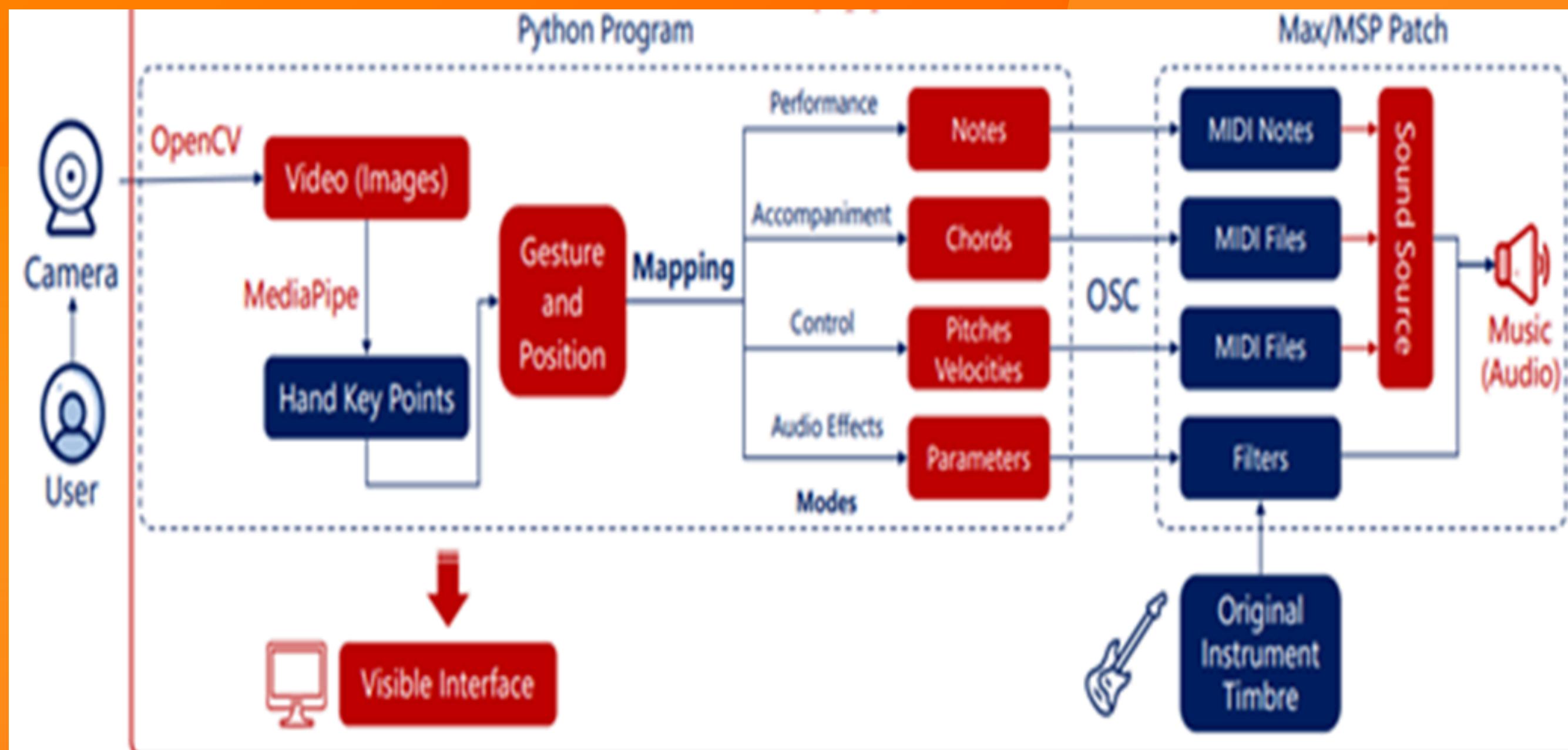
The main goal of this project is to design a system that can detect and recognize specific hand gestures in real-time using a standard webcam. Once detected, these gestures are mapped to chords within the D major scale, such as D Major, E minor, and F# minor. The system then sends appropriate MIDI messages to a synthesizer or digital audio workstation to play the corresponding chords. Achieving low-latency and accurate gesture recognition ensures a smooth musical experience. This project combines computer vision techniques, real-time processing, and MIDI synthesis to create an innovative musical instrument controlled solely by hand movements.



HARDWARE & SOFTWARE REQUIREMENTS

- 1. Standard WEB-CAM- for capturing video input**
- 2. Python 3.x -programming language**
- 3. MediaPipe&CVzone- for hand detection**
- 4. OpenCV -for video/camera processing**
- 5. Pygame-gesture to midi conversion**
- 6. Numpy-fast math for gesture data**

