Transcription of Guitar Chords from Acoustic Audio

Darrell A. Talavera, Erika Svetlana C. Nase, Leonard D. Pancho, and Adomar L. Ilao
Malayan Colleges Laguna, Cabuyao City, Laguna, Philippines
Email: {darrell.talavera, erika.nase, leonard.pancho}@live.mcl.edu.ph, alilao@mcl.edu.ph

***Abstract*—Guitar has been played since the 15th century until now. The traditional method used in transcribing guitar chords from a song is through determining the chords while listening to it. Through the use of modern technology in the field of Computer Science, transcribing guitar chords from acoustic audio can be attainable through classification method. Musical features were extracted by J48 decision from an audio file. The mean percentages of success using the prototype and Weka are 33.81% and 47.79% respectively. Further data analysis through t-test shows Weka selection of musical features is more relevant than the prototype.**

***Index Terms*—guitar, chords recognition, J48 decision trees, classification method, data mining, musical feature extraction**

I. Introduction

Music listening is one of most mystifying human behavior commonly associated to recreation [1]. More than recreation, music provides mood enhancement or distraction in everyday's life [2]. Others choose to listen to music based on context dependent [3]. Furthermore, playing music goes beyond recreation. It is known to raise self-esteem and brings enjoyment, challenge and empowerment [4].

The most popular musical instrument is guitar. Guitar is a versatile instrument that offers huge range of musical styles – rock music, country music and flamenco music all use the same instrument to create wildly different sounds [5]. It brings several benefits namely social benefits, personal benefits, professional benefits, mental benefits and physical health benefits [6]. Beyond guitar flexibility and mobility, it became one of the mostly played or learned instrument and sold around 1.4 Million US Dollar in 2011 [7]. Guitar became the most accepted instrument today possibly because it is practical and can accompany by singing voice easily.

Playing guitar requires different techniques to master namely frets shown in Fig. 1, finger positioning and strumming as shown in Fig. 2. It works knowing the chords. Chord is three or more musical notes played at the same time. The most common usage of chords occurs in popular music, usually played on guitar [8] as shown in Fig. 3. Fig. 3 shows varieties of chords ranging from major, minor, seventh, sixth, suspended and diminished chords.

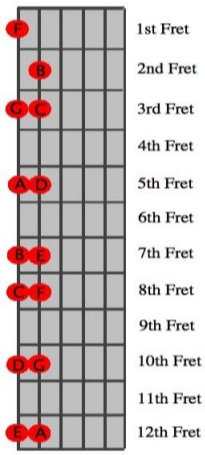


Figure 1. Guitar fret chords

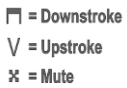


Figure 2. Strumming pattern

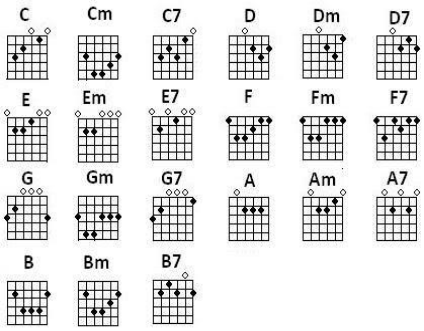


Figure 3. Basic guitar chords

Normally a song is differentiated by the type of genre. Whatever genre of music the song belongs, it is composed of more than one chord and varieties of strumming techniques.

However, some guitarists are fond of recognizing the chords of a specific song by themselves or wait a couple of times to be able to get the chords from their resources. Guitarist's experience can improve through availability of new technology in the form of media program that can be utilized to learn and construct their own music transcription [9].

Manuscript received March 11, 2020; revised July 3, 2020.

© 2020 J. Adv. Inf. Technol.
doi: 10.12720/jait.11.3.149-154

A single chord recognition is simple for beginner by watching finger position from a video recording. Availability of guitar video recording is based on popularity of the song. The process of converting video or audio recording to text is known as transcription. A technique that will serve useful to guitarist that has difficulties extracting of guitar chords to different musical style or genre. It deals with notation of musical piece to produce music score involving identification of particular instruments, pitches, rhythmic patterns [10] and timing as shown in Fig. 4.

