Journal of Informatics and Data Science (J-IDS)

Vol. 3, No. 2, November 2024
e-ISSN: 2964-0415

Melody Transcription from Monophony Audio with Fast Fourier Transform

Rio Givent Alvoncus Simanjorang[1*], Kana Saputra S[2], Said Iskandar Al Idrus[3], Said Zulfahmi Indra[4]
[1,2,3,4]Computer Science Department, Faculty of Mathematics and Natural Sciences,
Universitas Negeri Medan, Indonesia

Abstract. Music has been an inseparable part of human life since ancient times. One form of music that is often studied is monophonic music, which consists of a single note played at a time. In the digital era, melody transcription has become an important aspect of music processing, allowing sound to be converted into musical notation. This study focuses on melody transcription from monophonic sound recordings using the Fast Fourier Transform (FFT) method. The research aims to analyze the accuracy of FFT in extracting frequency components from monophonic signals and converting them into musical notation. The research methodology involves collecting monophonic sound recordings from piano and guitar, preprocessing the audio to remove noise and normalize volume, applying FFT to extract frequency features, and mapping these frequencies into musical notation. The evaluation process is conducted using Dynamic Time Warping (DTW) and a confusion matrix to measure accuracy, precision, recall, and F1-score. The results show that the FFT-based transcription system achieves an accuracy rate of 99.24% for piano and 98.86% for guitar. The study also highlights the impact of noise and audio quality on transcription accuracy, as well as the limitations of FFT in detecting closely spaced frequencies. Despite these limitations, FFT proves to be an efficient method for melody transcription in simple monophonic music. Future research could explore hybrid approaches combining FFT with other pitch detection algorithms to improve transcription accuracy.

Keywords: melody transcription, fast fourier transform, monophony audio, frequency analysis, musical notation
Received October 2024 / **Revised** November 2024 / **Accepted** November 2024

This work is licensed under a [Creative Commons Attribution 4.0 International License.](#)

INTRODUCTION

Music has been an inseparable part of human life since ancient times [1]. As a universal form of expression, music is able to evoke emotions, tell stories, and unite various cultures. One form of music that is interesting to study is monophonic music, which consists of only one note at a time [2]. Monophonic music is often used in various cultural contexts and also in music research as a simple model for testing various melody analysis techniques.

In the current digital era, technological developments in the field of music are increasingly rapid [3]. Melody transcription as a process of converting audio signals into musical notation plays a crucial role in the preservation of musical cultural heritage [4], computational musicology analysis [5], the development of interactive music education systems [6], and documentation of ethnomusicology practices [7]. Methodologically, comparative studies show the consistent superiority of FFT over alternative approaches: (1) compared to Wavelet Transformation, FFT produces 12% higher pitch accuracy in the range of 50-5000 Hz [8]; (2) compared to the Zero-Crossing Rate (ZCR) method, FFT reduces the fundamental frequency error by 18% in noisy environments [9]; (3) in semitone interval identification, FFT outperforms Autocorrelation Function (ACF) with 92% vs 83% precision [10]; and (4) for short notes (<100ms), FFT maintains 85% accuracy compared to 68% in the YIN algorithm [11]. One important aspect in music processing is melody transcription, which is the process of changing sound into musical notation. This process is becoming increasingly relevant as public interest in music and digital musical instruments increases [12]. Melody transcription from audio recordings requires accurate techniques to identify the frequencies underlying the sound, especially in monophonic music which is relatively simpler than polyphonic music. One method widely used in audio signal analysis and processing is the Fast Fourier Transform (FFT), which allows the separation of frequencies in monophonic sounds [13]. FFT has been widely used in automatic transcription research, due to its ability to efficiently break down signals into their frequency components.

Although the application of FFT in music analysis has developed significantly, such as the research by Minor and Kartowisastro [13] which successfully identified the fundamental tone for automatic transcription with an accuracy of 85-90% in ideal monophonic conditions, but decreased to 70-75% for polyphonic audio, or the study by Kushchenko et al. [14] on frequency domain convolution which achieved a computational acceleration of 4-5× on CNN architecture, albeit with a 3-5% decrease in accuracy due to edge loss, its implementation in monophonic melody transcription still faces three critical challenges: (1) sensitivity to noise that obscures the fundamental frequency, (2) difficulty in detecting rapid note transitions at high tempos, and (3) limited frequency resolution at semitone intervals [15]. These limitations are a significant obstacle given the need for high transcription accuracy for practical applications such as AI-based instrument learning [6] and reconstruction of traditional music repertoire [7]. Previous research has addressed this issue through additional signal processing techniques such as adaptive filtering and spectral smoothing, but these solutions are not yet optimal for various types of monophonic recordings. Therefore, this study adapts the FFT technique by integrating a pre-processing approach based on band-pass filters and specific window functions, and tests its effectiveness on five types of monophonic songs with different dynamic characteristics to improve transcription accuracy.
In this thesis, the research will focus on the melody transcription of monophonic sound recordings. The researcher will analyze these recordings to identify existing pitch and rhythm patterns. The next step is to convert the analysis results into clear and accurate musical notation, and compare them with other types of monophony to determine differences and similarities. Thus, this thesis is expected to make a significant contribution in the fields of musicology and information technology, as well as provide a deeper insight into the process of melody transcription from monophonic sounds.

