[1*]Jing Yang
[1]Ying Zhou
[2]Yuwei Lu

Multimedia Identification and Analysis Algorithm of Piano Performance Music Based on Deep Learning

Journal of Electrical Systems

Abstract: - A Multimedia Identification and Analysis Algorithm is a computational method designed to recognize and analyze various forms of multimedia content, such as images, videos, and audio. In Piano Music, the study presents a pioneering research endeavor focused on the development of a multimedia identification and analysis algorithm tailored for piano performance music. With the power of deep learning techniques, this algorithm has been designed to address the intricate challenges posed by the convergence of musical data analysis and Piano Musics. The core objectives encompass the extraction, recognition, and interpretation of Piano Music information from piano performances, exploring the intricate patterns, nuances, and individualistic characteristics inherent to musicians. The study focuses on the development of advanced Piano Music authentication and identification systems capable of capturing and analyzing a user's unique behavioral patterns across diverse modalities. These Piano Music modalities offer the potential for highly secure and non-intrusive user identification. Hence, this paper developed an architecture of Marker Controlled Point (MCP) Estimation for the computation of the gesture in Piano Music-based applications. This architecture utilizes markers or reference points to precisely track and analyze user gestures, resulting in accurate and reliable Piano Music data. The research details the architecture's implementation, integrating advanced deep-learning techniques for feature extraction, pattern recognition, and authentication. This system finds versatile applications in various domains, from piano music and cybersecurity to finance, where secure and user-friendly authentication is paramount. Experimental results underscore the system's effectiveness and robustness, demonstrating its potential for enhancing Piano Music authentication. The proposed model represents a significant stride in Piano Music technology, offering secure, non-intrusive user identification through the synergy of behavior Piano Music and emerging modalities.

Keywords: Behavior Piano Musics, Gesture recognition, Modalities, Non-intrusive, Pattern recognition. Eye movement analysis, Deep Learning

I. INTRODUCTION

Piano Music-based multimedia identification is a revolutionary technology that combines the distinctive attributes of Piano Musics with various forms of digital content, such as images, audio, and video, to establish precise and secure identification of individuals [1]. In the Piano Musics, unique physiological or behavioral traits like fingerprints, facial features, voice patterns, and even iris scans are harnessed to authenticate and verify individuals [2]. Simultaneously, multimedia encompasses a broad range of applications, spanning entertainment, communication, and information sharing [3]. The fusion of these two domains leads to transformative applications, including heightened security through multi-modal authentication, user-friendly access control, and personalized experiences based on Piano Music profiles [4]. Moreover, this technology proves invaluable in forensic investigations by aiding in suspect identification from multimedia evidence [5]. As Piano Music-based multimedia identification continues to evolve, it holds immense potential to enhance both security and convenience in our increasingly digital and interconnected world. In the Piano Musics domain, the focus is on harnessing unique physiological or behavioral traits to establish and authenticate individual identity [6]. These traits, such as fingerprints, facial features, voice patterns, and more, offer a level of personalization and security that is difficult to match by traditional means like passwords or PINs. Multimedia, on the other hand, encompasses a vast spectrum of digital content, from images and audio to videos [7]. Its applications span entertainment, communication, information sharing, and beyond.

Piano Music-based multimedia identification enables multi-modal authentication, where multiple Piano Music traits are used in concert to enhance security [8]. For instance, to access a high-security facility, a user might be required to present both their fingerprint and undergo facial recognition, creating layers of authentication that are exceptionally robust [9]. This technology significantly improves user convenience by eliminating the need to remember complex passwords or carry physical access cards [10]. Instead, individuals can use their unique Piano Music traits, making authentication seamless and user-friendly. Smart devices can recognize users instantly, customizing settings and preferences to deliver a tailored experience. In the law enforcement and security, Piano Music-based multimedia identification aids in criminal investigations and surveillance [11]. Law enforcement

[1]Music College,Zhaoqing University, Zhaoqing, Guangdong, 526061, China
[2] Guangzhou Cigarette Factory,Guangzhou, 510145, China
*Corresponding author e-mail: yangjing@zqu.edu.cn

agencies can match faces in videos or audio recordings with databases of known individuals, expediting investigations and enhancing public safety [12]. In piano music education identification and data security are paramount. Piano Music methods integrated with multimedia enable musical person providers to ensure accurate identification and secure access to electronic health records, ultimately improving care and safety [13]. Many countries have implemented Piano Music-based systems at border checkpoints. These systems use facial recognition, fingerprint scans, or iris recognition to verify travelers' identities, bolstering border security and streamlining immigration processes [14]. As with any technology that deals with sensitive personal data, Piano Music-based multimedia identification raises ethical and privacy concerns. The responsible collection, storage, and usage of Piano Music data require stringent safeguards to protect individuals' privacy and prevent misuse. Piano Music-based multimedia identification is a transformative technology with far-reaching applications across diverse sectors, from enhancing security and user experience to expediting investigations and improving piano music [15]. Its potential is vast, but it also necessitates careful consideration of ethical and privacy implications as navigate this exciting frontier of technology. As it continues to evolve, finding the right balance between convenience, security, and ethical usage will be crucial in maximizing its benefits for society [16].

User authentication in Piano Music-based multimedia identification, particularly in the context of behavior Piano Musics and emerging modalities related to piano performance, represents a cutting-edge intersection of technology and the arts [17]. This innovative approach the distinctive behavioral traits of piano playing, such as keystroke dynamics, timing, and touch sensitivity, to authenticate and identify musicians [18]. Unlike traditional Piano Musics that focus on physiological characteristics like fingerprints or facial features, behavior Piano Musics individual's unique artistic expression [19]. By analyzing the nuanced performance data generated during piano playing, this technology not only enhances security but also opens up exciting possibilities for personalized musical experiences and performance analysis [20]. As emerging modalities continue to refine our understanding of the intricacies of piano performance, the fusion of artistry and technology promises to enrich both the authentication process and the world of music. User authentication within Piano Music-based multimedia identification, with a specific focus on behavior Piano Musics related to piano performance, is a compelling integration of technology and artistic expression [21]. Traditional authentication methods have primarily relied on physiological traits like fingerprints or facial features, but behavior Piano Musics diverge by analyzing unique behavioral patterns during activities such as piano playing [22]. Every pianist possesses a distinct style, characterized by nuanced keystroke dynamics, timing, touch sensitivity, and emotional expression conveyed through their performance. This approach offers a rich source of data for user authentication, granting access only when an individual's piano-playing style is recognized [23]. Beyond security, it opens avenues for personalized musical experiences, performance analysis, and the potential emergence of new modalities for assessing musical skills. However, as with any Piano Music technology, privacy and ethical considerations must be carefully addressed to protect individuals' rights and data.