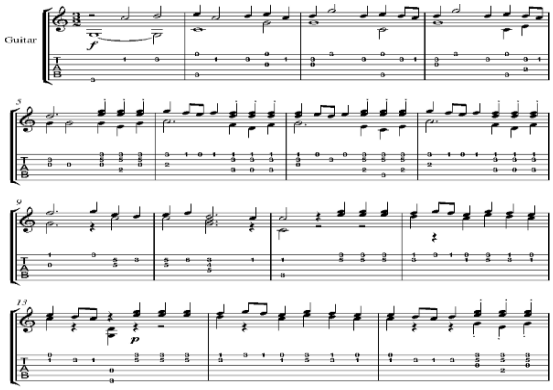


Figure 4. Guitar musical score

Basically transcription is done manually by skilled musician which requires time and sometimes inaccurate [11]. A new way of acquiring guitar chords for the guitarist is the concern of this study focus on chord recognition [12].

Fortunately, there are methods that can help transcribe guitar chords from an audio [13]. One technique is classification method. It was used to classify music genre, similarity analysis, music recommendation, performer identification, composer identification, and instrument identification [14]. It can be applied to do such activity through set of features or parameters to characterize each object, where these features should be relevant to the task at hand.

The study applied supervised classification, where a guitarist determined what class an object may be categorized and provided training set used to learn how to classify acquired objects [15] from an audio file.

The objective of the study is to provide an innovative resource in acquiring guitar chords of an acoustic guitar audio guided by specific objectives:

1) To extract musical features produced by the sound of guitar chords
2) To implement an algorithm based from a classification method that will match musical features of guitar chords
3) To test and evaluate the chord classifier.

Training set is strictly acoustic composed of single 6 strings and limited in four categories which are Major, Minor, Dominant 7th, and Minor 7th with 12-notes each which are A, A#, B, C, C#, D, D#, E, F, F#, G, G#.

Your goal is to simulate the usual appearance of papers in the. We are requesting that you follow these guidelines as closely as possible.

II. Methodology

The study followed the framework as reflected on Fig. 5. An acoustic guitar is used to generate guitar audio recorded from a microphone. The digital signal undergone through running-pass and band-pass filters to gather better results.

It requires several stages such as recording of basic guitar chords, extraction of musical features that will be stored in a CSV (Comma Separated Values) file; and, transcribing the chords from the audio file using decision tree created by J48 algorithm.

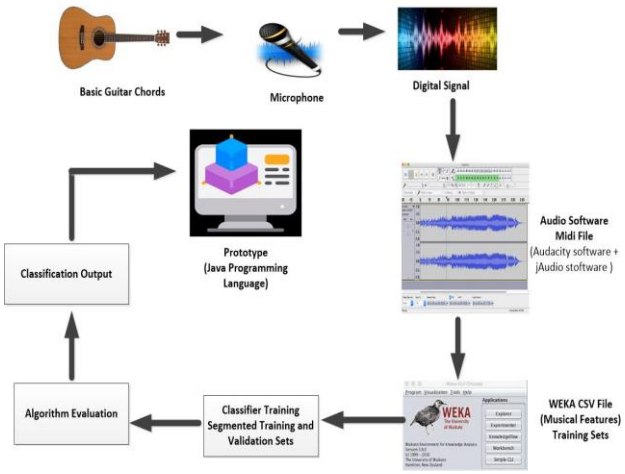


Figure 5. Implement research set-up

Guitar chords were recorded several times using Audacity software continuously. Afterwards, audio softwares aided to crop chord sounds available from a recorded audio file and extract musical features of sounds from different chords by Audacity software and jAudio software [16] which are vital to populate the needed training set. Audicity software was configured based on Table I and used to records the guitar chords in MIDI file format.