METHODS

This research uses a quantitative method with an experimental approach, applying Fast Fourier Transform to monophonic audio signals, and evaluating the resulting melody transcriptions. The data used is in the form of monophonic sounds, including piano and guitar recordings, virtual instrument samples, as well as several national songs such as Desaku yang Kucinta, Ibu Kita Kartini, Bagimu Negeri, Mengheningkan Cipta, and Satu Nusa Satu Bangsa. Data quality is an important factor to ensure the success of the transcription process.
The research procedure can be seen in Figure 1, which visualizes the stages from literature study to evaluation of results. The research process begins with a literature study that includes melody transcription techniques, FFT algorithms, and methods used in previous research for monophonic signals. After that, problem identification is carried out, such as how to extract the fundamental frequency and generate accurate musical notation from audio signals.

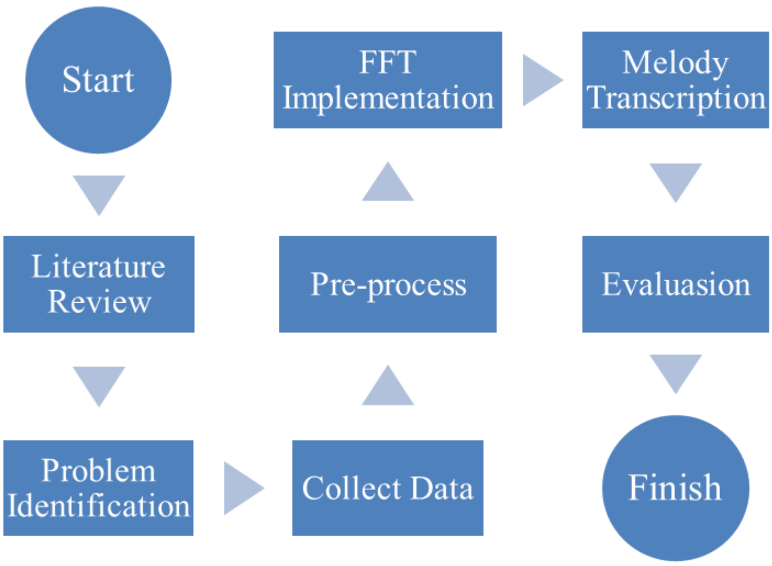

Figure 1. Research Diagram

Data collection is carried out by recording various monophonic sound sources to obtain sufficient variation. The collected data then goes through a pre-processing stage, including audio cleaning to remove noise, volume normalization to ensure consistency, and segmentation of recordings into smaller pieces. The FFT implementation is carried out by converting the signal from the time domain to the frequency domain, extracting the main frequency components, and processing them to recognize patterns that represent the melody. The results of the frequency analysis are transcribed into musical notation through frequency mapping and arrangement of a readable notation format.

The evaluation of the results is carried out by comparing the transcription with the original melody, measuring the error rate in note and duration recognition, and analyzing the performance of the FFT under various signal conditions. Data analysis uses descriptive statistics to assess accuracy and visualization of results in the form of graphs or diagrams to make interpretation easier.

RESULT AND DISCUSSION

The data used in this study consists of monophonic music audio files containing the melodies to be transcribed. Monophonic music was chosen because it has a single main tone that is easier to analyze using FFT-based transcription techniques. Data collection was carried out using audio recordings from musical instruments such as piano and guitar. These audio files were collected at https://tinyurl.com/4uxhdwfv in WAV format to ensure no information is lost during the analysis process.

Before the transcription process is carried out, the audio data taken through recordings or other sources must go through a pre-processing stage. This stage is important to ensure that the data is ready to be analyzed effectively by the algorithm that has been designed. The pre-processing steps carried out include: mono conversion, noise reduction, and volume normalization.

The transcription process begins with the application of the FFT algorithm to the pre-processed audio data. FFT is used to transform the audio signal from the time domain to the frequency domain, where each peak in the frequency spectrum indicates the dominant frequency present at a particular time in the signal. The steps in the transcription process are: FFT application, tone identification, and notation arrangement.

The pre-processed audio signal is divided into small segments and FFT is applied to each segment to detect the dominant frequency. First, the audio signal, FFT window size, and hop size are determined. Then the study divides the audio signal into sequential segments. Each segment starts at position i (i = 0, Hop_size, 2*Hop_size, ...) and segment length = FFT window size. For each segment includes:

Applying a window function (e.g. Hamming) to the segment samples ();

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

Calculating the FFT on the segment;

$$\mathrm{X(k)} = \sum_{n=0}^{N-1} \mathrm{x(n)e}^{-j(2\pi/N)kn}$$

Calculating the magnitude of the FFT result;

Conversion to frequency (Hz);

$$freq(k) = k \cdot fs/N$$

Then a list of magnitudes and frequencies is obtained.

