PREFACE

The 2nd International Conference on Applied Sciences and Smart Technologies (InCASST 2025) engages a theme of "Innovating for Sustainability: Digitalization, Green Energy, and Achieving Energy Independence for a Greener Future". The scope of the conference covers green technologies, environmental developments, and also digital societies. This event addresses various perspectives on how digital innovation in the green energy sectors change the global landscape towards energy independence and support more sustainable development. Accordingly, worldwide academics and practitioners were invited to either share their insights or exchange their ideas.

The dissemination of the ideas was expanded and became particular contribution to be offered in solving various problems, and earning the better quality of life. The initiation of the ideas was engaged by four distinguished keynote speakers, i.e. Professor Thomas Götz (Dept. of Mathematics, University of Koblenz-Germany), Prof. Dr. Kavita Sonawane (Dept. of Computer Engineering, SFIT, University of Mumbai-India), Ishak Hilton Pujantoro Tnunay, Ph.D (Beehive Drones/PT. Aerotek Global Inovasi-Indonesia), and also Andreas Prasetyadi, Ph.D. (Dept. of Mechanical Engineering, Universitas Sanata Dharma-Indonesia).

To overcome highly scientific impacts, and achieve a reputable scientific dissemination, a scientific board which is reliable and of high-caliber in the scope of conference was established. The determination of member of board was carefully carried out to meet the relevance to the scope of the conference. It further covers international members from several countries.

In addition, considering that publishing scientific papers involves several key steps to ensure the works are effectively presented, and reaches the intended audiences, series of procedures were carefully conducted during papers' review and selection. They cover contextual reviews of the received abstracts, editorial screening of its basic compliance and relevance, comprehensive reviews to address criteria of well-structured scientific manuscripts that must include clear introduction, methodology, results, discussions, and references. To overcome this hierarchal stages, a series of smart works were properly carried out by Prof. Peerapong Uthansakul, Ph.D., Dr. Eng. Ir. I Made Wicaksana Ekaputra, and Ir. Damar Widjaja, Ph.D.

Moreover, a smooth collaboration has been well established between Universitas Sanata Dharma and several other institutions as the co-hosts, i.e. Universitas Prasetiya Mulya, Universitas Surabaya, Universitas Pignatelli Triputra and also Universitas Atma Jaya Yogyakarta. Here also, a collaborative works with PT. Wijaya Kusuma Contractors was exhibited during this conference.

All in all, this prestige event was well organized by a generous team with a high commitment and rigorous determination. This committee was also fully supported by Faculty of Science and Technology, Universitas Sanata Dharma-Indonesia. We do hope future collaborations can be enhanced in pursuing the better world.

Yogyakarta, October 15th, 2025
Organizing Committee of InCASST 2025

Scientific Committee

- Prof. Hideki Kuwahara, Ph.D. (Sophia University, Japan)
- Prof. Dr. Willy Susilo ((Wollongong University, Australia)
- Prof. Dr. rer. nat. Herry Pribawanto Suryawan (Universitas Sanata Dharma - Indonesia)
- Prof. Ir. Sudi Mungkasi, Ph.D. (Universitas Sanata Dharma - Indonesia)
- Prof. Bambang Soelistijanto, Ph.D. (Universitas Sanata Dharma - Indonesia)
- Prof. Dr. Ir. Feri Yusivar, M.Eng. (Universitas Indonesia, Indonesia)
- Prof. Dr. Eng. Ir. Deendarlianto, S.T., M.Eng. (Gadjah Mada University, Indonesia)
- Dr. Uday Pandit Khot (St. Francis Institute of Technology, University of Mumbai, India)
- Dr. Farah Suraya Md Nasrudin (Universiti Teknologi MARA, Malaysia)
- Assoc. Prof. Dr. Shi Bai (Florida Atlantic University, U.S)
- Parkpoom Phetpradap, Ph.D. (Chiang Mai University, Thailand)
- Tri Hieu Le, Ph.D. (Ho Chi Minh University of Technology (HUTECH), Vietnam)
- Alex Baenaorama, Ph.D. (St. Paul University Dumaguete, Philipine)
- Ignasia Handipta Mahardika, Ph.D. (Sogang University, South Korea)
- Dr. Irma Saraswati, M.Sc. (Universitas Sultan Ageng Tirtayasa, Indonesia)
- Dr. Imamul Muttakin (Universitas Sultan Ageng Tirtayasa, Indonesia)
- Dyonisius Dony Ariananda, S.T., M.Sc., Ph.D. (Gadjah Mada University, Indonesia)
- Dr. I Gusti Ngurah Bagus Catrawedarma, S.T., M.Eng. (Politeknik Negeri Banyuwangi, Indonesia)