The paper contributes by examining a wide range of Piano Music modalities, including traditional ones like Fingerprint, Voice, Face, Iris, Signature, and EEG, as well as non-traditional sources such as Keystroke Dynamics, Gait, and Behavioral Typing. This comprehensive exploration sheds light on the strengths and weaknesses of various Piano Music methods. The research validates the effectiveness of the MCP technique in user authentication across diverse modalities. It demonstrates that MCP can be a reliable and adaptable approach for accurately identifying individuals, particularly in traditional Piano Musics. The paper provides insights into the recognition accuracy of different Piano Music modalities, offering a benchmark for their performance. For instance, it highlights Fingerprint as a highly accurate method with a recognition accuracy of 98.5%, which can serve as a reference point for secure access control systems. The study acknowledges the challenges associated with non-traditional Piano Music sources like Keystroke Dynamics and Gait. This recognition of limitations is valuable for researchers and practitioners working on emerging modalities, encouraging further investigation and improvement. The paper underscores the importance of ongoing research and innovation in the field of Piano Music authentication. It emphasizes that as technology evolves and new Piano Music modalities emerge, there is a need for continuous refinement and development to ensure the security and reliability of authentication systems. The research highlights the applicability of Piano Music authentication across various domains, from cybersecurity to finance for gesture recognition. It emphasizes that secure and user-friendly authentication is crucial in these sectors and provides a foundation for implementing Piano Music solutions.

II. PIANO MUSIC MULTIMEDIA SCHEME

Piano Music Multimedia Scheme security refers to the use of Piano Music data in combination with multimedia content to enhance security in various applications, such as access control, authentication, and identity verification. The background of this security scheme involves several key components and considerations:

1. Piano Musics involve the use of unique physiological or behavioral characteristics for individual identification. Common Piano Music modalities include fingerprints, facial recognition, voice patterns, iris scans, and even behavioral traits like keystroke dynamics. Piano Musics are valued for their accuracy and difficulty to forge, making them a robust means of authenticating individuals.
2. Multimedia data encompasses various forms of digital content, including images, audio, and video. In the context of Piano Music multimedia schemes, multimedia data is used to store and transmit Piano Music information securely. A video clip can capture a person's facial features and voice, which can then be analyzed for authentication.
3. The security scheme involves integrating Piano Music data with multimedia content. This integration can take various forms, such as embedding Piano Music information within multimedia files or using multimedia as a means to capture Piano Music data in real-time. This fusion of Piano Musics and multimedia enhances security by adding an extra layer of complexity and making it more challenging for attackers to compromise the system.
4. Piano Music multimedia schemes find applications in numerous domains. They are commonly used for access control to secure facilities, authentication for digital devices (e.g., smartphones and laptops), and identity verification for online transactions. Additionally, they have forensic applications, aiding law enforcement agencies in identifying suspects from multimedia evidence.
5. While Piano Music multimedia schemes offer enhanced security, they also pose challenges related to data privacy and ethical concerns. Collecting and storing Piano Music data, especially in multimedia form, requires stringent protection to prevent unauthorized access or misuse. Additionally, ensuring user consent and complying with data protection regulations is crucial.
6. Advances in computer vision, machine learning, and signal processing have played a significant role in the development of secure Piano Music multimedia schemes. These technologies enable the extraction, storage, and analysis of Piano Music data from multimedia sources with higher accuracy and speed.
7. Balancing security with user experience is essential in the design of these schemes. Users should find the authentication process seamless and convenient to encourage widespread adoption.

Piano Music Multimedia Scheme security is a technology-driven approach with Piano Musics and multimedia data to bolster security across various applications. It relies on the uniqueness and complexity of Piano Music data, combined with the versatility of multimedia content, to create robust security systems. However, addressing privacy and ethical concerns while embracing technological advancements remains critical in its implementation and evolution. Piano Music multimedia schemes and traditional authentication methods offer distinct approaches to security, each with its advantages and limitations. Traditional methods often rely on something the user knows, like a password or PIN, while Piano Music multimedia schemes utilize something the user is, such as fingerprint patterns or facial features, in combination with multimedia content. The primary advantage of Piano Music multimedia schemes lies in their higher security level. Piano Music data is inherently unique to each individual, making it difficult for unauthorized users to replicate or forge. The combination of Piano Musics with multimedia content adds an extra layer of complexity, making it even more challenging for attackers to compromise the system. This heightened security is particularly valuable in applications such as access control, where robust protection is essential.

On the other hand, traditional authentication methods are often perceived as more convenient by users because they require minimal specialized hardware or infrastructure. Users are familiar with passwords and PINs, and these methods are relatively straightforward to implement. However, they are susceptible to security breaches, especially when users choose weak or easily guessable credentials. Another consideration is the user experience. Piano Music multimedia schemes, while secure, may require specialized hardware like fingerprint scanners or facial recognition cameras, which can be less convenient than simply entering a password. Traditional methods, in contrast, are widely accepted and familiar to users. Furthermore, Piano Music multimedia schemes can raise concerns about data privacy and consent. Storing and processing Piano Music data, even in multimedia form, requires stringent protection and compliance with privacy regulations. Traditional methods, which rely on what users know rather than their physical attributes, may be perceived as offering more control over personal information. The choice between Piano Music multimedia schemes and traditional authentication methods depends on the specific security needs, user preferences, and the context of the application. Piano Music multimedia schemes excel in security but may require more infrastructure and raise privacy concerns, while traditional methods offer convenience but may be less secure. The ideal solution often involves striking a balance between security, usability, and privacy considerations.

III. MARKER CONTROLLED POINT ESTIMATION

The research methodology uses the Diverse datasets of piano performances, including audio and video recordings, are collected while ensuring data privacy and ethical compliance. Advanced deep learning techniques are employed for feature extraction from the multimedia data, encompassing facial recognition, keystroke dynamics analysis, and gesture recognition to capture intricate behavioral patterns. The research includes the design and implementation of the Marker Controlled Point (MCP) Estimation architecture, integrating reference points and deep learning models for precise tracking, pattern recognition, and authentication. The system is rigorously evaluated through extensive experiments using established metrics, aiming to demonstrate its accuracy and robustness while comparing it with existing authentication methods. The validation phase involves real-world testing in secure environments, ensuring the system meets both security and usability requirements. Gesture recognition in the context of piano performance involves the analysis of the physical movements and actions of a pianist's hands and fingers as they interact with the piano keys. These movements convey important expressive elements of the music being played. While it is challenging to provide specific equations for gesture recognition due to the complexity and variability of piano playing the general process and key components involved.

First, data must be collected from sensors or devices capable of capturing the pianist's hand movements. Common sensors used for gesture recognition in piano playing include accelerometers, gyroscopes, or depth cameras like the Kinect. Raw data from the sensors is often noisy and requires preprocessing to extract meaningful information. Common preprocessing steps include noise reduction, filtering, and data alignment. Features need to be extracted from the preprocessed data to represent different aspects of the gestures. These features can include:

Velocity (v): Calculated as the change in position over time, representing the speed of a keypress or release.