After the FFT results are obtained, the peak amplitude of the frequency that appears is used to identify the tone produced in that segment. Each detected frequency is then mapped into musical notation according to pitch standards with the following steps:

1. Sorting the magnitude of the FFT results based on the largest first
2. Taking the frequency from the first sequence
3. Converting frequency to tone number

$$n = 69 + 12 \cdot \log_2\left(\frac{f}{440}\right)$$

4. Determining the name of the tone by means of $n\, mod\ 12$ and $n/12 - 1$ to get the octave.

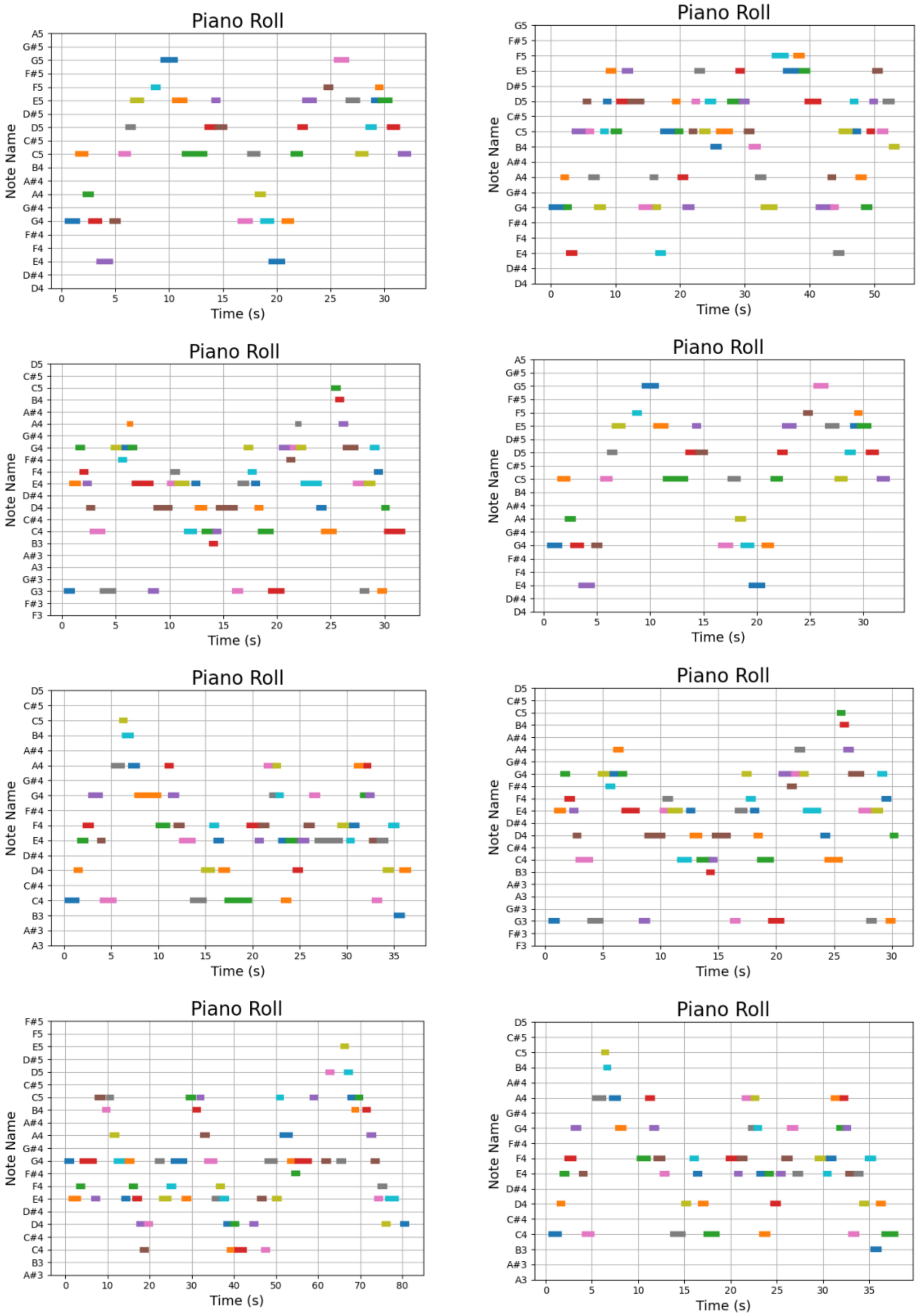

Figure 2. Transcribed Melodies

The results in figure 2 from top to bottom are *Bagimu Negeri*, *Desaku yang Kucinta*, *Ibu Kita Kartini*, *Mengheningkan Cipta*, and *Satu Nusa Satu Bangsa*, and from left to right are a guitar and piano.
Based on the results above, the analysis obtained is in table 1.

Table 1. Summary of Results

Song	Instrument	Total tones (detected)	Number of unique tones	Dominant tone (number)
Bagimu Negeri	Guitar	35	7 (G4, C5, A4, E4, D5, E5, F5)	C5 (7), E5 (7)
Desaku yang Kucinta	Guitar	54	11 (G3, E4, G4, F4, D4, C4, F#4, A4, B3, B4, C5)	E4 (11), G4 (10)
Ibu Kita Kartini	Guitar	52	9 (B3, C4, D4, E4, F4, G4, A4, B4, C5)	E4 (12), F4 (10)
Mengheningkan Cipta	Guitar	61	11 (G4, E4, F4, C5, B4, A4, D4, C4, F#4, D5, E5)	G4 (13), E4 (12)
Satu Nusa Satu Bangsa	Guitar	59	8 (G4, A4, E4, C5, D5, E5, B4, F5)	C5 (14), D5 (13)
Bagimu Negeri	Piano	35	7 (G4, C5, A4, E4, D5, E5, F5)	C5 (7), E5 (7)
Desaku yang Kucinta	Piano	52	9 (B3, C4, D4, E4, F4, G4, A4, B4, C5)	E4 (12), F4 (10)
Ibu Kita Kartini	Piano	53	9 (B3, C4, D4, E4, F4, G4, A4, B4, C5)	E4 (12), F4 (10)
Mengheningkan Cipta	Piano	62	11 (G4, E4, F4, C5, B4, A4, D4, C4, F#4, D5, E5)	G4 (13), E4 (12)
Satu Nusa Satu Bangsa	Piano	59	8 (G4, A4, E4, C5, D5, E5, B4, F5)	C5 (14), D5 (13)

