JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Dimensionality Reduction for Chord Recognition Features in Music Information Retrieval

[1]Joseph R

[1] Students MTech in Artificial Intelligence
[1]Reva Academy for Corporate Excellence (RACE),
[1]REVA University, Bengaluru, India

Abstract: Music spans a wide spectrum of styles, from classical and jazz to rock, pop, and folk. At the heart of understanding harmony lies chord recognition, a crucial task for both music analysis and performance. While modern audio processing techniques such as chromagrams, MFCCs, and pitch class profiles have made chord recognition feasible, they also produce large, high-dimensional feature spaces. These high-dimensional representations can introduce redundancy, slow down computation, and reduce classification accuracy due to overfitting.

To address these challenges, this study explores the role of dimensionality reduction in chord recognition. The proposed system processes audio recordings into feature vectors and applies three reduction techniques: Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), and Uniform Manifold Approximation and Projection (UMAP). These transformed features are then classified using machine learning models such as Random Forest, Gradient Boosting, Logistic Regression, and Support Vector Machines.

Experiments conducted on annotated chord datasets show that not all dimensionality reduction techniques contribute equally. While PCA and UMAP significantly degraded recognition accuracy, LDA improved classification results, achieving a peak accuracy of 94.21% with Random Forest. This demonstrates that supervised dimensionality reduction, which emphasizes class separability, is particularly effective for chord recognition tasks.

***IndexTerms – Music Information Retrieval, Chord Recognition, Dimensionality Reduction, PCA, LDA, UMAP, Machine Learning*.**

I. INTRODUCTION

Chord recognition is one of the fundamental tasks in MIR, serving as a crucial building block for higher-level applications such as musical chord transcription, harmonic analysis, recommendation systems, and genre classification. Accurate chord recognition requires the extraction of both harmonic and timbral features from audio signals. Commonly used features include chroma, which captures pitch class distributions; MFCCs, which describe spectral and timbral characteristics; and pitch class profiles, which encode harmonic content. While these features are highly informative, their combination produces large, high-dimensional feature spaces.

High-dimensional data introduces several challenges. It increases computational complexity, creates redundancy, and often leads to poor generalization due to the curse of dimensionality. As a result, dimensionality reduction becomes a critical step in MIR pipelines. The goal of dimensionality reduction is to transform high-dimensional feature vectors into compact representations that preserve discriminative information while removing redundancy.

Although classification algorithms for chord recognition have been widely studied, the systematic evaluation of dimensionality reduction methods in this domain remains limited. This paper investigates the role of three popular dimensionality reduction techniques such as PCA, LDA, and UMAP — in chord classification. By comparing them against a baseline with no reduction, we aim to understand whether reducing dimensionality can improve accuracy, efficiency, and robustness in chord recognition tasks.

II. OVERVIEW OF DIMENSIONALITY REDUCTION METHODS

Dimensionality reduction techniques are used to convert high-dimensional feature sets into more compact representations in order to improve classification performance and lower computational complexity. These techniques seek to increase the efficacy and efficiency of chord recognition models by preserving important information while eliminating noise and redundancy. Different strategies are offered by the widely used Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), and Uniform Manifold Approximation and Projection (UMAP) methods. These strategies range from variance maximization to class-separability optimization and non-linear structure preservation.

2.1 Principal Component Analysis (PCA)

PCA is an unsupervised linear projection method that reduces dimensionality by projecting data onto orthogonal components that maximize variance. While PCA is effective at compressing data and mitigating redundancy, it does not account for class labels. Consequently, PCA may discard information critical for distinguishing between different chord classes, potentially leading to reduced classification accuracy.

2.2 Linear Discriminant Analysis (LDA)

LDA is a supervised dimensionality reduction technique that explicitly incorporates class information. It maximizes the separation between classes by optimizing the ratio of between-class variance to within-class variance. Since chord recognition is inherently a classification task, LDA is particularly well-suited for this problem, as it reduces dimensionality while preserving chord-specific discriminative features.