Steering Committee

- Chairperson: Ir. Drs. Haris Sriwindono, M.Kom., Ph.D. (Universitas Sanata Dharma - Indonesia)
- Member:
  - Dr. Ir. I Gusti Ketut Puja (Universitas Sanata Dharma - Indonesia)
  - Prof. Dr. Frans Susilo, S.J. (Universitas Sanata Dharma - Indonesia)
  - Dr. Ir. Bernadeta Wuri Harini (Universitas Sanata Dharma - Indonesia)
  - Prof. Dr. Ir. Anastasia Rita Widiarti (Universitas Sanata Dharma - Indonesia)
  - Dr. Lusia Krismiyati Budiasih (Universitas Sanata Dharma - Indonesia)

Organizing Committee

- Chairman: Dr. Ir. Achilleus Hermawan Astyanto (Universitas Sanata Dharma - Indonesia)
- Member:
  - Dr. Sri Hartati Wijono, M.Kom. (Universitas Sanata Dharma - Indonesia)
  - Ir. Kartono Pinaryanto, S.T., M.Cs. (Universitas Sanata Dharma - Indonesia)
  - Regina Chelinia Erianda Putri, M.T. (Universitas Sanata Dharma - Indonesia)
  - Ir. Rosalia Arum Kumalasanti, M.T. (Universitas Sanata Dharma - Indonesia)
  - Michael Seen, M.Eng. (Universitas Sanata Dharma - Indonesia)
  - Ir. Robertus Adi Nugroho, M.Eng. (Universitas Sanata Dharma - Indonesia)
  - Eduardus Hardika Sandy Atmaja, Ph.D. (Universitas Sanata Dharma - Indonesia)
  - Gilang Arga Dyaksa, M.Eng. (Universitas Sanata Dharma - Indonesia)
  - Heryoga Winarbawa, M.Eng. (Universitas Sanata Dharma - Indonesia)

Editors

- Prof. Peerapong Uthansakul, Ph.D. (Suranaree University of Technology - Thailand)
- Assoc. Prof. Dr. Eng. Ir. I Made Wicaksana Ekaputra (Universitas Sanata Dharma - Indonesia)
- Assoc. Prof. Ir. Damar Widjaja, Ph.D. (Universitas Sanata Dharma - Indonesia)

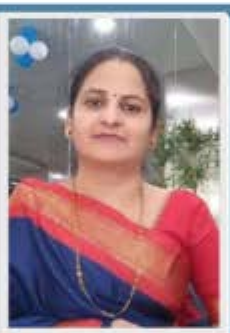
InCASST
International Conference on Applied Sciences and Smart Technologies

Innovating for Sustainability: Digitalization, Green Energy, and Achieving Energy Independence for a Greener Future

Yogyakarta, 15 October 2025 (Hybrid)

KEYNOTE SPEAKERS

Prof. Dr. Thomas Götz
University of Koblenz, Germany



Prof. Dr. Kavita Sonawane
SFIT, University of Mumbai, India

Ishak H.P. Tnunay, Ph.D.
Beehive Drones, PT. Aerotek Global Inovasi, Indonesia

Andreas Prasetyadi, Ph.D.
Universitas Sanata Dharma, Indonesia

Call for Papers

The 2nd International Conference on Applied Sciences and Smart Technologies (InCASST 2025)

SCOPES

Environmental Developments
- Environmental impact assessment and management
- Waste management and recycling
- Environmental biotechnology and microbiology
- Carbon capture and sequestration

Green Technologies
- Renewable energy technologies and systems
- Climate change and global warming
- Sustainable agriculture and land use practices
- Clean energy
- Energy Efficiency
- Water-Energy Nexus
- Green Materials