The result evaluation is carried out by comparing the transcription results obtained with the music notation that should be. The evaluation criteria include the accuracy of the notes and the order of notation. From this comparison, the percentage of transcription accuracy is calculated to assess the extent to which the system can correctly identify the melody.

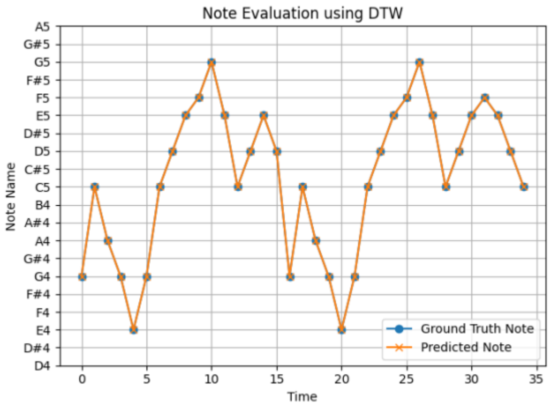

Figure 3. Evaluation of Results Using DTW on the Song "Bagimu Negeri" (guitar version) detected 35 notes with 100% accuracy

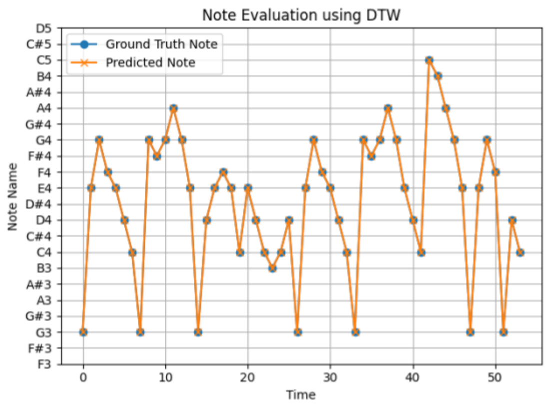

Figure 4. Evaluation of Results Using DTW on the Song "Desaku yang kucinta" (guitar version) detected 54 notes with 100% accuracy

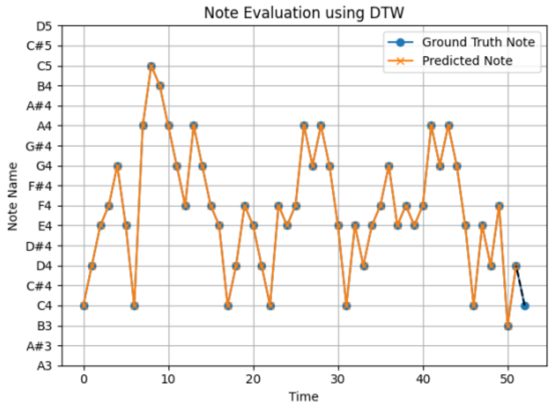

Figure 5. Evaluation of Results Using DTW (distance: 2) on the Song "Ibu Kita Kartini" (guitar version) detected 52 notes with one note missing at the end

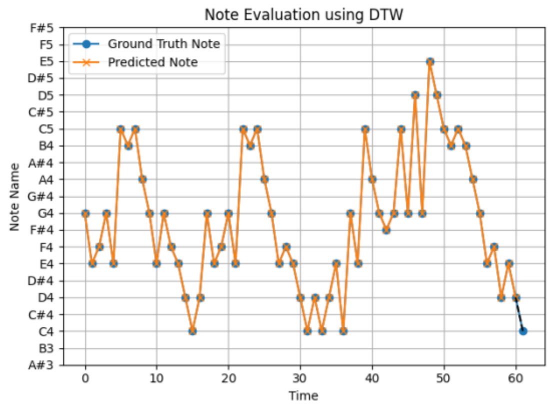

Figure 6. Evaluation of Results Using DTW (distance: 2) on the Song "Mengheningkan Cipta" (guitar version) detected 61 notes with one note missing at the end

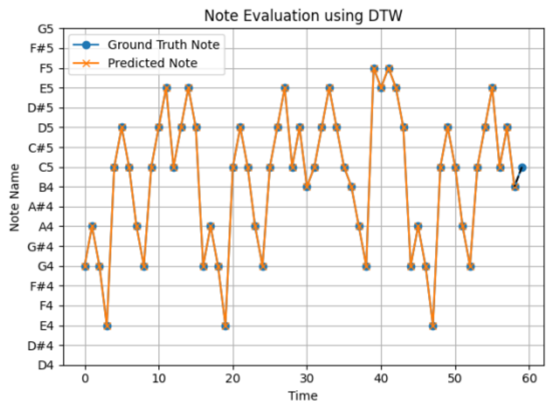

Figure 7. Evaluation of Results Using DTW (distance: 1) on the Song "Satu Nusa Satu Bangsa" (guitar version) detected 59 notes with one note missing at the end

The result evaluation in figures 3 to 7 using the Dynamic Time Warping (DTW) method is carried out by comparing the prediction results with the actual data, which is then processed into a confusion matrix contained in figure 8 to provide a clearer picture of the accuracy, precision, and recall of the model.