2.3 Uniform Manifold Approximation and Projection (UMAP)

UMAP is a non-linear manifold learning algorithm designed to preserve both local and global structure in the data. It has gained popularity for visualization and clustering tasks due to its ability to capture non-linear relationships. However, UMAP does not explicitly consider class labels during embedding, which may misalign with supervised classification objectives such as chord recognition.

III. RESEARCH METHOD

The research methodology involved processing audio recordings into labeled slices and transforming them into numerical feature representations for machine learning. Dimensionality reduction techniques were applied to compress these high-dimensional features into more compact forms. Classification models were then trained and evaluated on both the original and reduced feature sets to measure the effect of dimensionality reduction on chord recognition accuracy. The aim was to compare different reduction methods and assess their impact on efficiency and overall performance in chord recognition.

3.1 Flow Diagram

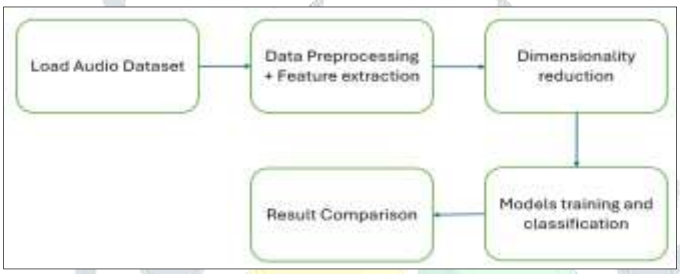

Fig 3.1: Flow Diagram

3.1.1 Dataset Construction and preprocessing:

The dataset consisted of audio recordings paired with annotation files. Audio was segmented into 1-second slices, each aligned with chord labels extracted from annotations. Only simple chord labels (major and minor) were retained, while complex chords (e.g., seventh, suspended, diminished) were excluded to ensure label consistency giving a total of 24 chord classes.

3.1.2 Feature Extraction:

From each slice, features were computed including chroma, MFCCs, and chroma-CENS. To capture both average tendencies and variability, the mean and standard deviation of each feature were calculated across frames. This produced high-dimensional feature vectors representing the harmonic and spectral content of each slice.

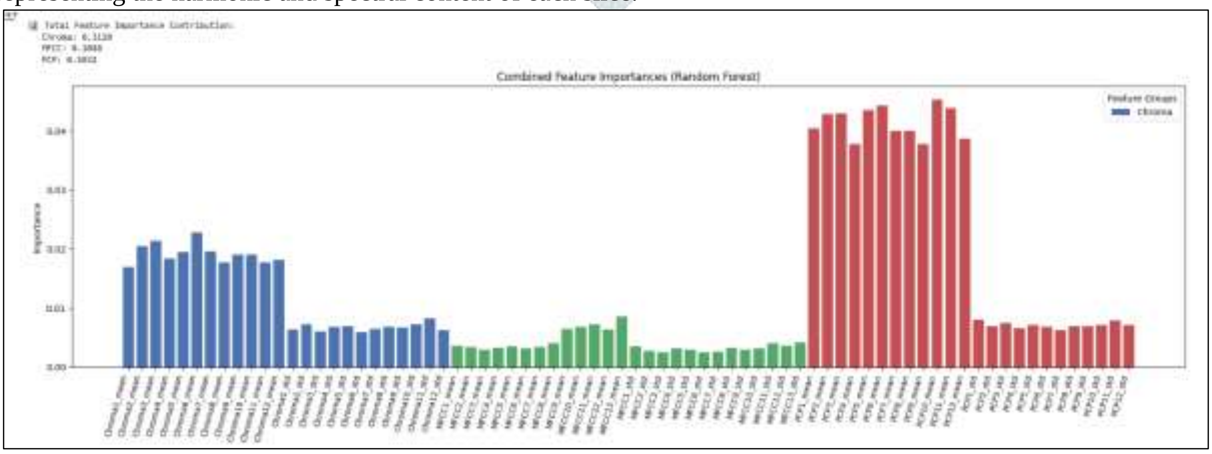

Fig:3.2 Feature extraction with 74 Baseline features

3.1.3 Dimensionality Reduction:

To address the redundancy in feature vectors, three dimensionality reduction techniques were applied: PCA (20 components), LDA (up to 20 components, constrained by the number of chord classes), and UMAP (20 components). A baseline condition with no reduction was also included for comparison. As shown in fig 3.2.