TABLE I. Audacity Reverb Configuration

Settings	Value
Room Size (%)	95
Pre-delay (ms)	20
Reverberance (%)	55
Damping (%)	50
Tone Low (%)	50
Tone High (%)	55
Wet Gain (db)	-1
Dry Gain (db)	-1
Stereo Width (%)	75

While jAudio software is capable to extracts 33 core features namely: Power Spectrum, Magnitude Spectrum, Spectral Variability, Spectral Centroid, Partial-Based Spectral Centroid, Partial-Based Spectral Smoothness, Compactness, Spectral Roll-off Point, Spectral Flux, Partial-Based Spectral Flux, Method of Moments, Area of Moments and MFCC, Zero Crossings, RMS, Relative Difference Function, Fraction of Low-Energy Frames,

© 2020 J. Adv. Inf. Technol.

LPC, Beat Histogram,, Strongest Beat, Beat Sum, Strength of Strongest Beat, Strongest Frequency via Zero Crossings, Strongest Frequency via Spectral Centroid, Strongest Frequency via FFT Maximum [17], Average Spectral Flux, Beat Histogram Bin Labels, Compactness, FFT Bin Frequency Labels, Root Mean Square Variability, Spectral Centroid Variability, Strongest Frequency Variability, Zero Crossings Derivative, Zero Crossing Variability [18] and import the said features to an CSV file format directly compatible to WEKA.

WEKA is an open-source machine learning tool written in JAVA equipped with learning algorithms sufficient to run dataset. It was applied to different field of data mining study which requires data preprocessing, clustering, classification, regression, visualization and features selection [19].

The critical step is to identify the best features out of the 32 extracted by jAudio. The technique to draw the optimal subset from listed features is Feature Selection. It can be able to reduce the irrelevant and redundant features and choose the best features that represent the dataset. The study will implement ConsistentctSubsetEval with Best First search and Ranker techniques to determine best features before J48 algorithm [20].

J48 algorithm implements tree pruning. Tree pruning simplifies options and correct possible over fitting of data. It classifies data by excessive rules and generalized rules through decision tree. The internal nodes of the decision tree denote the different attributes and the range of values an attributes can have in the observed training set.

The study is designed in quantitative approach through experimental design. It applies One Group Two-Way Posttest Design approach. WEKA features will be served as the classifier data sets while prototype results will be considered as the Post-Test datasets. Prototype was written in Java programming language utilizing computer specification depicted in Table II.

TABLE II. COMPUTER HARDWARE SPECIFICATION

OS: Windows 10
Processor: Dual Core from Intel or AMD at 2.6GHz
Memory: 4GB
Graphics: nVidia GeForce 8600/9600GT, ATI/AMD Radeon HD 2600/3600
Hard drive: 500MB available space

The initial phase involves recording of chords from the acoustic guitar through Audacity Software. The recording requires 10 times stroke for each chords (A, A#, B, C, C#, D, D#, E, F, F#, G, G#.). Each note is fundamentally within 82.41 to 1244.51 Hertz and prominent spectral components are beneath 2 kHz as illustrated in Fig. 6.

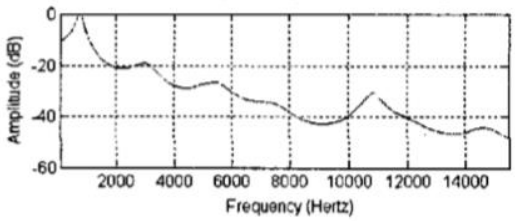


Figure 6. 6-String, 22-fret guitar spectral envelope [21]

Afterwards, jAudio software was used to extract musical features of different chords and became the dataset of experiment. Test set was analyzed in WEKA to identify the best musical features and generated a decision tree by J48 algorithm.

An input audio file is slice to single strum and transcribe to a guitar chord with the aid of the test set. Each transcribed chord will be captured to a text file.

Prototype performance evaluation comprised of statistical tool such as success rate and t-test to identify whether the difference from pre-selected features between the prototype and WEKA data mining tool were significant.