Digital Society
- IoT and AI-based sustainability solutions
- Data analysis and predictive modeling for environmental sustainability
- Sustainable transportation and mobility solutions
- Data and distributing computing

REGISTRATION FEE

Offline Presenter
- Early bird : IDR 2,650,000 (USD 185)
- Late : IDR 2,900,000 (USD 205)

Online Presenter
- Early bird : IDR 2,500,000 (USD 175)
- Late : IDR 2,750,000 (USD 195)

Offline Non Presenter DR 150,000 (USD 15)

IMPORTANT DATES

- Early bird registration deadline : 14 July 2025
- Full paper submission deadline : 29 Aug 2025
- Acceptance notification : 12 Sept 2025
- Late registration deadline : 15 Sept 2025
- Conference day : 15 Oct 2025

PAYMENT METHODS

- Bank Transfer
  Acc. Number : 8466045484
  Acc. Name : Michael Seen
  Bank Name : Bank Central Asia (BCA)
- Paypal : incasst@usd.ac.id
  Swift Code : CENAIDJA

CONTACT PERSON

Eduardus Hardika Sandy Atmaja, Ph.D. +62 895 3817 54488
Dr. Sri Hartati Wijono, M.Kom. +62 811 2646 471

Scan here

Co Hosted by

UBAYA

UNIVERSITAS PIGNATELLI TRIPUTRA

Sponsored by

wkc wijaya kusuma contractors

Published by

E3S Web of Conferences

IJASST International Journal of Applied Sciences and Smart Technologies

incasst_usd   incasst@usd.ac.id   s.usd.ac.id/incasst

Statement of Peer review

In submitting conference proceedings to *Web of Conferences*, the editors of the proceedings certify to the Publisher that

1. They adhere to its **Policy on Publishing Integrity** in order to safeguard good scientific practice in publishing.
2. All articles have been subjected to peer review administered by the proceedings editors.
3. Reviews have been conducted by expert referees, who have been requested to provide unbiased and constructive comments aimed, whenever possible, at improving the work.
4. Proceedings editors have taken all reasonable steps to ensure the quality of the materials they publish and their decision to accept or reject a paper for publication has been based only on the merits of the work and the relevance to the journal.

Title, date and place of the conference

The 2nd International Conference on Applied Sciences and Smart Technologies (InCASST 2025)
October 15th, 2025
Yogyakarta

Proceedings editor(s):

1. Prof. Peerapong Uthansakul, Ph.D. (Guest Editor)
2. Assoc. Prof. Dr. Eng. Ir. I Made Wicaksana Ekaputra (Editor)
3. Assoc. Prof. Ir. Damar Widjaja, Ph.D. (Editor)

Date and editor's signature

Dec 18th, 2025

E3S Web of Conferences

Volume 687 (2026)

The 2nd International Conference on Applied Sciences and Smart Technologies (InCASST 2025)

Yogyakarta, Indonesia, October 15, 2025

P. Uthansakul, I M.W. Ekaputra and D. Widjaja (Eds.)

- Green Technologies & Digital Society

E3S Web of Conferences
eISSN: 2267-1242

https://doi.org/10.1051/e3sconf/202668702014

An Evaluation of the Harmonic Product Spectrum for Neural Network-Based Chord Recognition

Linggo Sumarno[1*]

[1]Sanata Dharma University, Electrical Engineering Study Program, 55282 Maguwoharjo, Yogyakarta, Indonesia,

Abstract. The low-power computing systems have come under great demand today. This is due to the increasing energy consumption of connected devices and digital infrastructure. In the domain of chord recognition, there is a challenge to find feature representation methods that are computationally low while still preserving a high level of accuracy. In this study, the effectiveness of the Harmonic Product Spectrum (HPS) as a feature representation method for neural network chord recognition is evaluated. This chord recognition can be targeted for a small-scale and low-power system. Experiments were carried out using eight different HPS levels, where increasing the HPS level corresponded to a proportional reduction in the input size of the neural network. Based on the experimental results, it was shown that using HPS level 7, the chord recognition system could achieve an accuracy of up to 97.14%. These results indicate that HPS level 7 can provide the optimal trade-off between computational efficiency and accuracy.