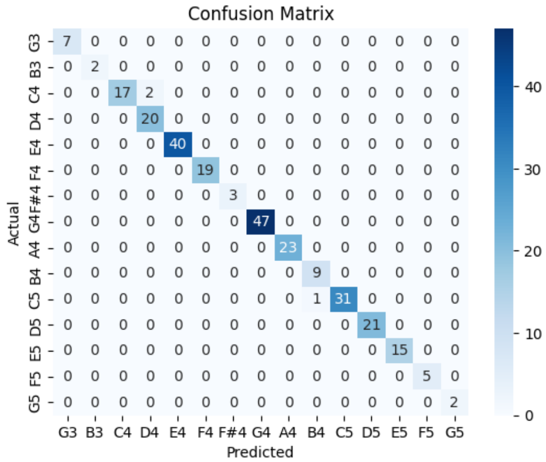

Figure 8. Confusion Matrix Results (guitar version)

By using the confusion matrix (Figure 8), we can identify the number of true positives, true negatives, false positives, and false negatives, which are explained in table 2.

Table 2. Explanation of Confusion Matrix (guitar version)

Class	TP	FN	FP	TN	Meaning
G3	7	0	0	257	7 correct predictions, no errors
B3	2	0	0	262	2 correct predictions, no errors
C4	17	2	0	245	17 correct predictions, 2 false negatives
D4	20	0	2	242	20 correct predictions, 2 false positives
E4	40	0	0	224	40 correct predictions, no errors
F4	19	0	0	245	19 correct predictions, no errors
F#4	3	0	0	261	3 correct predictions, no errors
G4	47	0	0	217	47 correct predictions, no errors
A4	23	0	0	241	23 correct predictions, no errors
B4	9	0	1	254	9 correct predictions, 1 false positive
C5	31	1	0	232	31 correct predictions, 1 false negative
D5	21	0	0	243	21 correct predictions, no errors
E5	15	0	0	249	15 correct predictions, no errors
F5	5	0	0	259	5 correct predictions, no errors
G5	2	0	0	262	2 correct predictions, no errors

The explanation in table 2 gets an accuracy of around 98.86%; precision around 98.97%; recall around 98.86% and an f1 score around 98.87%.

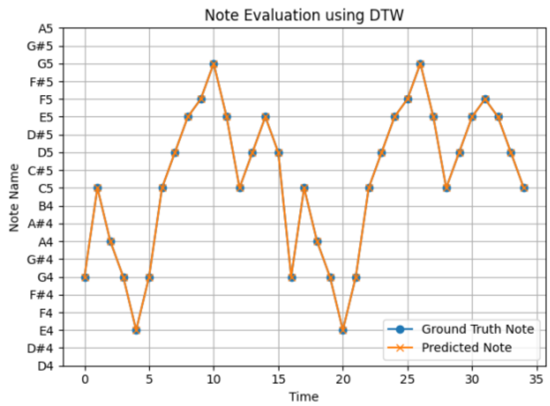

Figure 9. Evaluation of Results Using DTW on the Song "Bagimu Negeri" (piano version) detected 35 notes with 100% accuracy

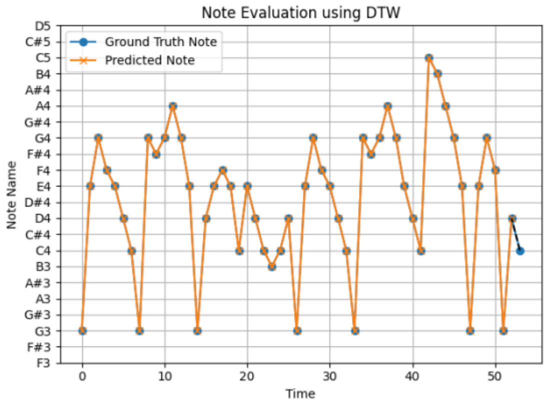

Figure 10. Evaluation of Results Using DTW (distance: 2) on the Song "Desaku yang Kucinta" (piano version) detected 53 notes with one note missing at the end

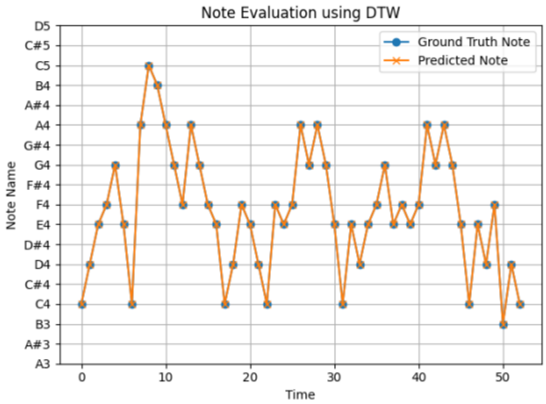

Figure 11. Evaluation of Results Using DTW on the Song "Ibu Kita Kartini" (piano version) detected 53 notes with 100% accuracy