III. RESULTS AND CONCLUSION

In order to have the desired results of the experiment, it underwent to this methodology process: recording of chords, cropping recorded chords to one second time length, determining the features to be used for the comparison of chords, extraction of selected features of each chord, and comparing chords with the use of decision tree classifier using J48 algorithm.

A. Data Collected

The first process of the data collection is the recording of acoustic guitar chords, dreadnought, and a recorder in a closed room. A total of 48 chords from Major, Minor, Dominant 7th, and Minor 7th with 12-notes each which are A, A#, B, C, C#, D, D#, E, F, F#, G, G# were collected leading to 1857 of instances each chords. Instances were the result of the one second cropped recorded chord through the use of Audacity. The one second time interval for the recorded chords was based on the ADSR (Attack Decay Sustain Release) of sounds [22] as illustrated in Fig. 7 which becomes the training set.

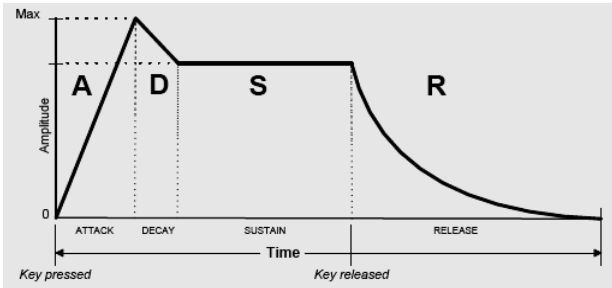


Figure. 7. Schematic of ADSR

Training set contains 10 instances of each chord from the pool of 1857 chords with a total of 480 chords. Each chord will be processed by jAudio that was setup to 88 features and was trimmed down to selected 2 sets of features manually which 29 listed in Table III and 32 number of features in Table IV respectively. The feature selection will determine which features are vital to the study.

TABLE III. LIST OF 29 FEATURES SELECTED MANUALLY

Features (in overall standard deviation)	Number
Spectral Centroid	1
Spectral Flux	1

Compactness	1
Spectral Variability	1
Root Mean Square	1
Mel-frequency Cepstral Coefficients (MFCC)	13
Linear Predictive Coeffecients (LPC)	9
Partial Based Spectral Centroid	1
Partial Based Spectral Flux	1

TABLE IV. LIST OF 32 FEATURES SELECTED MANUALLY

Features (in overall standard deviation)	Number
Spectral Centroid	1
Spectral Flux	1
Compactness	1
Spectral Variability	1
Root Mean Square	1
Zero Crossings	1
Strongest Frequency via Zero Crossings	1
Strongest Frequency via FFT Maximum	1
Mel-frequency Cepstral Coefficients (MFCC)	13
Linear Predictive Coeffecients (LPC)	9
Partial Based Spectral Centroid	1
Partial Based Spectral Flux	1

Initially, the study obtained five training sets resulting from five sets of features each having 10 instances per chords. The datasets were in the CSV (Comma Separated Values) file format for WEKA compatibility.

WEKA through Ranker and Best First attribute evaluator with ConsistentctSubsetEval resulted 19, 31, and 50 as the best features that can be used in transcribing the captured chord as depicted in Table V, Table VI and Table VII respectively.

TABLE V. LIST OF 19 FEATURES FROM WEKA

Features	Number
Root Mean Square Overall Standard Deviation	1
MFCC Overall Standard Deviation	2
LPC Overall Standard Deviation	1
Derivative of Spectral Centroid Overall Average	1
MFCC Overall Average	12
LPC Overall Average	1
Strongest Frequency Via FFT Maximum Overall Average	1

TABLE VI. LIST OF 31 FEATURES FROM WEKA

Features	Number
Root Mean Square Overall Standard Deviation	1
Mel-frequency Cepstral Coefficients (MFCC) Overall Standard Deviation	13
Linear Predictive Coefficients 2 (LPC) Overall Standard Deviation	1
Derivative of Spectral Centroid Overall Average	1
Strongest Frequency Via FFT Maximum Overall Average	1
Mel-frequency Cepstral Coefficients (MFCC) Overall Average	13
Linear Predictive Coefficients 2 (LPC) Overall Average	1