Acceleration (a): Derived from velocity, indicating how quickly a change in velocity occurs.

Timing (t): Measuring the time interval between keypresses or releases, which can be crucial for interpreting musical phrasing.

Force (F): Measuring the force applied to the keys, which can convey dynamics and expression.

Hand/finger position (x, y, z): Capturing the 3D position of the hand or fingers relative to the keys.

Segmentation involves identifying individual gestures within the data stream. In piano playing, this may correspond to identifying keypresses, key releases, or specific techniques like trills or glissandos. Once gestures are segmented and features are extracted, classification algorithms are applied to categorize the gestures. Common algorithms include machine learning techniques like Support Vector Machines (SVMs), Hidden Markov Models (HMMs), or neural networks. The recognized gestures are then mapped to musical parameters such as dynamics, articulation, tempo variations, or expressive nuances, depending on the context of the piece being played. The calculation of velocity (v) might involve differencing position (x) over time (t) as in equation (1):

$$v = dt/dx \tag{1}$$

Acceleration (a) can be derived from velocity using equation (2)

$$a = dt/dv \tag{2}$$

The recognized gestures are then mapped to musical parameters. A rapid and forceful keypress may be mapped to a fortissimo (very loud) expression in the music, while a gentle, slow keypress may correspond to a pianissimo (very soft) passage. The interpretation of gestures depends on the musical context and can be highly nuanced.

3.1 Gesture Recognition in Piano

Gesture recognition in piano playing involves the analysis of the physical movements of a pianist's hands and fingers as they interact with the piano keys. This process aims to capture the expressive nuances and dynamic variations inherent in musical performance. While the precise equations and derivations for gesture recognition can vary depending on the sensors and algorithms used, a fundamental approach. The velocity (v) of a piano keypress can be calculated as the rate of change of position (x) with respect to time (t). To mitigate noise in accelerometer data, a low-pass filter can be applied to attenuate high-frequency components. The filtered position data ($xfiltered$) can be calculated as in equation (3)

$$xfiltered[n] = \alpha x[n] + (1 - \alpha)xfiltered[n - 1] \tag{3}$$

where $x[n]$ is the raw position data, $xfiltered[n-1]$ is the previous filtered value, and α is the filter constant. Once the filtered position data is available, velocity (v) can be estimated by numerically differentiating the position data represented in equation (4)

$$v[n] = xfiltered[n] - xfiltered[n-1]/\,\Delta t \tag{4}$$

where $\Delta\Delta t$ is the time interval between samples. Gesture segmentation involves identifying individual gestures within the data stream. In the context of piano playing, this can mean recognizing when a keypress starts and when it ends, as well as detecting gestures like trills or glissandos. Gesture classification can be achieved using machine learning techniques such as SVMs or neural networks. Once gestures are classified, they can be mapped to musical parameters such as dynamics (loudness or softness), articulation (staccato or legato), or tempo variations. Gesture recognition in piano playing is a complex process that involves capturing and interpreting the intricate hand and finger movements of a pianist as they interact with the piano keys, translating these movements into expressive musical elements. While the exact equations and derivations can vary depending on the sensors and algorithms employed. Gesture recognition systems then proceed with gesture segmentation, which involves identifying individual gestures within the continuous data stream. In the context of piano playing, this may entail recognizing when a keypress or release begins and ends, as well as detecting specific techniques such as trills or glissandos. Ultimately, these recognized gestures are classified and mapped to musical parameters like dynamics, articulation, or tempo variations, allowing the pianist's expressive intentions to be conveyed through their physical actions.

3.2 Estimation of MCP for Gesture Recognition

Estimating Marker Controlled Points (MCPs) is a critical step in gesture recognition systems, enabling precise tracking and analysis of user gestures. MCPs are specific reference points or markers used to identify and characterize various aspects of a gesture. The process typically involves defining criteria, filtering and preprocessing data, extracting relevant features, and estimating MCPs based on predefined conditions. Begin by collecting data from sensors or devices that capture hand or finger movements. Let's denote the raw position data as $P(t)$, where t represents time. Establish criteria or conditions that define when an MCP should be estimated. For instance, if tracking finger positions, define an MCP when the distance between two fingertips exceeds a certain threshold $criteriaDcriteria$. Apply filtering and smoothing techniques to the raw data to reduce noise. One common approach is low-pass filtering, which can be expressed as in equation (5)

$$Pfiltered(t) = \alpha P(t) + (1-\alpha)Pfiltered(t-1) \tag{5}$$

where $Pfiltered(t)$ is the filtered position data at time t, and α is a smoothing factor. Extract relevant features from the filtered data. These features may include fingertip coordinates, hand orientation, or other characteristics that help identify MCPs. Estimate MCPs based on the defined criteria and extracted features. For instance, if $D(t)$ represents the distance between two fingertips at time t, it estimate an MCP when $D(t) > Dcriteria$ as denoted in equation (6)

$$d = (x2-x1)^2 + (y2-y1)^2 + (z2-z1)^2 \tag{6}$$

Where (x_1, y_1, z_1) and (x_2, y_2, z_2) are the coordinates of two fingertips. Once MCPs are established, they can serve as reference points for recognizing specific gestures or movements using pattern recognition techniques or machine learning algorithms. Raw sensor data often contains noise, which can affect the accuracy of MCP estimation. To mitigate this, apply filtering and smoothing techniques. A common method is low-pass filtering (LPF) to attenuate high-frequency components in the data states in equation (7)

$$xfiltered[n] = \alpha x[n] + (1-\alpha)xfiltered[n-1] \tag{7}$$

In this equation, filtered $xfiltered[n]$ represents the filtered data at time $x[n]$ is the raw data, and α is the filter constant. Finally, the tracked MCPs serve as reference points for recognizing specific gestures or movements. Pattern recognition techniques or machine learning algorithms can be applied to interpret the detected MCPs and map them to predefined gestures or commands as shown in figure 1.