1 Introduction

The growing electricity use from digital infrastructure and connected devices has increased the global need for energy conservation. Beyond environmental implications, energy efficiency has also emerged as both an economic and also functional necessity. As a consequence, energy-saving strategies can no longer be addressed solely at the macro level. They must be addressed at the micro level via the design of optimized algorithms, efficient models, and energy-efficient hardware [1].

One practical approach for energy savings achievement is the development of low-power small systems. Devices based on the TinyML platform [2] can enable on-device processing. Therefore, they can reduce the need for energy-intensive communication with data centers. Recent literature highlights the importance of combining low-power hardware design with software optimizations—such as quantization, pruning, and lightweight architectures—to enable neural network inference on power-constrained devices [3].

* Corresponding author: lingsum@usd.ac.id

In the domain of musical audio processing, chord recognition traditionally relies on spectral feature representation methods that are sensitive to the harmonic structure of the signal. Techniques such as the Enhanced Pitch Class Profile (EPCP) [4] often employ operations that emphasize harmonic components—most notably through the Harmonic Product Spectrum (HPS) [5]—as a preliminary step for mapping to chord classes. HPS can serve as a computationally inexpensive choice [6] for extracting pitch and harmonic information prior to classification.

Although modern neural network architectures, including Conformer- and Transformer-based models, have improved chord recognition accuracy across large datasets [7], their computational complexity and power requirements are often incompatible with the constraints of low-power small systems. This creates a clear research need to explore computationally efficient feature representations that can be paired with compact neural networks to achieve an optimal trade-off between accuracy and energy consumption.

In this context, the present study is a preliminary investigation that evaluates the use of the HPS as a feature representation method for neural network-based chord recognition, targeting low-power small systems. The main contribution lies in an empirical assessment of the computationally efficient HPS method for representing chord features in neural network-based recognition systems. The findings are expected to provide practical guidance for developing sustainable chord recognition solutions aligned with one of the core green technology principles, namely energy efficiency.

2 Methodology

2.1 Developed system

Figure 1 illustrates the chord recognition system developed for this study, presenting the complete block diagram of its components. It is important to note that the system was implemented using the Python programming language with the PyTorch package. The following subsections provide a detailed explanation of each block shown in Figure 1.

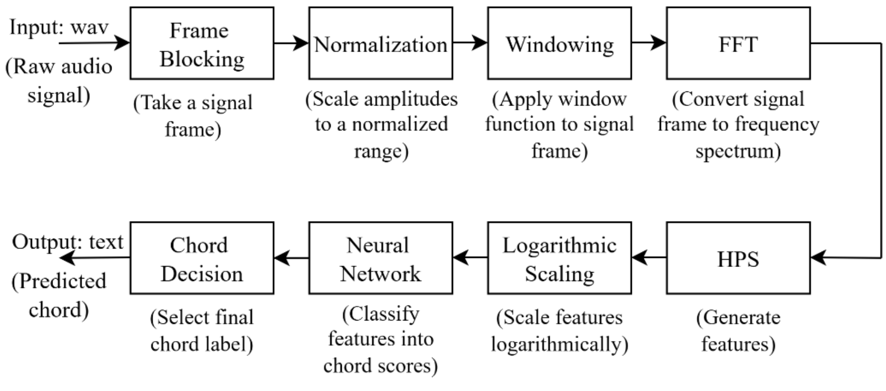


Fig. 1. Chord recognition system used in this study.

2.1.1 Input and output

The system input is an isolated guitar chord recording in the form of an audio file in .wav format with a fixed sampling rate of 5000 Hz (mono) and a duration of 2 seconds. The recordings were conducted using a Yamaha CPX-500-II guitar in a quiet environment where each recording used only one variation in playing the guitar chords. It were obtained from the dataset that available at https://github.com/lingsum/Chord-DST-DWT. The system output is a text label—C, D, E, F, G, A or B—indicating the chord recognized by the system.

2.1.2 Frame blocking

The signal from the input file is segmented into a fixed-length signal frame, a process referred to as frame blocking. In this study, the signal frame length was set to 2^n samples. This choice was made because the subsequent Fast Fourier Transform (FFT) stage employed a radix-2 FFT algorithm, which is computationally more efficient when the signal frame length is a power of two. Specifically, $2^{13} = 8192$ samples were taken from the total 10,000 samples available in each .wav file. Each .wav file contains a 2-second recording sampled at 5000 Hz.