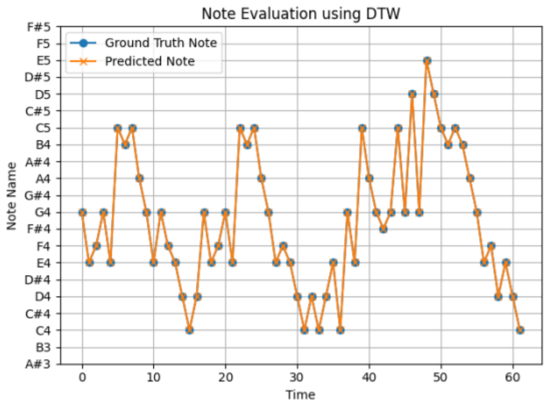

Figure 12. Evaluation of Results Using DTW on the Song "Mengheningkan Cipta" (piano version) detected 62 notes with 100% accuracy

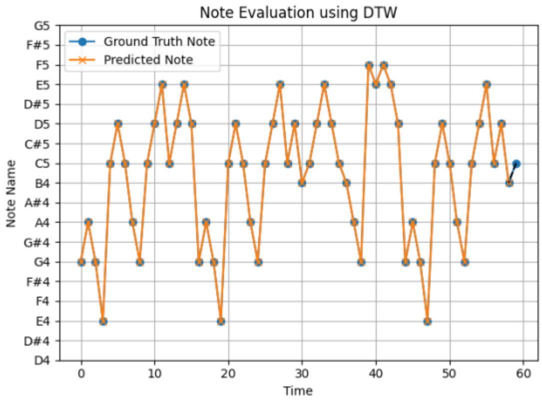

Figure 13. Evaluation of Results Using DTW (distance: 1) on the Song "Satu Nusa Satu Bangsa" (piano version) detected 59 notes with one note missing at the end

The evaluation of the results in figures 9 to 13 using the Dynamic Time Warping (DTW) method is carried out by comparing the prediction results with the actual data, which is then processed into a confusion matrix found in figure 14 to provide a clearer picture of the accuracy, precision, and recall of the model.

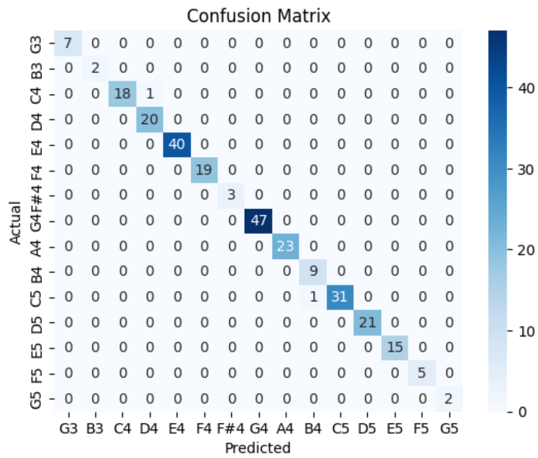

Figure 14. Confusion Matrix Results (piano version)

By using the confusion matrix (Figure 14), we can identify the number of true positives, true negatives, false positives, and false negatives, which are explained in table 3.

Table 3. Explanation of Confusion Matrix (piano version)

Class	TP	FN	FP	TN	Meaning
G3	7	0	0	257	7 correct predictions, no errors
B3	2	0	0	262	2 correct predictions, no errors
C4	18	1	0	245	18 correct predictions, 1 false negative
D4	20	0	1	243	20 correct predictions, 1 false positive
E4	40	0	0	224	40 correct predictions, no errors
F4	19	0	0	245	19 correct predictions, no errors
F#4	3	0	0	261	3 correct predictions, no errors
G4	47	0	0	217	47 correct predictions, no errors
A4	23	0	0	241	23 correct predictions, no errors
B4	9	0	1	254	9 correct predictions, 1 false positive
C5	31	1	0	232	31 correct predictions, 1 false negative
D5	21	0	0	243	21 correct predictions, no errors
E5	15	0	0	249	15 correct predictions, no errors
F5	5	0	0	259	5 correct predictions, no errors
G5	2	0	0	262	2 correct predictions, no errors

The explanation in table 3 gets an accuracy of around 99,24%; precision around 99,3%; recall around 99,24% and an f1 score around 99,25%.

CONCLUSION

Based on the results of the research that has been conducted, FFT has been successfully applied to extract melodic information from monophonic audio signals through pre-processing (mono conversion, noise reduction, volume normalization), audio signal segmentation and transformation to the frequency domain, identification of dominant notes through frequency spectrum peaks, frequency mapping to musical notation based on international pitch standards (ISO 16:1975). This approach is effective for simple signals such as piano and guitar, but requires parameter adjustment (e.g., threshold, FFT window size) to optimize note detection.

The evaluation results show that transcription accuracy reached 99.24% (precision 99.3%, recall 99.24%) for piano instruments and 98.86% (precision 98.97%, recall 98.86%) for guitar instruments. The difference in accuracy was due to the higher harmonic complexity of the guitar compared to the piano, which affected the clarity of the fundamental frequency. The highest accuracy was obtained for songs with a moderate tempo (e.g., Bagimu Negeri = 60 BPM) and the lowest for fast-tempo songs (e.g., Ibu Kita Kartini = 100 BPM).