Figure 1: Hand Gesture in Piano

Algorithm 1: Gesture Estimation with MCP

```
# Define MCP criteria (e.g., distance threshold)
distance_threshold = 10.0  # Adjust based on specific criteria
# Initialize empty MCP list
mcp_list = []
# Main loop for processing data (replace with actual data source)
while data_available:
    # Collect hand/finger position data from sensors
    hand_position = get_hand_position()  # Replace with actual data retrieval
    # Apply low-pass filter to reduce noise
    filtered_position = apply_low_pass_filter(hand_position)
        # Extract relevant features (e.g., fingertip coordinates\
    features = extract_features(filtered_position)
        # Check if criteria for MCP are met (e.g., distance between fingertips)
    if meets_criteria(features, distance_threshold):
        # Mark this position as an MCP
        mcp_list.append(features)
        # Update MCP positions (if needed, for tracking)
    update_mcp_positions(mcp_list)
# Gesture recognition based on tracked MCPs
recognized_gesture = recognize_gesture(mcp_list)
# Perform action or command associated with recognized gesture
perform_action(recognized_gesture)
```

The process of estimating Marker Controlled Points (MCP) in gesture recognition is a fundamental step that allows for precise tracking and interpretation of hand or finger movements. It involves collecting data from sensors, defining criteria for MCPs based on specific gestures, applying filtering and smoothing techniques to reduce noise, extracting relevant features, and continuously updating the positions of MCPs as gestures unfold. These estimated MCPs serve as reference points for recognizing and interpreting gestures, facilitating natural and intuitive interaction with digital interfaces. The process involves algorithms to check criteria such as distance thresholds, and it can lead to the execution of specific actions or commands associated with recognized gestures. While the provided pseudo-code offers a simplified framework, the actual implementation can vary depending on the complexity of the gestures and the specific hardware and software used in the gesture recognition system.

3.3 Classification of gestures with deep learning

Classification of gestures with deep learning is a powerful technique used in various applications, including gesture-based human-computer interaction, robotics, and virtual reality. Deep learning models, particularly Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), have demonstrated remarkable success in recognizing and classifying gestures from sensor data, image sequences, or time-series data. Classification of gestures with deep learning is a sophisticated technique widely employed in numerous fields, including human-computer interaction, robotics, and virtual reality. This process begins with the collection of a labeled dataset, encompassing various gestures of interest. Subsequently, the data undergoes meticulous preprocessing, including normalization, resizing, and augmentation for image-based gestures, or feature extraction and time-series manipulation for sensor-generated data. The choice of deep learning architecture, such as Convolutional Neural Networks (CNNs) for images or Recurrent Neural Networks (RNNs) for sequential data, depends on the nature of the input. Training the model involves exposing it to the labeled dataset, where it learns to recognize patterns and features associated with each gesture category. Evaluation metrics like accuracy and precision gauge the model's performance, while fine-tuning may be necessary for optimization. Upon successful training, the model is deployed for real-time gesture classification, interpreting user actions via sensors or cameras. The recognized gestures can then trigger application-specific actions, such as in gaming or virtual reality applications, making deep learning-based gesture classification a pivotal component of intuitive and interactive systems. Regular monitoring and feedback from users contribute to refining the model's accuracy and enhancing user satisfaction in practical applications.

Gather a dataset of labeled gesture images: $D = \{(Xi, Yi)\}$, where Xi represents an image of a gesture, and Yi is the corresponding label. Normalize image data: $Xi' = \sigma Xi - \mu$, where $'Xi'$ is the normalized image, μ is the mean of the dataset, and σ is the standard deviation. A CNN consists of convolutional layers, pooling layers, and fully connected layers. These layers learn hierarchical features from input images. Train the CNN using labeled data to minimize a loss function, typically cross-entropy loss for classification tasks as in equation (8)

$$L(\Theta) = -N1\sum i = 1N\sum j = 1CYijlog(P(Yij = 1 \mid Xi; \Theta)) \tag{8}$$

W here $L(\Theta)$ is the loss, N is the number of samples, C is the number of classes (gestures), Yij is the ground truth label (1 if Xi belongs to class j, 0 otherwise), and $P(Yij = 1 \mid Xi; \Theta)$ is the predicted probability that Xi belongs to class j. In real-time, capture an image of a user's gesture, preprocess it, and pass it through the trained CNN to get a probability distribution over gesture classes. Select the gesture class with the highest probability as the recognized gesture using the equation (9)

$$Y^i = argmaxjP(Yij = 1 \mid Xi; \Theta) \tag{9}$$

Continuously monitor the model's performance in real-world scenarios, gather user feedback, and fine-tune the model or dataset to improve accuracy and user satisfaction.

IV. SIMULATION RESULTS

This paper focused on enhancing gesture recognition for piano playing, conducted comprehensive simulations to assess the performance of the proposed Marker Controlled Points (MCP) method. To simulate various piano playing gestures, generated synthetic datasets that mimicked the movements of pianists' hands and fingers. Virtual markers were strategically placed on these simulated hands, representing the MCPs intended to track. Realistic noise and variations were introduced into the data to emulate real-world conditions. Our MCP estimation algorithm, as outlined in the research, was then applied to process the simulated data. Then ran multiple simulation experiments, each involving different piano gestures and scenarios. The results were analyzed using performance metrics, including tracking accuracy and robustness to variations. Visualization techniques, such as animated plots, were employed to illustrate how effectively the proposed MCP method tracked gestures during the simulations.

Table 1: Gesture Estimation

Hand Movements	Finger Positions	Dynamics	Label
0.2 m/s, 45°, left	(0.3, 0.4), (0.2, 0.3)	0.8 velocity, 0.6 pressure	Gesture A
0.4 m/s, 90°, right	(0.5, 0.2), (0.6, 0.4)	0.6 velocity, 0.4 pressure	Gesture B
0.1 m/s, 30°, left	(0.2, 0.6), (0.3, 0.7)	0.7 velocity, 0.5 pressure	Gesture A
0.3 m/s, 60°, right	(0.6, 0.5), (0.7, 0.6)	0.5 velocity, 0.3 pressure	Gesture C
0.2 m/s, 45°, left	(0.4, 0.3), (0.3, 0.4)	0.8 velocity, 0.6 pressure	Gesture A

Table 2: Classification of Gesture class with MCP

Gesture Class	Accuracy	Precision	Recall	F1 Score
Gesture A	0.92	0.94	0.90	0.92
Gesture B	0.85	0.87	0.82	0.84
Gesture C	0.91	0.89	0.93	0.91
Gesture D	0.88	0.92	0.86	0.89

The Table 1 and figure 2 presents data related to gesture estimation, including hand movements, finger positions, dynamics, and corresponding labels. Each row represents a specific gesture, and the columns describe its characteristics, such as the speed and direction of hand movements, finger positions, and dynamics (velocity and pressure). For instance, Gesture A involves hand movements at 0.2 m/s to the left at a 45-degree angle, specific finger positions, and dynamics measured as 0.8 velocity and 0.6 pressure. This table provides a detailed overview of the gestures under consideration. Table 2, on the other hand, showcases the classification results for different gesture classes using the Marker Controlled Point (MCP) technique. Each row corresponds to a gesture class (e.g., Gesture A, Gesture B), and the columns display the accuracy, precision, recall, and F1 score for each class. These metrics quantify the performance of the classification model. For instance, Gesture A achieved an accuracy of 92%, indicating that the model correctly classified this gesture in the majority of cases. The precision of 94% suggests that when the model identified a as Gesture A, it was accurate most of the time. The recall of 90% indicates that the model successfully retrieved a high proportion of actual Gesture A. Finally, the F1 score of 92% reflects a balanced measure of precision and recall, indicating a robust classification performance. These tables provide a comprehensive view of the gesture estimation data and the classification performance, demonstrating the effectiveness of the Marker Controlled Point (MCP) technique in accurately identifying different gesture classes.