2.1.3 Normalization

Each frame obtained from the frame blocking process was normalized to mitigate amplitude variability effects during recording. The normalization process is mathematically expressed as:

$$x_{norm}(n) = \frac{x(n)}{\max(|x(n)|)} \quad (1)$$

where $x_{norm}(n)$ is the normalized signal frame, and $x(n)$ is the signal frame obtained from the frame blocking process.

2.1.4 Windowing

The windowing process was applied to reduce spectral leakage effects. These kind of effects occur as a result of frame blocking process. In windowing process, each signal frame was multiplied by a window. This study employed the Hamming window. This kind of window is a commonly used window in various audio and speech signal processing applications [8-9]. The Hamming window $w(n)$ is defined as:

$$w(n) = 0.54 - 0.46\cos\left(\frac{2\pi n}{N-1}\right) \quad (2)$$

where N is the signal frame length after frame blocking process. The windowing process result $x_w(n)$ is then given by:

$$x_w(n) = x_{norm}(n) \, . \, w(n) \quad (3)$$

2.1.5 Fast Fourier Transform (FFT)

The Fast Fourier Transform (FFT) was applied to convert the time-domain signal into the frequency domain signal. This transformation is an efficient method for computing the Discrete Fourier Transform (DFT):

$$X(k) = \sum_{n=0}^{N-1} x_w(n)\, e^{-\frac{j2\pi kn}{N}} \quad (4)$$

where $X(k)$ is the DFT of windowing process result $x_w(n)$. In this study, FFT was used to construct a chord representation based on frequency spectrum. Only the left half of the FFT output was used in the subsequent processing, as the FFT result is symmetric, and the left half contains sufficient information. Additionally, the DC (Direct Current) component $X(0)$ was set to zero because it corresponds to the zero-frequency term, which is unrelated to the fundamental frequencies of chords.

2.1.6 Harmonic Product Spectrum

HPS process is primarily employed to substantially reduce the amount of data used for classification by the neural network. From a signal processing perspective, HPS is computed by multiplying the magnitude spectrum of a signal by the magnitude spectra obtained from down sampled versions of the same signal. This study adopts the HPS formulation originally introduced by Noll [5], but rewritten as follows:

$$HPS(k) = \prod_{m=1}^{v} |X(k \cdot m)| \tag{5}$$

where $HPS(k)$ is the result of HPS process, and $v = 1, 2, 4, \ldots, 2^L$, with $L = 0, 1, 2, \ldots$ denotes the HPS level. There are two important notes regarding this implementation. Firstly, for $L = 0$, the HPS process is bypassed. Secondly, in the practical programming implementation, if a value in $HPS(k)$ is detected to overflow, it is replaced with a large constant value of 1×10^5. Figure 2 illustrates the practical HPS process for a signal length of 8.

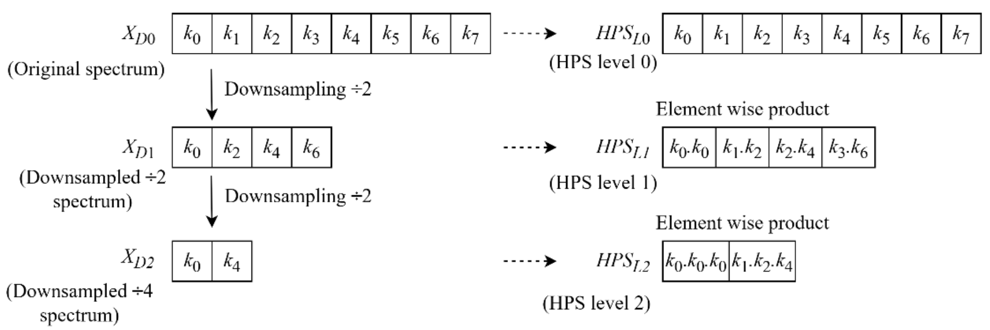


Fig. 2. Illustration of the practical HPS process for a signal length of 8.