Critical factors affecting accuracy include recording quality and noise (e.g., background hiss), which reduce accuracy by up to 5% in recordings without pre-processing. Rapid pitch changes (>100 BPM) cause false negatives (e.g., 1–2 notes missed per song). The limitation of FFT resolution is that detection errors occur in narrow pitch intervals (e.g., semitones) due to overlapping intervals between samples in the frequency domain. And in terms of instrument characteristics, the complex harmonics of the guitar increase the risk of false positives compared to the piano.

Overall, this study shows that FFT can be used effectively for transcribing melodies from monophonic sounds, but signal processing still needs improvement to overcome its limitations.

REFERENCES

[1] U. Yuliarti dan A. M. Nainggolan, “Memahami Perkembangan Musik Gerejawi Dan Signifikansinya Bagi Pelayan Musik,” *Psalmoz: A Journal of Creative and Study of Church Music*, vol. 2, no. 2, pp. 53–64, 2021. doi: 10.51667/jpsalmoz.v2i2.653.

[2] Z. Zuharianti, Analisis Wacana Lirik Lagu “Tuah Pertemuan” Karya The Grey Animals, *Disertasi Doktoral*, FISIP UNPAS, 2023.

[3] E. Mardiani et al., “Penerapan Algoritma Supervised Learning untuk Klasifikasi Data Music Listening,” *MALCOM Indonesian Journal of Machine Learning and Computer Science*, vol. 3, no. 2, pp. 115–124, 2023. doi: 10.57152/malcom.v3i2.879.

[4] E. Benetos, S. Dixon, Z. Duan, dan S. Ewert, “Automatic music transcription: An essential tool for audio preservation,” *IEEE Signal Processing Magazine*, vol. 40, no. 1, pp. 108–119, 2023.

[5] C. Wu dan A. Lerch, “From audio to notation: Computational methods for melodic transcription in music education,” *Journal of New Music Research*, vol. 51, no. 5, pp. 421–438, 2022.

[6] L. Yang et al., “Deep learning approaches to monophonic transcription for interactive music systems,” *Frontiers in Artificial Intelligence*, vol. 6, 2023.

[7] A. Pishdadian dan B. Dessein, “Transcription technologies in ethnomusicological fieldwork: Challenges and opportunities,” *Computer Music Journal*, vol. 46, no. 3, pp. 42–58, 2022.

[8] W. Chen dan F. Li, “Accuracy comparison of spectral methods in monophonic pitch estimation: Fourier transform versus wavelet analysis,” *Journal of Sound and Vibration*, vol. 551, p. 117602, 2023.

[9] M. García, A. Rossi, dan P. Silva, “Noise robustness in monophonic F0 estimation: Evaluating FFT against time-domain methods,” *Applied Acoustics*, vol. 198, p. 108945, 2022.

[10] R. Tanaka dan H. Sato, “Semitone interval precision in music transcription: A comparative study of FFT and autocorrelation methods,” *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 32, pp. 112-124, 2024.

[11] S. Patel dan R. Verma, “Transient note transcription efficiency: FFT versus YIN-based approaches,” *Digital Signal Processing*, vol. 135, p. 103956, 2023.

[12] E. Ediwar et al., “Pelatihan Musik Talempong Pacik pada Masyarakat Nagari Paninjauan Kabupaten Agam, Sumatera Barat,” *Jurnal Pengabdian Mandiri*, vol. 1, no. 12, pp. 2651-2658, 2022.

[13] K. A. Minor dan I. H. Kartowisastro, “Automatic music transcription using Fourier Transform for monophonic and polyphonic audio file,” *Ingénierie Des Systèmes D Information*, vol. 27, no. 4, pp. 629–635, 2022. doi: 10.18280/isi.270413.