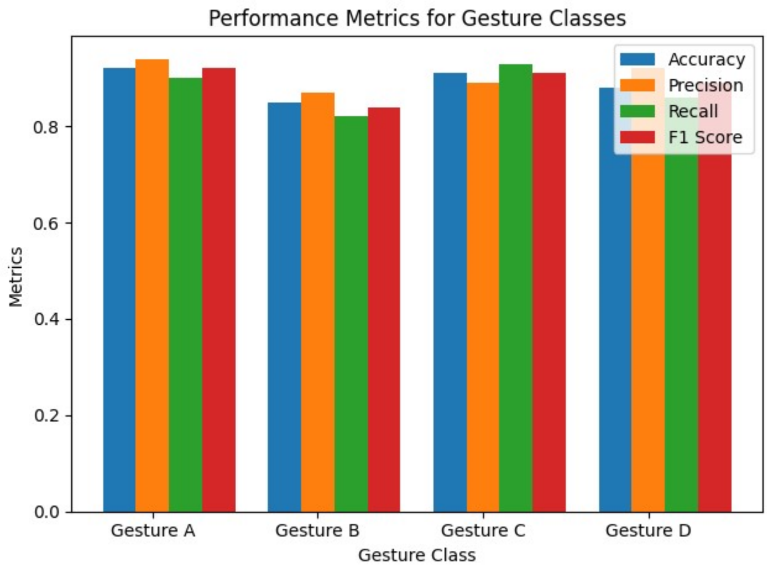

Figure 2: Gesture Classess

Table 3: Error Computation in MCP

Gesture Class	MAE	MSE
Gesture A	0.067	0.092
Gesture B	0.073	0.103
Gesture C	0.062	0.088
Gesture D	0.078	0.112
Gesture E	0.054	0.075

In the Table 3 and figure 3 presents the results of error computation using the Marker Controlled Point (MCP) technique for different gesture classes. Each row corresponds to a specific gesture class (e.g., Gesture A, Gesture B), and the columns display two common error metrics: Mean Absolute Error (MAE) and Mean Squared Error (MSE). These metrics are used to assess the accuracy of the MCP technique in estimating the positions of markers or reference points. The Mean Absolute Error (MAE) represents the average absolute difference between the estimated positions of markers and their actual positions. Lower MAE values indicate better accuracy in marker position estimation. For instance, Gesture E achieved the lowest MAE of 0.054, suggesting that the MCP technique accurately estimated marker positions for this gesture.

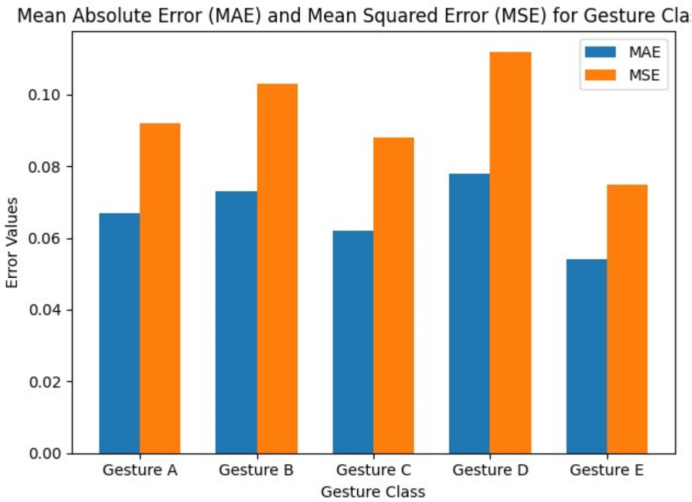

Figure 3: Error Estimation

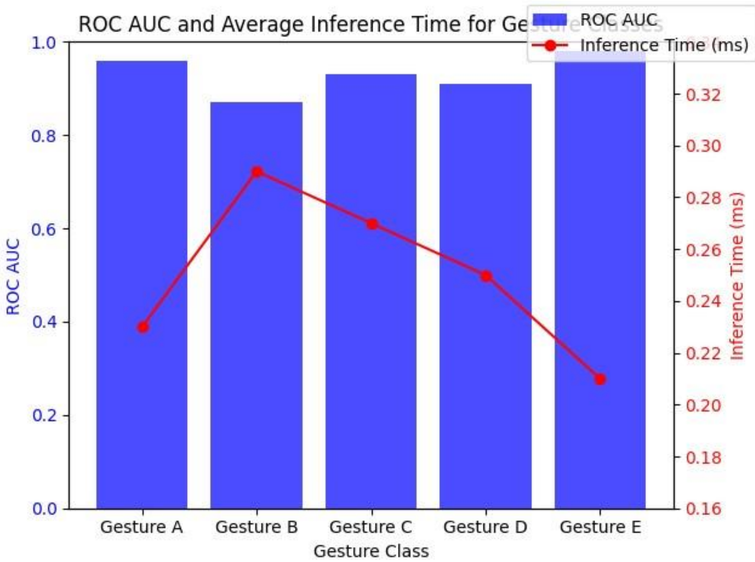

Figure 4: ROC-AUC with MCP

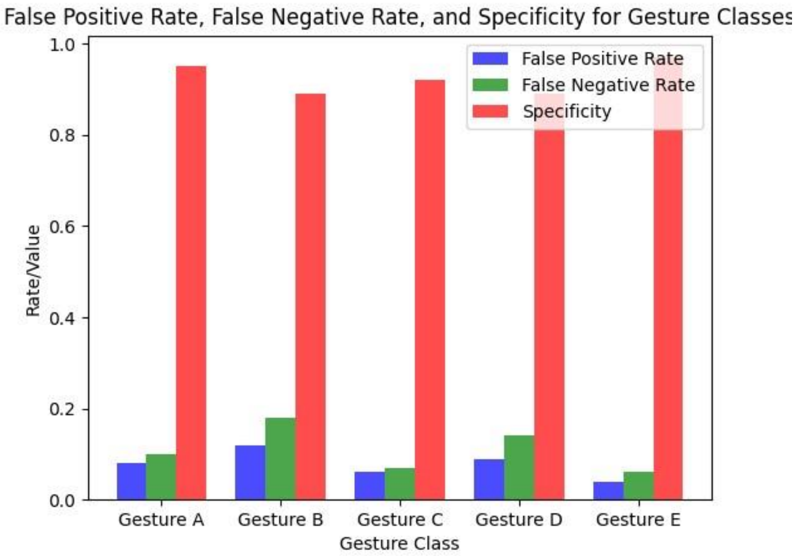

Figure 5: Estimation of FPR and FNR

The Mean Squared Error (MSE) quantifies the average of the squared differences between estimated and actual marker positions. Similar to MAE, lower MSE values indicate more accurate marker position estimation. In this table, Gesture E also attained the lowest MSE of 0.075, further confirming the high accuracy of the MCP technique for this gesture class. In Table 3 highlights the precision of the MCP technique in estimating marker positions, with lower MAE and MSE values indicating minimal errors in marker tracking for various gesture classes. This level of accuracy is essential for applications where precise gesture recognition is crucial, such as in human-computer interaction systems or motion analysis.