2.1.7 Logarithmic scaling

The logarithmic scaling is applied to reduce the relative differences between the peaks in the HPS output. It is noted that the HPS process can produce extreme spectral peaks; applying the logarithm "flattens" these peaks while preserving finer spectral details. The logarithmic transformation is defined as:

$$S_{log}(k) = \log_{10}(1 + HPS(k)) \tag{6}$$

where $S_{log}(k)$ is the result of logarithmic scaling process, and $HPS(k)$ is the result of HPS process. The addition of 1 in equation (6) ensures that the function domain remains positive. In this context, the domain refers to the magnitude domain of the signal, which inherently contains only positive values.

2.1.8 Neural network

The logarithmic output of the HPS process was used as the input feature for a Multilayer Perceptron (MLP) neural network with a single hidden layer. The choice of a single hidden layer was made as part of a strategy to minimize the neural network's size.

The neural network structure is defined as follows:

a. Input layer size corresponds to the length of the HPS feature vector.
b. Hidden layer size corresponds to the average of the input layer size and output layer size [10].
c. Output layer size corresponds to the seven chord classes (C, D, E, F, G, A and B).

A sigmoid activation function was used in the hidden layer neurons. It was chosen because it is simple, non-linear, and globally differentiable. Non-linearity allows the network to learn complex patterns, while global differentiability ensures that gradients are available across the entire domain, supporting effective training [11].

The network was trained using the Adam optimization algorithm, a widely used method in neural network training because of its efficiency and adaptability [12-13]. In this study, the Adam optimizer was run with its default learning rate of 0.001 [14]. A batch size of 5 was also used, as it is well-suited for training with relatively small datasets [15].

2.1.9 Chord decision

The neural network output is a vector containing the output values (scores) of each output neuron. In this case, each neuron is labeled according to a specific chord (C, D, E, F, G, A, or B). The recognized chord is determined by selecting the chord corresponding to the output neuron with the highest value.

2.2 Training and testing

The training process was performed using a total of 70 chord datasets, with each chord class containing 10 samples. The testing process was performed using a total of 140 chord datasets, with each chord class containing 20 samples.

3 Results, analysis and discussion

3.1 Results

The experimental results observing the effect of varying HPS levels and the number of training epochs on the recognition accuracy of the neural network are presented in Table 1. Meanwhile, Table 2 presents the relationship between HPS level variation and the neural network input size.

3.2 Analysis

3.2.1 Analysis of training and testing results (Table 1)

HPS levels 0–6
At these levels, recognition accuracy reached 100% across all training epochs (20, 40, 60, 80, and 100). This indicates that the neural network was able to learn the data patterns

perfectly, even with a relatively small number of epochs (20). Such performance suggests that the essential information in the input features remained intact and was easily learnable by the neural network.

HPS level 7
Accuracy decreased to 84.29% at 20 epochs but improved significantly, reaching 97.14% at 80–100 epochs. This trend suggests that a slight loss of essential information occurred at this level, requiring more training epochs for the model to approach maximum performance.

Table 1. Training and testing results of the neural network for various HPS levels and numbers of epochs. Results shown: Accuracy (%).

HPS level	**Number of epochs**				
	20	40	60	80	100
0	100.00	100.00	100.00	100.00	100.00
1	100.00	100.00	100.00	100.00	100.00
2	100.00	100.00	100.00	100.00	100.00
3	100.00	100.00	100.00	100.00	100.00
4	100.00	100.00	100.00	100.00	100.00
5	100.00	100.00	100.00	100.00	100.00
6	100.00	100.00	100.00	100.00	100.00
7	84.29	91.43	96.43	**97.14**	97.14
8	62.14	67.14	70.71	72.86	73.57

Table 2. Relationship between HPS level and neural network input layer size.

HPS level	0	1	2	3	4	5	6	**7**	8
Neural network input layer size	4096	2048	1024	512	256	128	64	**32**	16

HPS level 8
The lowest performance among all evaluated levels, with an initial accuracy of 62.14% at 20 epochs and only reaching 73.57% at 100 epochs. Despite the improvement with additional epochs, the results remain substantially lower than the preceding levels. This performance drop indicates that excessive dimensionality reduction led to a significant loss of essential information.