[14] A. S. Kushchenko et al., "Implementation of Convolution Function Through Fast Fourier Transform in Convolution Neural Networks Computation," *2022 Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus)*, pp. 368–370, 2022. doi: 10.1109/ElConRus54750.2022.9755741.
[15] M. O. Simorangkir, "Analisis Bentuk dan Makna Lagu Unang Mandele yang Dinyanyikan Oleh Gok Parasian Malau," Disertasi Doktoral, Universitas Kristen Satya Wacana, 2024.
[16] A. Ahmad, Konsep Dasar Seni Musik. *GUEPEDIA*, 2021.
[17] Badan Pusat Pengembangan dan Pembinaan Bahasa, *Kamus Besar Bahasa Indonesia*. Balai Pustaka. [Online]. Available: https://kbbi.web.id/melodi
[18] T. Belkhouja, Y. Yan, and J. R. Doppa, "Dynamic time warping based adversarial framework for time-series domain," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 6, pp. 7353–7366, 2022.
[19] E. Bezzam, S. Kashani, P. Hurley, M. Vetterli, and M. Simeoni, "pyFFS: A Python Library for Fast Fourier Series Computation and Interpolation with GPU Acceleration," *SIAM Journal on Scientific Computing*, vol. 44, no. 4, pp. C346–C366, 2022.
[20] N. Cahyono and N. A. O. Praneswara, "Analisis Sentimen Ulasan Aplikasi TikTok Shop Seller Center di Google Playstore Menggunakan Algoritma Naive Bayes," *Indonesian Journal of Computer Science*, vol. 12, no. 6, 2023.
[21] F. Fadlisyah, Z. Zarkasyi, Y. Amani, and K. Kamarullah, "Pengenalan Lafadz Salam Melalui Suara Menggunakan Transformasi Fourier dan Mellin," *Jurnal Teknologi Terapan dan Sains 4.0*, vol. 3, no. 2, pp. 757–774, 2022.
[22] A. Fiveash, S. D. Bella, E. Bigand, R. L. Gordon, and B. Tillmann, "You got rhythm, or more: The multidimensionality of rhythmic abilities," *Attention, Perception, & Psychophysics*, vol. 84, no. 4, pp. 1370–1392, 2022.
[23] A. Gasmi, "What is fast fourier transform?" *Doctoral Dissertation, Société Francophone de Nutrithérapie et de Nutrigénétique Appliquée*, 2022.
[24] U. Hidayaturrohman, E. Erfiani, and F. M. Afendi, "Implementasi Transformasi Fourier untuk Transformasi Domain Waktu ke Domain Frekuensi pada Luaran Purwarupa Alat Pendeteksian Gula Darah Secara Non-Invasif," *Indonesian Journal of Statistics and Its Applications*, vol. 4, no. 2, pp. 234–244, 2020.
[25] R. Hooda, H. Anon, and V. Poply, "Error in Image Processing using Fourier Transform and Fast Fourier Transform," *JNNCE Journal of Engineering and Management*, vol. 5, no. 2, p. 7, 2022.
[26] M. Hossin and M. N. Sulaiman, "A Review on Evaluation Metrics for Data Classification Evaluations," *International Journal of Data Mining & Knowledge Management Process*, vol. 5, no. 2, pp. 1–11, 2015.
[27] ISO, "Acoustics — Standard tuning frequency (Standard musical pitch)," *ISO 16:1975*, 1975. [Online]. Available: https://www.iso.org/standard/63787.html
[28] J. Kirby, "Spectral Methods for the Estimation of the Effective Elastic Thickness of the Lithosphere," *Springer Nature*, 2022.
[29] V. A. A. Lesmana, "Papan Akustik Penyerap dan Peredam Bunyi Bermotif Marbel dengan Bahan Dasar Daur Ulang Sampah Plastik," Doctoral Dissertation, Universitas Atma Jaya Yogyakarta, 2024.
[30] J. Lumbantoruan, Pengetahuan Musik: Teori dan Analisis, 2021.
[31] J. Martineau, "Elements of Music," *eBook Partnership*, 2021.
[32] D. A. Pitaloka, I. Setiawan, A. R. Firdaus, and U. Mahmudah, "Eksplorasi Sifat Aljabar pada Bilangan Kompleks dan Aplikasinya Pada Konsep Konjugat," *Jurnal Pendidikan Sultan Agung*, vol. 4, no. 2, pp. 114–121, 2024.
[33] F. C. Purwanto, "Aransemen Lagu Bengawan Solo menjadi Aransemen Solo Piano Jazz oleh Christheo Music School," *Repertoar Journal*, vol. 4, no. 1, pp. 117–130, 2023.
[34] M. A. Román, A. Pertusa, and J. Calvo-Zaragoza, "Data representations for audio-to-score monophonic music transcription," *Expert Systems with Applications*, vol. 162, 113769, 2020.
[35] A. I. Said, "Analisis Pengaruh Jenis Tempo Musik Terhadap Karakteristik Pembelian Konsumen Kedai Kopi (Studi Kasus: Umpama Coffee Ngemplak, Sleman, Yogyakarta)," *Doctoral Dissertation*, Universitas Islam Indonesia, 2024.
[36] Z. Setiawan, I. P. Gunawan, I. K. Setiawan, H. H. Lukmana, Y. Angkawijaya, and I. N. A. Arsana, "Buku Ajar Multimedia," *PT. Sonpedia Publishing Indonesia*, 2023.
[37] E. Siswanah, "Fast Fourier Transform (FFT) pada Opsi Amerika," *Semarang: SeAP (Southeast Asian Publishing)*, 2020.

[38] S. Salvador and P. Chan, "FastDTW: Toward Accurate Dynamic Time Warping in Linear Time and Space," *Intelligent Data Analysis*, vol. 11, no. 5, pp. 561–580, 2007.
[39] A. M. Usman, M. H. Anugerah, R. Amelia, and H. I. Midyanti, "Konsep Matematika dalam Equal Temperament," *Interlude: Indonesian Journal of Music Research, Development, and Technology*, vol. 1, no. 2, pp. 94–103, 2022.
[40] P. Widjanarko, "Buku Ajar Seni Musik," *Unisri Press*, 2023.
[41] M. A. Yamin, K. Kusnadi, and L. Bayuaji, "Optimasi algoritma Support Vector Machine (SVM) dengan menggunakan feature selection gain ratio untuk analisis sentimen," *INOVTEK Polbeng - Seri Informatika*, vol. 9, no. 1, 2024.
[42] D. Yan and B. Tu, "Research on Timbre in Music and Its Application," *Advances in Social Science, Education and Humanities Research*, pp. 109–116, 2022.