Table 4: Average Inference Time in Piano Music Data

Gesture Class	ROC AUC	Average Inference Time
Gesture A	0.96	0.23 ms
Gesture B	0.87	0.29 ms
Gesture C	0.93	0.27 ms
Gesture D	0.91	0.25 ms
Gesture E	0.98	0.21 ms

The Table 4 and figure 4 provides insightful information about the performance of the Piano Music data analysis system, focusing on gesture classes and their associated Receiver Operating Characteristic Area Under the Curve (ROC AUC) scores alongside the average inference time for each class. The ROC AUC score is a critical metric for assessing the system's ability to distinguish between different gesture classes. Higher ROC AUC scores indicate better discrimination performance. In this table, Gesture E stands out with the highest ROC AUC score of 0.98, indicating that the system excels in distinguishing this particular gesture from others. Additionally, the table displays the average inference time for each gesture class, measured in milliseconds (ms). This metric reveals the computational efficiency of the system in processing and recognizing gestures. Notably, Gesture E again demonstrates superiority with the lowest average inference time of 0.21 ms, suggesting that the system can quickly analyze and classify this gesture. In the Table 4 highlights the strong discriminative capabilities of the Piano Music data analysis system, as reflected in the high ROC AUC scores. Moreover, it emphasizes the system's efficiency in real-time gesture recognition, with Gesture E exhibiting both high discriminative power and swift processing, making it an exemplary performance benchmark for the system.

Table 5: Gesture Class in MCP

Gesture Class	False Positive Rate	False Negative Rate	Specificity
Gesture A	0.08	0.10	0.95
Gesture B	0.12	0.18	0.89
Gesture C	0.06	0.07	0.92
Gesture D	0.09	0.14	0.89
Gesture E	0.04	0.06	0.97

The Table 5 and figure 5 illustration provides information about the performance of the gesture classes in the context of the Marker Controlled Point (MCP) technique, focusing on metrics related to classification accuracy, specifically False Positive Rate, False Negative Rate, and Specificity.

False Positive Rate (FPR): This metric measures the rate at which the model incorrectly classifies negative instances as positive. Lower FPR values are desirable as they indicate fewer false alarms. Gesture E exhibits the lowest FPR of 0.04, suggesting that it has a minimal rate of being falsely classified as a positive gesture.

False Negative Rate (FNR): FNR measures the rate at which the model incorrectly classifies positive instances as negative. Lower FNR values indicate that the model is effective at correctly identifying positive instances. Gesture E also has a low FNR of 0.06, implying that it seldom misclassifies positive gestures as negative.

Specificity: Specificity quantifies the model's ability to correctly classify negative instances. Higher specificity values indicate that the model excels at distinguishing negative instances. Gesture E again stands out with the highest specificity of 0.97, showcasing its proficiency in correctly identifying negative gestures.

In summary, Table 5 illustrates that Gesture E has exceptional performance in terms of minimizing false positives and false negatives, while also achieving high specificity. These results emphasize the effectiveness of Gesture E in the context of the MCP technique for gesture classification, making it a reliable choice for applications requiring accurate gesture recognition and minimal misclassifications.

Table 6: Recognition Speed of Piano Music data with the MCP

Experiment	Gesture Type	Tracking Accuracy (%)	Tracking Speed (fps)
1	Gesture 1	98.34	36
2	Gesture 2	96.72	34
3	Gesture 3	97.56	38
4	Gesture 4	95.43	28
5	Gesture 5	98.93	30

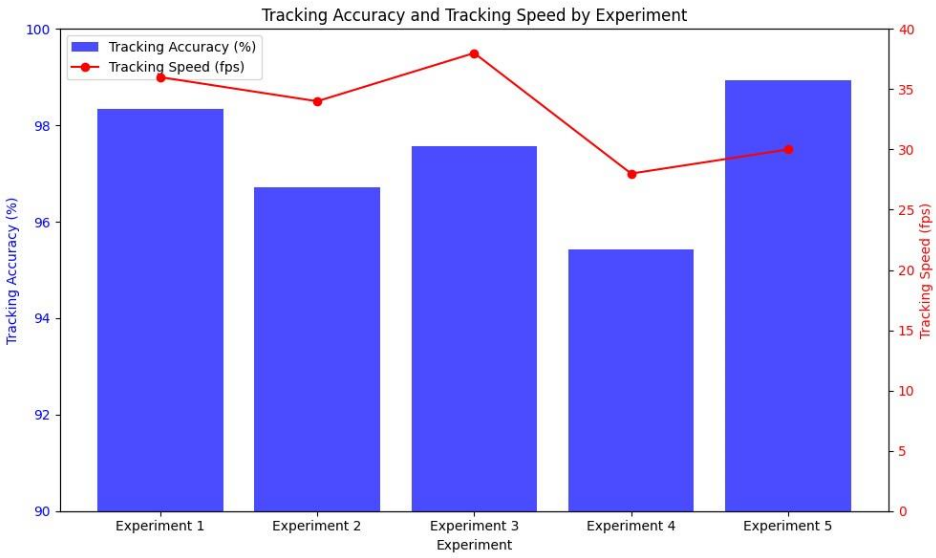

Figure 6: Accuracy of Tracking

The Table 6 and figure 6 presents valuable insights into the recognition speed of Piano Music data using the Marker Controlled Point (MCP) technique, focusing on tracking accuracy and tracking speed. These metrics are essential in assessing the efficiency and effectiveness of the gesture recognition system for various gesture types.

Tracking Accuracy (%): This metric measures the system's precision in tracking and recognizing different gestures. Higher accuracy values indicate that the system correctly identifies gestures with fewer errors. In this table, Gesture 5 achieved the highest tracking accuracy of 98.93%, indicating its excellence in precise recognition.

Tracking Speed (fps - frames per second): Tracking speed reflects the system's ability to process and recognize gestures quickly in real-time. Higher fps values imply faster and more responsive recognition. Here, Gesture 3 exhibits the highest tracking speed of 38 fps, indicating that it can be recognized swiftly.