TABLE VII. LIST OF 50 FEATURES FROM WEKA

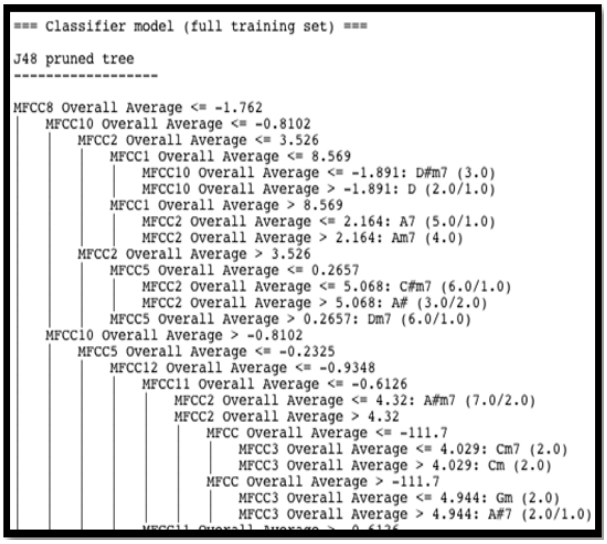
Features	Number
Spectral Centroid Overall Standard Deviation	1
Derivative of Spectral Centroid Overall Standard Deviation	1
Derivative of Spectral Centroid Overall Average	1
Spectral Variability	1
Spectral Variability Overall Standard Deviation	1
Derivative of Spectral Variability Overall Average	1
Root Mean Square Overall Standard Deviation	1
Root Mean Square Overall Average	1
Derivative of Root Mean Square	1
Zero Crossings	1
Derivative of Zero Crossings Overall Average	1
Strongest Frequency via FFT Maximum Overall Standard Deviation	1
Mel-frequency Cepstral Coefficients (MFCC) Overall Standard Deviation	13
Mel-frequency Cepstral Coefficients (MFCC) Overall Average	13
Linear Predictive Coeffecients (LPC)	9
Derivative of Strongest Frequency Via Zero Crossing Overall Average	1
Derivative of Strongest Frequency Via Spectral Centroid Overall Average	1
Strongest Frequency via FFT Maximum Overall Average	1

Training set resulted five sets namely training set with 19, 29, 31, 32 and 50 features respectively. Each feature has 10 instances per chords. While datasets were in the form of CSV file. In order to test the accuracy of the test set that was created, WEKA creates data model using J48 classifier.

While testing J48 classifier accuracy was done through the prototype and WEKA. The study testing was implemented in this structure: Test the classifier with the different training sets to the test set of chords with a total of 1377 chords.

B. *Results of Testing*

After using the decision tree generated by WEKA as shown in Fig. 8, the summary of results of different training sets tested in the test sets using the prototype are shown in Table VIII where the correctly classified chords are averaged to get the success rate.



```
=== Classifier model (full training set) ===

J48 pruned tree
------------------

MFCC8 Overall Average <= -1.762
    MFCC10 Overall Average <= -0.8102
        MFCC2 Overall Average <= 3.526
            MFCC1 Overall Average <= 8.569
                MFCC10 Overall Average <= -1.891: D#m7 (3.0)
                MFCC10 Overall Average > -1.891: D (2.0/1.0)
            MFCC1 Overall Average > 8.569
                MFCC2 Overall Average <= 2.164: A7 (5.0/1.0)
                MFCC2 Overall Average > 2.164: Am7 (4.0)
        MFCC2 Overall Average > 3.526
            MFCC5 Overall Average <= 0.2657
                MFCC2 Overall Average <= 5.068: C#m7 (6.0/1.0)
                MFCC2 Overall Average > 5.068: A# (3.0/2.0)
            MFCC5 Overall Average > 0.2657: Dm7 (6.0/1.0)
    MFCC10 Overall Average > -0.8102
        MFCC5 Overall Average <= -0.2325
            MFCC12 Overall Average <= -0.9348
                MFCC11 Overall Average <= -0.6126
                    MFCC2 Overall Average <= 4.32: A#m7 (7.0/2.0)
                    MFCC2 Overall Average > 4.32
                        MFCC Overall Average <= -111.7
                            MFCC3 Overall Average <= 4.029: Cm7 (2.0)
                            MFCC3 Overall Average > 4.029: Cm (2.0)
                        MFCC Overall Average > -111.7
                            MFCC3 Overall Average <= 4.944: Gm (2.0)
                            MFCC3 Overall Average > 4.944: A#7 (2.0/1.0)
```