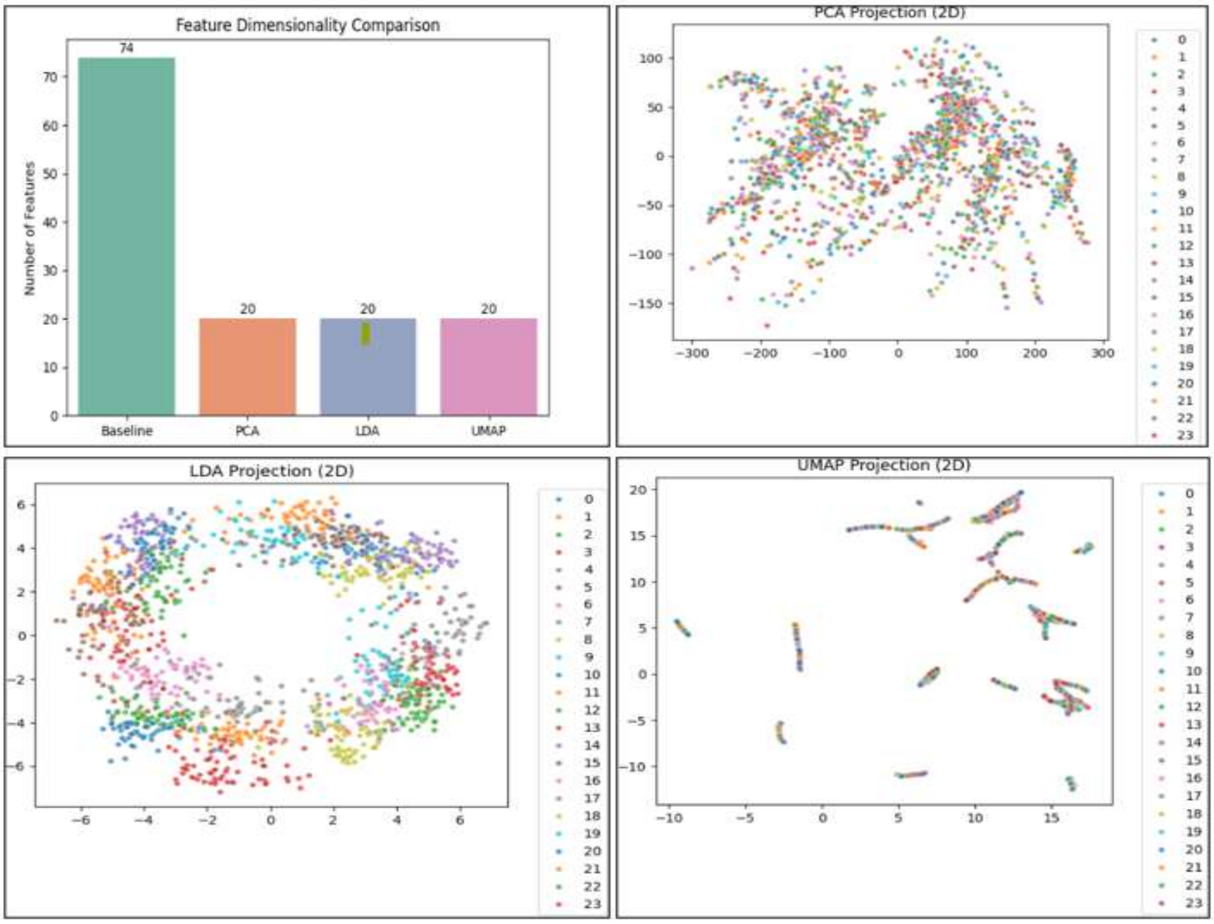

Fig 3.2: Dimension reductions.

3.1.4 Classification Models:

For this study, four widely used machine learning classification models were employed to analyze the musical features and predict chord labels: Random Forest, Gradient Boosting, Logistic Regression, and Support Vector Machines (SVM).