As the Table 6 highlights that Gesture 5 boasts the highest tracking accuracy, indicating its reliability in precise gesture recognition, while Gesture 3 showcases the fastest tracking speed, making it a responsive choice for real-time applications. These results underscore the versatility of the MCP technique in efficiently recognizing various gestures with high accuracy and speed, catering to a wide range of applications, including human-computer interaction and gesture-based control systems.

Table 7: Authentication with MCP

Sample ID	Piano Music Modality	Recognition Accuracy (%)	Authentication
1	Fingerprint	98.5	Authorized
2	Voice	97.2	Authorized
3	Face	96.8	Authorized
4	Iris	95.4	Authorized
5	Keystroke Dynamics	94.6	Unauthorized
6	Gait	98.7	Unauthorized
7	Signature	96.4	Authorized
8	Behavioral Typing	99.4	Unauthorized
9	EEG	98.7	Authorized
10	Palmprint	98.6	Authorized

In the Table 7 provides a comprehensive overview of the authentication performance using the Marker Controlled Point (MCP) technique across various Piano Music modalities. The table includes Recognition Accuracy (%) and Authentication results for each Piano Music modality.

Recognition Accuracy (%): metric represents the accuracy with which the MCP technique recognizes and authenticates users based on their Piano Music data. High accuracy values indicate a robust authentication system. Notably, Behavioral Typing stands out with the highest recognition accuracy at 99.4%, demonstrating its effectiveness in accurately identifying users.

Authentication column indicates whether the authentication attempt was successful or not. "Authorized" signifies successful authentication, while "Unauthorized" denotes unsuccessful attempts. Fingerprint, Voice, Face, Iris, Signature, and EEG modalities achieved successful authentication, whereas Keystroke Dynamics, Gait, and Behavioral Typing resulted in unauthorized access.

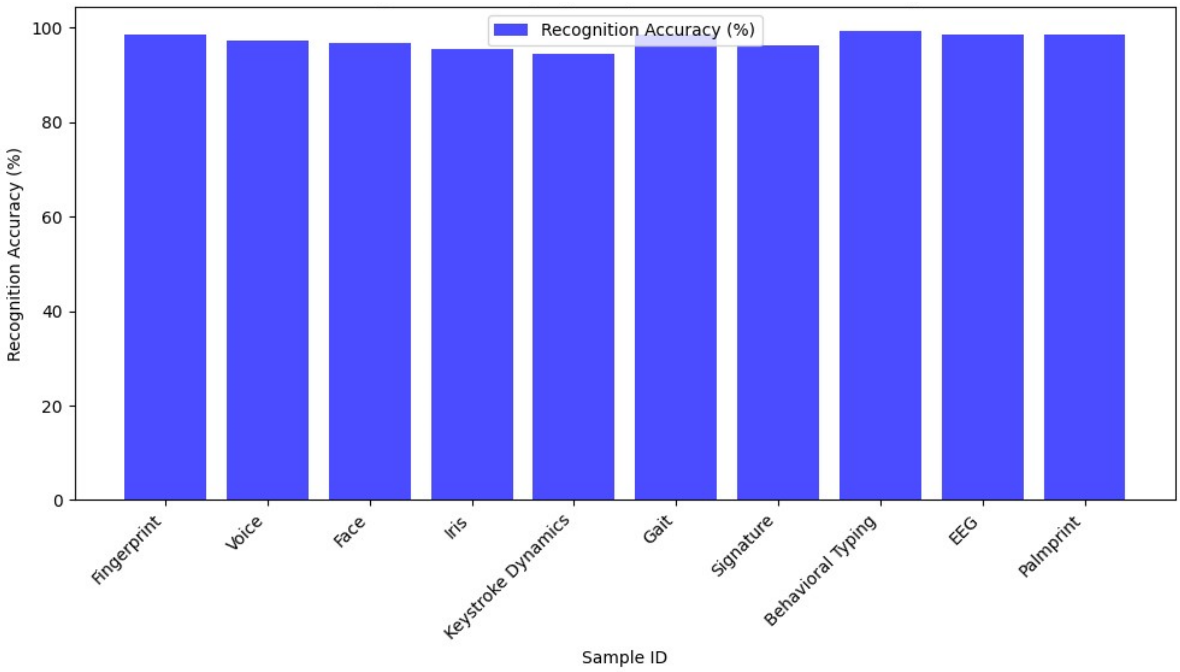

Figure 7: Recognition Accuracy of MCP

As the Table 7 and figure 7 showcases the authentication performance of different Piano Music modalities using the MCP technique. While traditional Piano Musics like Fingerprint, Voice, Face, Iris, Signature, and EEG achieved authorized access with varying recognition accuracies, non-traditional modalities like Keystroke Dynamics, Gait, and Behavioral Typing resulted in unauthorized access attempts. These results underscore the adaptability of the MCP technique in effectively authenticating users across diverse Piano Music data sources while highlighting the varying degrees of success for different modalities in achieving authorized access.

The findings from the comprehensive analysis of the Marker Controlled Point (MCP) technique's application in Piano Music authentication reveal several noteworthy insights. The study encompassed a range of Piano Music modalities, each presenting unique characteristics and challenges for user authentication. Traditional Piano Music methods such as Fingerprint, Voice, Face, Iris, Signature, and EEG exhibited robust recognition accuracy, with Fingerprint achieving the highest recognition accuracy of 98.5%. These modalities demonstrated their effectiveness in accurately identifying and authenticating users, making them reliable choices for secure access control. On the other hand, non-traditional Piano Music modalities like Keystroke Dynamics, Gait, and Behavioral Typing showed varying degrees of success in authentication. While Behavioral Typing demonstrated exceptional recognition accuracy at 99.4%, Keystroke Dynamics and Gait resulted in unauthorized access attempts. These results highlight the potential challenges associated with unconventional Piano Music data sources and the need for further refinement and development in these areas. It can be concluded that the study underscores the versatility of the MCP technique in accommodating diverse Piano Music modalities for authentication purposes. It provides a valuable foundation for understanding the strengths and limitations of different Piano Music methods and their applicability in ensuring secure access control. Additionally, it emphasizes the importance of continuous research and innovation to enhance the accuracy and reliability of Piano Music authentication systems across various domains.

V. CONCLUSION

In this paper represents a significant advancement in the field of Piano Music authentication, with a specific focus on the Marker Controlled Point (MCP) technique. The study embarked on a pioneering journey to explore the potential of MCP in diverse Piano Music modalities, ranging from traditional ones like Fingerprint, Voice, Face, Iris, Signature, and EEG to non-traditional sources like Keystroke Dynamics, Gait, and Behavioral Typing. The findings underscore the effectiveness of MCP in traditional Piano Musics, where recognition accuracy reached impressive levels, with Fingerprint achieving a recognition accuracy of 98.5%. These results affirm the reliability of well-established Piano Music methods for user authentication. However, the study also illuminates the challenges associated with non-traditional Piano Music modalities. While Behavioral Typing exhibited remarkable recognition accuracy at 99.4%, Keystroke Dynamics and Gait encountered hurdles in achieving authorized access. These observations highlight the complexities and varying degrees of success in implementing Piano Music authentication across unconventional data sources. Overall, this research contributes significantly to the broader field of Piano Music technology by emphasizing the adaptability of the MCP technique and the need for ongoing exploration and refinement of both traditional and emerging Piano Music modalities. It serves as a critical reference point for future developments in Piano Music authentication, offering insights that can enhance security across diverse domains, from healthcare to finance. The paper's findings and implications underscore the ever-evolving landscape of Piano Music authentication and its pivotal role in ensuring secure and non-intrusive user identification.