3.2.2 Relationship between HPS level and neural network input layer size (Table 2)

The relationship
As shown in Table 2, there is a trend: as the HPS level increases, the neural network input layer size decreases significantly. Specifically, at level 0, the neural network has 4,096 inputs, while at level 8, the neural network has only 16 inputs. This trend confirms that HPS is a dimensionality reduction method that progressively compresses a higher dimensional representation into a much lower dimension representation.

Impact of dimensionality reduction

Dimensionality reduction offers several advantages, including reduced computational load, faster training times, and lower memory requirements. However, when taken to the extreme, such as at level 8, it can lead to significant essential information loss, which in turn causes a noticeable decrease in recognition accuracy.

3.3 Discussion

3.3.1 Relationship Between Feature Complexity and Accuracy

For HPS Levels 0–6, the data representation retains sufficient complexity and completeness, enabling the model to achieve maximum accuracy with ease. At Level 7, the representation is less complex, with minor essential information loss preventing the model from reaching absolute maximum accuracy. At Level 8, the representation becomes oversimplified, leading to significant essential information loss and reduced pattern recognition capability, even with increased training epochs.

3.3.2 Role of training epochs

Increasing the number of epochs improved accuracy for HPS levels 7 and 8. However, the effect was limited. This indicates that the accuracy constraints were not caused by underfitting due to insufficient training but rather by the lack of essential information in the input features.

3.3.3 Practical implications

HPS Level 7 offers the most practical balance between efficiency and accuracy. The use of Level 7 results in only a marginal reduction in accuracy compared to the substantial computational savings achieved.

3.3.4 Study Limitations

This study has several limitations. These limitations are described as follows.

a. The dataset used was limited to isolated chord recordings from a single guitar. This restriction reduced dataset variability in terms of timbre, instrument type, and performance style. Essentially, this limitation was intended to create an isolated condition for evaluating the HPS.
b. The dataset size was relatively small compared to common practices in neural network applications. The dataset consisted of 70 training samples (10 per class) and 140 testing samples (20 per class). This limitation was chosen because the goal of the study was not to develop a chord recognition system that ready for real-world deployment, but rather to conduct a controlled proof of concept.
c. Performance was evaluated solely on the accuracy metric. This metric is essentially straightforward and widely recognized in feasibility studies. Many chord recognition studies also begin with accuracy before introducing more complex performance metrics.
d. This study did not include comparisons with other feature representation methods (e.g., Mel-Frequency Cepstral Coefficients (MFCC), chroma features, Enhanced Pitch Class Profile (EPCP)), or benchmarks against existing chord recognition systems. This was intended to avoid broadening the scope and to maintain full focus on evaluating HPS as a feature representation.

From the four limitations described above, it can be said that these limitations were methodological choices, in order to keep this study focused. Nevertheless, this study still provides meaningful preliminary evidence of HPS's capability as a feature representation method.

4 Conclusion and future studies

4.1 Conclusion

Based on the experimental results, the application of the Harmonic Product Spectrum (HPS) method demonstrates that increasing the HPS level consistently reduces the neural network input dimensionality. At HPS levels 0–6, essential information is preserved, enabling the model to achieve perfect accuracy (100%). HPS level 7 retains most of the essential information, allowing the model to reach an accuracy of 97.14%, albeit requiring a longer training time. In contrast, HPS level 8 results in substantial loss of essential information, leading to a pronounced drop in accuracy.

From an efficiency perspective, HPS effectively reduces computational load. However, excessive dimensionality reduction can significantly degrade accuracy. Empirically, HPS level 7 offers the optimal trade-off between efficiency and accuracy, where the accuracy loss is relatively minor compared to the substantial computational savings achieved.

4.2 Future studies

Based on the limitations used in evaluating the Harmonic Product Spectrum (HPS), there are several directions for future studies. These directions are described as follows.

a. The dataset is expanded not only to isolated chords on a single guitar but also to include various instruments, diverse timbres, different playing styles, and longer chord progressions. In addition, the dataset is also expanded to complex polyphonic music recordings involving many instruments. This dataset expansion is intended to test whether the results obtained in the isolated condition can be extended to more realistic contexts.
b. The dataset size is increased significantly. This is to address concerns related to overfitting and the generalization ability of the neural network.
c. Evaluation uses not only the accuracy metric but also other metrics, such as confusion matrix, per-class accuracy, and precision/recall. These other metrics will provide more detailed insights regarding the specific strengths and weaknesses of each class.
d. The implementation of comparisons with other feature representation methods (e.g., MFCC, chroma features, EPCP), as well as benchmarking against baseline models or existing chord recognition systems. This can highlight the unique advantages of HPS compared to those other feature representation methods, as well as provide context for the findings in this study.

