1. INTRODUCTION

During the quarantine caused by the COVID-19 pandemic, people are trying to find a new hobby to fill in their free time during the pandemic, so a lot of these people got into playing the guitar[1],[2],[3],[4],[5]. Many of these new players rely on videos from the internet or guitar tablatures to hone their playing skills or to learn new songs. Unfortunately, novice players often struggle with the complexities of finger positioning, struggling to translate chord diagrams into the accurate placement of the fingers. The inherent challenge in acquiring these fundamental techniques can be a significant challenge to overcome. A lot of systems has been proposed to overcome these challenges. One of the many ways is to use computer vision to detect the chord performed by the player. This itself comes with its own problems. Previous works on chord recognition using computer vision still face problems with recognizing similarly shaped chords and even distinctly shaped ones as the number of chords to be recognized grow up[6],[7]. With the number of even the most basic chords there are, using computer vision as suggested by previous works becomes less effective. We aim to overcome these problems by using a new approach that leverages a keypoint detection model.

There has been multiple works that can be implemented to perform guitar chord recognition. One work suggests by using ResNet18 to perform recognition[6]. This comes with its drawbacks being that it often misclassify similarly shaped chords even with hyperparameters tunings and with unfreezing it's last few layers which did help to improve it's accuracy. This work is also trained only on pictures of 5 different people, which can affect the robustness of the model. Furthermore, the author had noted by saying that the model's robustness could be improved by training the model on a more diverse dataset composed of examples with more diversity in hand color, size, image angle and background. Another work utilized a pre-trained deep learning neural network based on GoogLeNet[7]. It also faces the same problem of lacking image data that causes overfitting. The work also shows that as the number of category increases, the accuracy goes down linearly. Both works also suggests that future works may overcome this problem by utilizing a more sophisticated model that can better recognize the played chords while doing it in real time to create a more realistic use of the model. In summary these approaches suffer from several major limitations:

• **Misclassification of Similar Chords:** The application of ResNet18 in guitar chord recognition, as suggested by one work[6] and the use of a pre-trained deep learning neural network based on GoogLeNet as suggested by another[7], tends to misclassify chords with similar shapes.

• **Limited Dataset Diversity:** A notable drawback in the ResNet18-based work is the reliance on a training dataset composed of images from only five different individuals[6]. This lack of diversity in the dataset, specifically in hand color, size, image angle, and background, poses a challenge to the robustness of the model.

• **Challenges with Scalability**: Both works highlight a common challenge related to scalability. As the number of chord categories increases, the accuracy of the models tends to decrease [6],[7]. This scalability issue underscores the need for future works to explore more sophisticated models capable of real-time chord recognition, ensuring a more realistic application of the models in practical scenarios.

To solve the aforementioned problem, we introduce a new approach to assist beginners in their guitar learning journey. We present a real-time guitar chord recognition system that uses computer vision and machine learning techniques. By leveraging a webcam-based feed, this system aims to accurately detect and identify the chords being played, providing support and guidance to new guitarists. It is important to notice that even the most sophisticated model cannot be the be-all and end-all solution in this problem. This is because of the hardware limitiations of the devices that this model will most likely to be used in. According to the recent hardware surveys[8], while most of the hardwares are capable of running the detection algorithm, some of them may have difficulties to achieve the desired performance, especially the older ones. Previous works also shows that while the number of categories of chords increases, the accuracy tends to go down. In order to achieve a fair performance across multiple hardwares, we keep our model architecture simple and not overly complicated[9].

To summarize, the contributions from this paper are as follows:

• **Two-Step Methodology Integration:** We introduce a two-step methodology that synergizes keypoint detection models and machine learning algorithms. This integration leverages on the strengths of each phase, with keypoint detection identifying finger positions and a machine learning model performing chord classification based on either raw keypoint outputs or inter-keypoint distances.

• **Real-time Chord Recognition:** Our proposed method operates in real-time, providing immediate feedback to guitar learners. By leveraging Google's Mediapipe framework for keypoint detection, the system accurately identifies and tracks critical finger positions on the guitar fretboard, facilitating chord recognition and classification.

• **Experiments on an Inter-Keypoint Distance Classifier:** We provide experiments and analysis on the implementation of neural network-based inter-keypoint classifiers in some scenarios. The outcomes of these experiments contribute valuable insights into the performance and effectiveness of the classifiers, providing a solid foundation for understanding their practical implications and potential applications.

The remainder of this paper is organized as follows. An exploration on existing works related to this study in Section 2, details of the proposed method in Section 3, Evaluation of our proposed method in section 4 and finally, Section 5 concludes the paper with possible future works.