1. **Random Forest (RF):** An ensemble-based model that constructs multiple decision trees and aggregates their outputs to improve prediction accuracy and control overfitting. Its ability to handle large feature sets and complex, non-linear relationships make it well-suited for musical data, where interactions between features can be intricate.

2. **Gradient Boosting (GB):** Another ensemble approach that builds trees sequentially, each one focusing on the errors of the previous model. Gradient Boosting is effective in capturing subtle patterns in high-dimensional data, making it valuable for distinguishing between similar chord classes.

3. **Logistic Regression (LR):** A linear model commonly used for classification tasks. While simpler than ensemble methods, Logistic Regression provides interpretability and acts as a strong baseline, especially when features are well-engineered and linearly separable.

4. **Support Vector Machines (SVM):** SVMs are powerful classifiers that maximize the margin between different classes in high-dimensional space. Using kernel functions, SVMs can model complex, non-linear decision boundaries, which is beneficial for musical chord classification where feature relationships are rarely strictly linear.

3.1.5 Evaluation Metrics:

To evaluate the performance of the classification models, several standard metrics were used: accuracy, precision, recall, and F1-score**.**

1. **Accuracy:** Measures the overall proportion of correctly classified instances. While useful for an overall snapshot, accuracy can be misleading in cases where chord classes are imbalanced.

2. **Precision:** Indicates the proportion of correctly predicted instances among all instances predicted for a particular class. High precision ensures that when a model predicts a specific chord, it is likely to be correct.

3. **Recall (Sensitivity):** Represents the proportion of correctly predicted instances out of all actual instances of a class. High recall indicates that the model can identify most instances of a chord, reducing missed detections.

4. **F1-Score:** The harmonic mean of precision and recall, providing a single metric that balances both concerns. This is particularly important in chord classification, where some chords may be harder to distinguish, and a model must maintain a trade-off between over-predicting and under-predicting classes.

Collectively, these metrics provide a comprehensive evaluation, capturing not only the overall predictive power of the models but also their ability to correctly classify individual chord classes, which is critical for accurate music analysis.

IV. RESULTS:

The baseline performance using the full feature set was strong across all models, as shown in Table 4.1 and the accompanying figure. Random Forest achieved the highest accuracy at 93.42%, followed by Gradient Boosting (92.11%), Logistic Regression (91.32%), and SVM (88.68%). These results indicate that the original features were highly discriminative for chord classification, providing a solid foundation for evaluating dimensionality reduction methods.

Table 4.1: Accuracy Comparison

Model	Baseline	PCA	LDA	UMAP
Random Forest	93.42%	21.84%	**94.21%**	3.95%
Gradient Boosting	92.11%	16.05%	90.79%	3.16%
Logistic Regression	91.32%	17.63%	90.79%	1.32%
SVM	88.68%	16.84%	92.11%	1.05%

When dimensionality reduction was applied, performance varied considerably. PCA caused a sharp decline in accuracy for all models (16–22%), highlighting its limitation in supervised tasks since it ignores class labels. LDA, being supervised, improved results beyond baseline levels: Random Forest with LDA reached 94.21%, and SVM achieved 92.11%, demonstrating that LDA preserves class separability. In contrast, UMAP embeddings yielded very low accuracy (1–4%), showing that while UMAP is effective for visualization, it does not align with supervised chord classification objectives. These trends are clearly illustrated in figure 4.1, which compares model accuracies across all dimensionality reduction techniques.