Electrical Engineering Study Program at Sanata Dharma University has supported this study.

References

1. A. Tabbakh, L.A. Amin, M. Islam, G.M.I. Mahmud, I.K. Chowdhury, and M.S.H. Mukta, Towards sustainable AI: a comprehensive framework for Green AI, Discov. Sustain., 5:408 (2024) https://doi.org/10.1007/s43621-024-00641-4
2. S. Heydari and Q. H. Mahmoud, Tiny Machine Learning and On-Device Inference: A Survey of Applications, Challenges, and Future Directions, Sensors, **25**(10), 3191 (2025) https://doi.org/10.3390/s25103191
3. M. Faheem, Energy Efficient Neural Architectures for TinyML Applications, Int. J. Sci. Res. Mod. Technol., **4**(5), 45-50 (2025) https://doi.org/10.38124/ijsrmt.v4i5.531
4. K. Lee, Automatic Chord Recognition from Audio Using Enhanced Pitch Class Profile, in Proceedings of International Computer Music Conference, New Orleans, 306-313 (2006) http://hdl.handle.net/2027/spo.bbp2372.2006.064
5. A.M. Noll, Pitch Determination of Human Speech by the Harmonic Product Spectrum, the Harmonic Sum Spectrum and a Maximum Likelihood Estimate, in Proceedings of Symposium on Computer Processing in Communications, Vol. 19, Polytechnic Press, New York, 779-797 (1970) https://cir.nii.ac.jp/crid/1570572699087005312
6. D. Coccoluto, V. Cesarini and G. Costantini, OneBitPitch (OBP): Ultra-High-Speed Pitch Detection Algorithm Based on One-Bit Quantization and Modified Autocorrelation, Appl. Sci. **13**(14), 8191 (2023) https://doi.org/10.3390/app13148191
7. M. W. Akram, S. Dettori, V. Colla and G.C. Buttazzo, ChordFormer: A Conformer-Based Architecture for Large-Vocabulary Audio Chord Recognition, arXiv (2025) https://doi.org/10.48550/arXiv.2502.11840
8. X. Yang, A study of the influence of audio signal processing technology on the expression of music aesthetics in piano performance, Appl. math. nonlinear sci., **10**(1), 1-18 (2025) https://doi.org/10.2478/amns-2025-0671
9. J. Saini, and R. Mehra, Power Spectral Density Analysis of Speech Signal using Window Techniques, Int. J. Comput. Appl. **131**(14), 33-36 (2015) https://doi.org/10.5120/ijca2015907549
10. J. Heaton, Introduction to Neural Networks with Java (2nd ed.) (Heaton Research Incorporated, Chesterfield, 2008)
11. P. Gopalani, A. Mukherjee, Global convergence of SGD on two layer neural nets, IMA J. Inf. Inference, 14(1) (2025) https://doi.org/10.1093/imaiai/iaae035
12. I.K.M Jais, A.R. Ismail A.R. and Nisa, S. Q, Adam Optimization Algorithm for Wide and Deep Neural Network, Kno. Eng. Da. Sc. **2**(1), 41-46 (2019) https://doi.org/10.17977/um018v2i12019p41-46
13. A. Bhattacharjee, Improving the Adaptive Moment Estimation (ADAM) stochastic optimizer through an Implicit-Explicit (IMEX) time-stepping approach, J. Mach. Learn. Model. Comput. **5**(3), 47-68 (2024) https://doi.org/10.1615/JMachLearnModelComput.2024053508
14. D.P. Kingma and J.L. Ba, Adam: A Method for Stochastic Optimization, in Proceedings of 3rd International Conference on Learning Representation, San Diego, May 7-9 (2015) https://doi.org/10.48550/arXiv.1412.6980
15. D. Masters and C. Luschi, Revisiting Small Batch Training for Deep Neural Networks, arXiv (2018) https://doi.org/10.48550/arXiv.1804.07612