Figure 8. Decision tree generated by J48 algorithm

© 2020 J. Adv. Inf. Technol.

TABLE VIII. RESULTS OF THE TESTING USING THE PROTOTYPE

Training Set No.	No. of Features	Correctly Classified	Success Rate
1	19	482	33.47%
2	29	479	27.01%
3	31	518	35.97%
4	32	507	35.21%
5	50	539	37.43%

After the actual testing of the prototype, WEKA was utilized to test the success rate of the test sets over the 1377 chords as depicted in Table IX.

TABLE IX. RESULTS OF TESTING USING WEKA

Training Set No.	Number of Features	Correctly Classified	Success Rate
1	19	756	52.5%
2	29	578	40.14%
3	31	733	50.9%
4	32	620	43.06%
5	50	754	52.36%

C. *Interpretation of Results*

Fig. 9 shows performance of the prototype where training sets namely 19, 31 and 50 features provided by WEKA have a higher success rate than 2 training sets namely 29 and 32 features manually selected. It also illustrates that as the chords features increases, it success rate is positively affected.

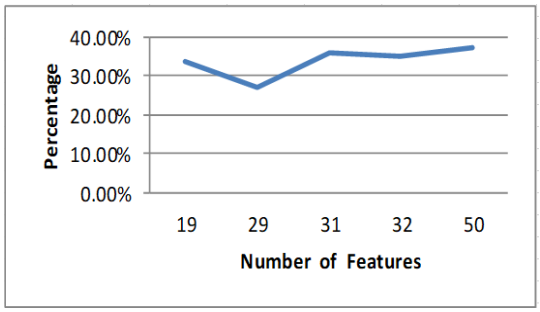


Figure 9. Result of prototype testing

Fig. 10 shows same findings as Fig. 8 Three training sets namely 19, 31 and 50 features provided by WEKA have a higher success rate than 2 training sets namely 29 and 32 features manually selected. But improvement on success rate from all training sets using WEKA shows an increase of 18% to 36% as depicted in Fig. 11.

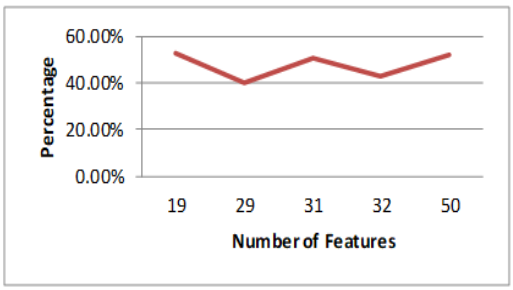


Figure 10. Result of WEKA testing

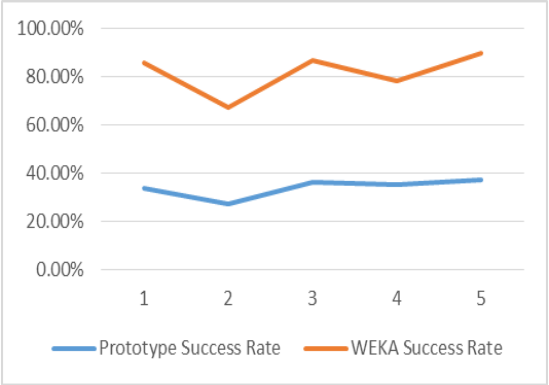


Figure 11. Accuracy rate of WEKA and prototype

TABLE X. COMPUTATION PROTOTYPE AND WEKA USING T-TEST

	Mean	Mean Difference	t-value	p-value
Prototype	33.818	-13.974	-4.422	0.003
WEKA	47.792			

T-test result as shown in Table X shows that the mean difference of the prototype's success rate and WEKA's success rate is -13.974 is significant in favor of WEKA based on the p-value of 0.003 which is less than 0.05.