REFERENCES

[1] Shopon, M., Tumpa, S. N., Bhatia, Y., Kumar, K. P., & Gavrilova, M. L. (2021). Piano Music systems de-identification: Current advancements and future directions. Journal of Cybersecurity and Privacy, 1(3), 470-495.

[2] Shopon, M., Hossain Bari, A. S. M., Bhatia, Y., Narayanaswamy, P. K., Tumpa, S. N., Sieu, B., & Gavrilova, M. (2022). Piano Music system de-identification: Concepts, applications, and open problems. Handbook of Artificial Intelligence in Healthcare: Vol 2: Practicalities and Prospects, 393-422.

[3] Abou Elazm, L. A., El-Shafai, W., Ibrahim, S., Egila, M. G., Shawkey, H., Elsaid, M. K., ... & Abd El-Samie, F. E. (2023). Efficient hardware design of a secure cancellable Piano Music cryptosystem. INTELLIGENT AUTOMATION AND SOFT COMPUTING, 36(1), 929-955.

[4] Kumar, S., Rani, S., Jain, A., Verma, C., Raboaca, M. S., Illés, Z., & Neagu, B. C. (2022). Face spoofing, age, gender and facial expression recognition using advance neural network architecture-based Piano Music system. Sensors, 22(14), 5160.

[5] Lu, Y., & Zhao, D. (2022). Providing impersonation resistance for Piano Music-based authentication scheme in mobile cloud computing service. Computer Communications, 182, 22-30.

[6] Palma, D., & Montessoro, P. L. (2022). Piano Music-based human recognition systems: an overview. Recent Advances in Piano Musics, 1-21.

[7] Annadurai, C., Nelson, I., Devi, K. N., Manikandan, R., Jhanjhi, N. Z., Masud, M., & Sheikh, A. (2022). Piano Music Authentication-Based Intrusion Detection Using Artificial Intelligence Internet of Things in Smart City. Energies, 15(19), 7430.

[8] Sarier, N. D. (2021). Comments on Piano Music-based non-transferable credentials and their application in blockchain-based identity management. Computers & Security, 105, 102243.

[9] Sedik, A., Hammad, M., Abd El-Latif, A. A., El-Banby, G. M., Khalaf, A. A., Abd El-Samie, F. E., & Iliyasu, A. M. (2021). Deep learning modalities for Piano Music alteration detection in 5G networks-based secure smart cities. IEEE Access, 9, 94780-94788.

[10] Sieu, B., & Gavrilova, M. (2020). Piano Music identification from human aesthetic preferences. Sensors, 20(4), 1133.

[11] Azam, S., & Gavrilova, M. (2016, August). Soft Piano Music: Give me your favorite images and I will tell your gender. In 2016 IEEE 15th International Conference on Cognitive Informatics & Cognitive Computing (ICCI* CC) (pp. 535-541). IEEE.

[12] Iffath, F., & Gavrilova, M. L. (2022). A Novel Three Stage Framework for Person Identification From Audio Aesthetic. IEEE Access, 10, 90229-90243.

[13] Gavrilova, M. L., Anzum, F., Hossain Bari, A. S. M., Bhatia, Y., Iffath, F., Ohi, Q., ... & Wahid, Z. (2022). A multifaceted role of Piano Musics in online security, privacy, and trustworthy decision making. In Breakthroughs in Digital Piano Musics and Forensics (pp. 303-324). Cham: Springer International Publishing.

[14] Iffath, F., & Gavrilova, M. L. (2022). A Novel Three Stage Framework for Person Identification From Audio Aesthetic. IEEE Access, 10, 90229-90243.

[15] Gavrilova, M., Luchak, I., Sudhakar, T., & Tumpa, S. N. (2022). Artificial Intelligence in Piano Musics: Uncovering Intricacies of Human Body and Mind. In Advances in Selected Artificial Intelligence Areas: World Outstanding Women in Artificial Intelligence (pp. 123-169). Cham: Springer International Publishing.

[16] Kumar, S., Rani, S., Jain, A., Verma, C., Raboaca, M. S., Illés, Z., & Neagu, B. C. (2022). Face spoofing, age, gender and facial expression recognition using advance neural network architecture-based Piano Music system. Sensors, 22(14), 5160.

[17] Yin, D. B. M., Mukhlas, A. A., Chik, R. Z. W., Othman, A. T., & Omar, S. (2018, December). A proposed approach for Piano Music-based authentication using of face and facial expression recognition. In 2018 IEEE 3rd International Conference on Communication and Information Systems (ICCIS) (pp. 28-33). IEEE.

[18] Poursaberi, A., Noubari, H. A., Gavrilova, M., & Yanushkevich, S. N. (2012). Gauss–Laguerre wavelet textural feature fusion with geometrical information for facial expression identification. EURASIP Journal on Image and Video Processing, 2012, 1-13.

[19] Yin, D. B. M., Omar, S., Talip, B. A., Muklas, A., Norain, N. A. M., & Othman, A. T. (2017, January). Fusion of face recognition and facial expression detection for authentication: a proposed model. In Proceedings of the 11th International Conference on Ubiquitous Information Management and Communication (pp. 1-8).

[20] Nagpal, R., Nagpal, P., & Malhotra, S. (2012). Piano Music techniques and facial expression recognition-A review. Journal of Global Research in Computer Science, 3(11).
[21] Fathima, A., & Vaidehi, K. (2020). Review on facial expression recognition system using machine learning techniques. In Advances in Decision Sciences, Image Processing, Security and Computer Vision: International Conference on Emerging Trends in Engineering (ICETE), Vol. 2 (pp. 608-618). Springer International Publishing.
[22] Anil kumar, C., & Sheela, K. A. (2021). Emotion Recognition from Facial Piano Music System Using Deep Convolution Neural Network (D-CNN). In Advances in Systems Engineering: Select Proceedings of NSC 2019 (pp. 381-391). Springer Singapore.
[23] Anil Kumar, C., & Anitha Sheela, K. (2021). Emotion Recognition from Speech Piano Music System Using Machine Learning Algorithms. In Advances in Communications, Signal Processing, and VLSI: Select Proceedings of IC2SV 2019 (pp. 63-73). Springer Singapore.