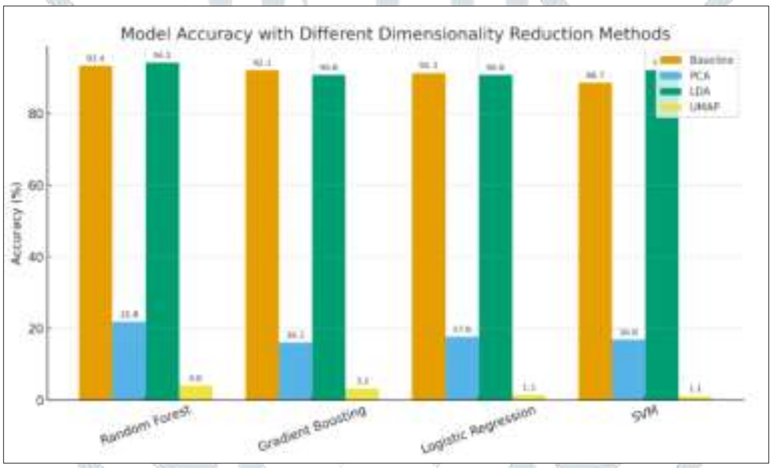

Fig:4.1: Model Accuracy with Different Dimensionality Reduction methods

V. CONCLUSION:

This study highlighted the critical role of dimensionality reduction in chord recognition for Music Information Retrieval (MIR). A 74-dimensional baseline of Chroma, MFCC, and PCP features, dimensionality was reduced to 20 or fewer features using PCA, LDA, and UMAP. While the baseline feature set already produced strong classification performance, it carried redundancy and computational overhead. PCA and UMAP, despite reducing dimensionality, significantly degraded accuracy because they did not incorporate class information. In contrast, LDA effectively preserved chord-specific discriminative features, improving performance across models and achieving the highest accuracy of 94.21% with Random Forest. These results demonstrate that dimensionality reduction is most beneficial when aligned with the classification objective, with supervised methods like LDA proving particularly effective for chord recognition tasks.

For future work, several directions can be explored. Hybrid approaches that combine supervised and unsupervised methods could balance variance preservation with class separability. Deep autoencoders and task-specific neural embeddings may provide more powerful compressed representations tailored to MIR tasks. Expanding experiments to include larger datasets and more complex chord classes (e.g., seventh, suspended, diminished) will enhance generalization and robustness. Finally, optimizing dimensionality reduction for real-time applications could lead to practical tools for music transcription, education, and interactive performance.

REFERENCES

[1]Velliangiri, S., Alagumuthukrishnan, S., & Joseph, S. I. T. (2019). A Review of Dimensionality Reduction Techniques for Efficient Computation. *Procedia Computer Science*, 165, 104–111. https://doi.org/10.1016/j.procs.2020.01.079

[2]You, S. D., & Hung, M. J. (2021). Comparative Study of Dimensionality Reduction Techniques for Spectral–Temporal Data. *Information*, 12(1), Article 1. https://doi.org/10.3390/info12010001

[3]Abiodun, O., Ajibade, S., & Sunoloso, F. (2024). The Impact of Dimensionality Reduction Techniques on Machine Learning Algorithm Efficiency. *Preprint*. Retrieved from https://www.researchgate.net/publication/386451975

[4]Pál, T., & Várkonyi, D. T. (2020). Comparison of Dimensionality Reduction Techniques on Audio Signals. *ResearchGate Preprint*. https://www.researchgate.net/publication/348416323

[5]Silipo, R., Adae, I., Hart, A., & Berthold, M. (2014). Seven Techniques for Dimensionality Reduction. *KNIME Whitepaper*. Retrieved from http://www.sigkdd.org/kdd-cup-2009-customer-relationship-prediction

[6]McInnes, L., Healy, J., & Melville, J. (2018). UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *arXiv preprint* arXiv:1802.03426. https://doi.org/10.48550/arXiv.1802.03426

[7]Yu, J., Ye, N., Du, X., & Han, L. (2022). Automated English Speech Recognition Using Dimensionality Reduction with Deep Learning Approach. *Wireless Communications and Mobile Computing*, 2022, Article ID 3597347. https://doi.org/10.1155/2022/3597347

[8]Ahmad, N., & Bou Nassif, A. (2022). Dimensionality Reduction: Challenges and Solutions. *ITM Web of Conferences*, 43, 01017. https://doi.org/10.1051/itmconf/20224301017

[9]Van der Maaten, L. J. P., Postma, E. O., & Van den Herik, H. J. (2007). Dimensionality Reduction: A Comparative Review. *Journal of Machine Learning Research*, 10, 66-71.