As a general conclusion, the process presented to transcribe guitar chords from an audio file able to labels each frame of A, A#, B, C, C#, D, D#, E, F, F#, G, G# using J48 classifier generated by WEKA. J48 classifier is capable to transcribe guitar chords from acoustic audio provided that there is a comprehensive experiment of selecting good features. Ambiguities on related chords must be analyzed better by increasing training sets due to different time interval of cropping recorded chords in order to be compatible in other songs that have fast-strumming measure. Furthermore, integration of attack-decay concept to automatically slice audio frame be able to improve prototype functionality.

IV. FUTURE WORKS

Several consideration can be done in order to increase the success rate of detecting the guitar chord. The researchers recommend inclusion of Mel-frequency Cepstral Coefficients (MFCC) as part of the features in the dataset because as they have concluded in our study, they are the strongest features comprising the dataset. Increase the number of recordings of each chord to ensure that there are different variations of the same chords is part of the training set. Introduce variation of time interval in cropping recorded chords in order to be compatible in other songs that have fast strumming measure. The prototype can improve its functionality if there is a module that automatically slices an audio file according to the attack-decay concept of chords. Consider using other feature extraction tools such as jSymbolic, which is also an open-source Music Information Retrieval (MIR) tool, to provide appropriate features aligned to

© 2020 J. Adv. Inf. Technol.

acoustic guitar sounds provided that the study focuses on MIDI audio format. In implementing the J48 decision tree generated by WEKA into the prototype, consider the weights of attributes included in the decision tree.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Ilao initially identified the research topic and methodology. Pancho performed the chord dataset collection. All authors evaluated different chord features using J48 algorithm. The classifier was integrated by Nase and Talavera to developed prototype. Pancho, Nase and Talavera determined the accuracy of the model from collected datasets. All authors participated and approved the final paper revision.

ACKNOWLEDGMENT

Authors would like to express their greatest gratitude to the financial support given by the Malayan Colleges Laguna, Philippines.