SYSTEM ARCHITECTURE

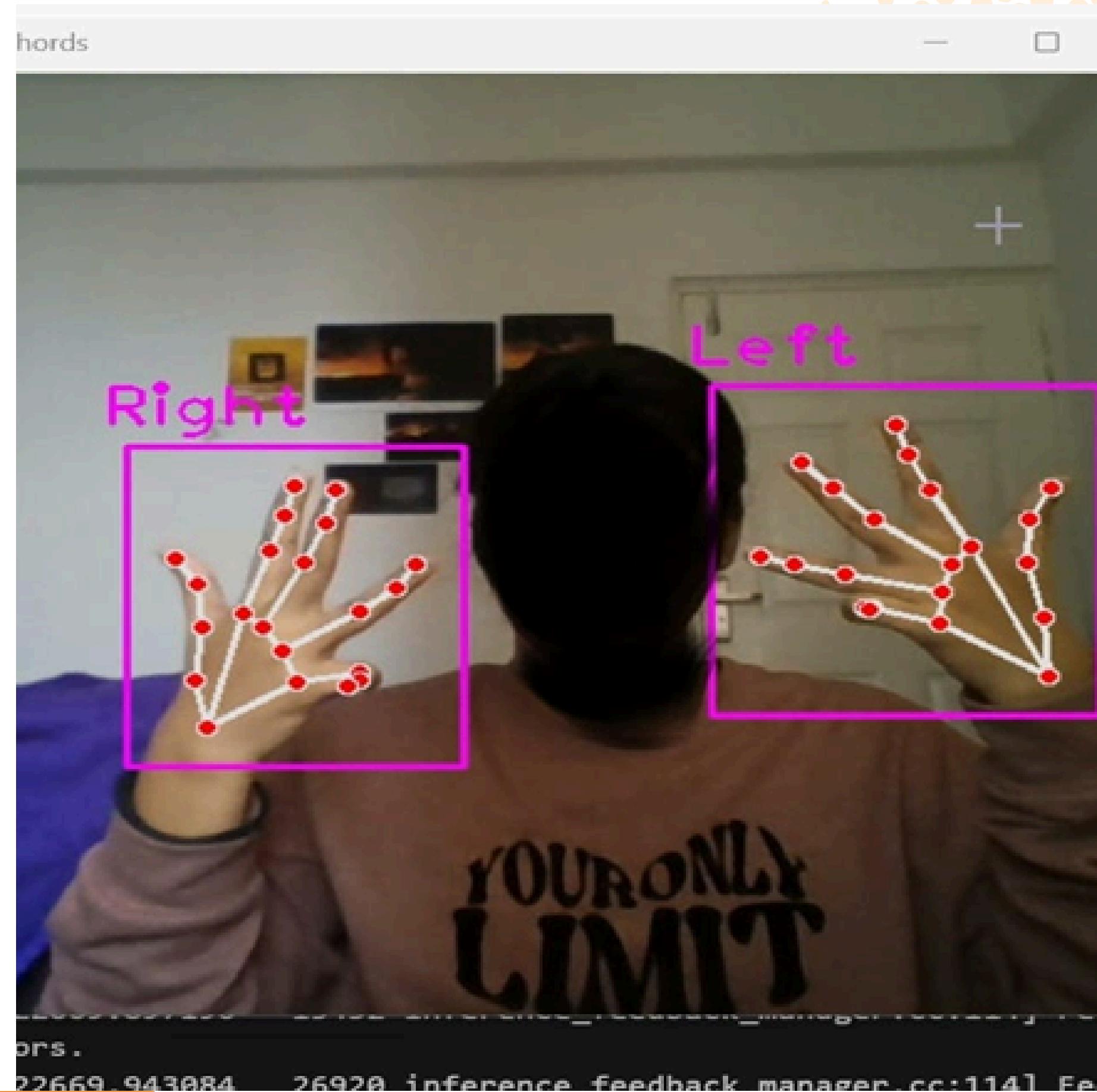


GESTURE RECOGNITION LOGIC

- Checks finger states: up, down, extended, bent.**
- Compares distances between landmarks (e.g., tip and base).**
- Predefined patterns for each chord gesture.**
- Simple if-else conditions used for classification.**

CHORD MAPPING

- Each gesture corresponds to a D scale chord:
- D major, E minor, F# minor, G major, A major, B minor, C# dim.
- Chord = combination of 3 notes sent via MIDI.
- Gesture switch changes the chord in real-time.



SOFTWARE WORKFLOW

- 1.Capture frame via webcam.**
- 2.Detect hand using MediaPipe.**
- 3.Classify gesture into chord.**
- 4.Generate & send MIDI signal.**
- 5.Update UI accordingly.**





BENEFITS OF MIDI

- No physical contact needed.
- Affordable hardware (just a webcam).
- Great for disabled users.
- Interactive and visually engaging.
- Easy to modify and extend.



APPLICATIONS

- Live music performance.
- Music therapy and education.
- Installation art and interactive exhibits.
- Hands-free MIDI control in studios.

Air

Johann Sebastian Bach
from Orchestral Suite No.3,BWV1068
Arranged by Shigeo Ida / Jean Redus

The musical score for 'Air' by Johann Sebastian Bach, arranged for piano, is displayed. The score is in common time with a key signature of one sharp (F#). The tempo is marked as quarter note = 31. The music is divided into ten measures. Measure 1 starts with a piano dynamic and a bass line. Measures 2-3 show a melodic line in the treble staff with grace notes. Measures 4-5 continue the melodic line. Measures 6-7 show a return to the bass line. Measures 8-9 show a continuation of the melodic line. Measure 10 concludes the piece. Various dynamics are used throughout, including piano (p) and mezzo-forte (mp).



LIMITATIONS

- Needs good lighting for tracking.
- Background noise may confuse hand detection.
- Only one hand used in current version.
- Accuracy depends on hand distance from webcam.



FUTURE IMPROVEMENTS

- Add support for both hands.
- Use ML for more robust gesture classification.
- Enable custom chord training by user.
- Build a mobile app version.



THANK YOU !