REFERENCES

[1] S. Thomas, S. Peter, S. Christine, and H. David, "The psychological functions of music listening," *Frontiers in Psychology*, vol. 4, no. 511, 2013.
[2] J. Lutz, "Music, memory and emotion," *Journal of Biology*, vol. 7, no. 21, 2008.
[3] G. Alinka and L. Alexandra, "Music preference in adulthood: Why do we like the music we do?" in *Proc. 9th International Conference on Music Perception Cognition*, 2006.
[4] E. A. Scott, "Lifelong learning: Adults' perceptions of learning a musical instrument," *Innovation and Empowerment: SNU-Tulsa Research Journal*, vol. 4, 2012.
[5] M. Brain. (2000). How acoustic guitars work. [Online]. Available: http://entertainment.howstuffworks.com/guitar.htm
[6] F. A. Baker, "Review of music, health, and well-being," *Psychomusicology Music Mind & Brain*, vol. 23, no. 2, pp. 131-133, 2013.
[7] 1.4 billion guitar players. [Online]. Available: https://www.quora.com/How-many-people-in-the-world-play-guitar
[8] L. Coffey. (2010). What is a musical chord? [Online]. Available: http://www.elpin.com/tutorials/musicalchord.php
[9] L. Alcabasa and N. Marcos, "Simple audio processing approaches for guitar chord distinction," in *Proc. the DLSU Research Congress*, 2016.
[10] M. Cory and F. Ichiro, "JSymbolic: A feature extractor for MIDI files," in *Proc. ICMC*, 2006.
[11] S. Goldstein, and Y. Moses, "Guitar music transcription from silent video," in *Proc. BMVC*, 2018, p. 309.
[12] T. Gagnon, S. Larouche, and R. Lefebvre, "A neural network approach for preclassification in musical chords recognition," in *Proc. Thirty-Seventh Asilomar Conference on Signals, Systems and Computers*, 2004, vol. 2, pp. 2106-2109.
[13] A. Barbancho, "Automatic transcription of guitar chords and fingering from audio," *IEEE Transactions on Audio, Speech, and Language Processing*, 2012.
[14] M. Cory, F. Rebecca, M. Daniel, L. Beinan, and F. Ichiro, "ACE: A framework for optimizing music classification," in *Proc. ISMIR 6th International Conference on Music Information*, 2005, pp. 42-49.
[15] R. L. White. (1996). Methods for classification. [Online]. Available: http://sundog.stsci.edu/rick/SCMA/node2.html
[16] C. McKay, "jAudio: Towards a standardized extensible," *Theory and Practice of Computation*, 2006.
[17] C. McKay, *Automatic Music Classification with jMIR*, Montreal: McGill University, 2010.
[18] C. McKay, "jAudio: Towards a standardized extensible audio music feature extraction system," course paper, McGill University, Canada, 2009.
[19] S. B. Jagtap, "Census data mining and data analysis using WEKA," arXiv preprint arXiv:1310.4647, 2013.
[20] A. R. Onik, N. F. Haq, L. Alam, and T. I. Mamun, "An analytical comparison on filter feature extraction method in data mining using J48 classifier," *International Journal of Computer Applications*, vol. 124, no. 13, 2015.
[21] T. Gagnon, S. Larouche, and R. Lefebvre, "A neural network approach for preclassification in musical chords recognition," in *Proc. the Thirty-Seventh Asilomar Conference on Signals, Systems & Computers*, 2003, vol. 2, pp. 2106-2109.
[22] F. Mazhar, "Automatic guitar chord detection," master of science thesis, Department of Signal Processing, Tampere University of Technology, 2011.

Copyright © 2020 by the authors. This is an open access article distributed under the Creative Commons Attribution License (CC BY-NC-ND 4.0), which permits use, distribution and reproduction in any medium, provided that the article is properly cited, the use is non-commercial and no modifications or adaptations are made.

Adomar L. Ilao is a faculty member of Malayan Colleges Laguna under the College of Computer and Information Science. He is a member of International Innovative Scientific & Research Organization (IISRO) and Information Association of Computer Science and Information Technology (IACSIT). Furthermore, he is an active paper reviewer for local and international conferences. His passion revolves in the field of Computer Vision, Data Mining, Mobile Application Development, Gamification and Natural Processing Language publications. Several papers were indexed by SCOPUS and IEEE.

Darrell A. Talavera graduated with a degree in Bachelor of Science in Computer Science graduate of Malayan Colleges Laguna with academic award of Cum Laude. During his college days, he is a consistent dean lister and officer of college and institution organization. As of now, he worked as mobile application developer and software engineer in different companies in the Philippines and Canada. He is part of previous research paper in computer vision published in Springer Link using Haar Classifier to identify car sticker.

Leonardo Pancho is a graduate of Bachelor of Science in Computer Science graduate of Malayan Colleges Laguna. He is a Full Stack .Net Developer with 6 years of experience in IT industry. He is an experience multi-platform applications using ASP.NET, SQL, Web services, JavaScript, CSS, Cordova, and Mobile developer since 2013. He also managed to develop and enhance project applications using automated programs created in C# language. He is involved from a previous published focus on computer vision applied in automated car sticker detection using Haar Classifier published in Springer.

Erika Svetlana C. Nase is a graduate of Bachelor of Science in Computer Science graduate of Malayan Colleges Laguna with an academic award of Cum Laude. After graduation, she worked as senior software engineer in one of system development company in the Philippines. Moreover, she published a paper in Springer emphasis on automated car sticker detection. It was tested in her alma mater during her third year college days.