Early Diagnosis of Atrial Fibrillation Episodes: Comparative Analysis of Different Matrix Architectures

Naseha Wafa Qammar 1,4, Alfonsas Vainoras 2, Zenonas Navickas 1, Gediminas Jaruševičius 2 and Minvydas Ragulskis 1

1 Department of Mathematical Modelling, Kaunas University of Technology, 51368 Kaunas, Lithuania
2 Institute of Cardiology, Lithuanian University of Health Sciences, 44307 Kaunas, Lithuania
* Correspondence: naseha.qammar@ktu.edu

Abstract: This study presents three different matrix architectures for the analysis of ECG parameters, aimed at detecting atrial fibrillation episodes. The evaluation involves a cohort of 15 individuals, utilizing these matrix architectures across various orders. The findings reveal that the matrix norm delivers significantly better results compared to the large discriminant of the matrix. Detailed analysis of the spatial expansion of each matrix structure indicates that the PMLD architecture excels in terms of expandability compared to the MA1 and MA2 matrices. Consequently, third- and fifth-order PMLD matrix architectures are employed for classification techniques, demonstrating enhanced sensitivity with increased matrix order. These results are validated through the classification of several test candidates, confirming the efficacy of the proposed method. The study suggests that the developed approach holds substantial potential for clinical diagnostics in the early detection of atrial fibrillation.

Keywords: atrial fibrillation episodes; three different matrix architectures; sensitivity; expandability

1. Introduction

Atrial fibrillation (AF) is a prevalent cardiac arrhythmia characterized by irregular and frequently accelerated heartbeats. It increases the risk of cardiovascular mortality, severe cardiovascular events, heart failure, ischemic heart disease, sudden cardiac death, and stroke [1,2]. Patients often suffer symptoms with AF such as exhaustion, palpitations, shortness of breath, and chest discomfort. In some instances, individuals also experience these with asymptomatic or silent AF, meaning that an individual may not exhibit any visible symptoms related to AF. Thus, it is crucial to mention that silent AF or undiagnosed AF can cause life-threatening thromboembolic consequences due to its intermittent nature [3,4]. The combination of rhythm irregularities in the AF makes the diagnosis and detection difficult and thus requires efficient and effective methods and tools for its early detection [2,4,5].

Various heart health monitoring instruments are available on the market for the detection of AF, and one such tool is AliveCor [2]. However, it is important to mention that this equipment is distinguished by a significant expense, making it unattainable for most individuals. Moreover, the achievement of reliable outcomes with AliveCor relies on the precise and prompt identification of AF symptoms. Other methods for diagnosing AF include the use of single-lead electrocardiogram (ECG) monitors, Holter monitoring, mobile telemetry monitoring, and implanted loop recorders [2]. These heart health monitoring methods may be classified as either expensive or may necessitate the use of specialized methods that must be used in conjunction with substantial signal processing algorithms to provide high-quality and reliable results.
The 12-lead electrocardiogram (ECG) is often regarded as the most dependable technique for identifying a fibrillation waveform [6,7]. When diagnosing atrial fibrillation (AF) using the electrocardiogram (ECG), physicians look for abnormalities in the pattern of ECG waveforms. This includes irregularities in the RR interval and the absence of P waves, which are replaced by irregular fibrillation waves (f waves) that vary in shape and size. There is a huge body of research that studies atrial fibrillation symptoms for early detection. It has been shown in one of the studies that hospitalized patients with atrial fibrillation (AF) were older and had more health issues, such as diabetes and heart disease, compared to outpatients [8]. Additionally, another study showed that patients with type 2 diabetes and hyperuricemia are 9.66 times more likely to have AF, as analyzed using a generalized linear model of the Poisson family [9]. AF is a growing health concern influenced by aging and cardiovascular risk factors, progressing from paroxysmal (episodes lasting <7 days) to persistent (≥7 days) or permanent forms, with yearly progression rates varying widely [10]. Hohnloser et al.’s study found Apixaban promising for AF patients up to 140 kg, but careful monitoring is needed for those at extreme weights [11]. Another study revealed that 3.9% of lung cancer patients in outpatient oncology clinics had recently diagnosed AF, with high rates of major bleeding and cardiovascular events significantly impacting mortality in early-stage cancer patients [12]. Diastolic dysfunction, often overlooked in coronary heart disease, can lead to AF through increased atrial afterload, stretching, and pressure, causing left atrial enlargement and triggering the condition [13]. Finally, coronary embolism, a rare cause of acute STEMI, can be triggered by conditions like AF [14]. These studies and many others show that the detection of AF is indeed important as it may cause many diseases and can be fatal too if not treated on time. The distinctive electrocardiogram (ECG) pattern assists clinicians in distinguishing atrial fibrillation (AF) from abnormal sinus rhythm and other types of cardiac arrhythmias [2,15,16].

The electrocardiogram (ECG) signals exhibit a complex behavior because of a range of physiological and pathological factors [17–19]. The analysis of the ECG signal is a complex task due to its intricate nature. Nevertheless, researchers are always involved in a proactive endeavor to develop intelligent, efficient, and effective methods for the purpose of clinical diagnosis [20]. Machine learning techniques have been extensively employed for the early detection of cardiac diseases. The utilization of deep learning algorithms is a prevalent practice in various domains, often employed for the purpose of executing tasks such as feature extraction and dimensionality reduction to extract pertinent information for clinical diagnosis and monitoring [21,22]. In [23], a real-time deep learning model has combined the CNN and RNN for AF detection in long-term ECG recordings and has also achieved high sensitivity. Additionally, methods such as reconstructed phase space analysis [24], Lyapunov exponents [25,26], detrended fluctuation analysis (DFA) [30], recurrence plots and Poincaré plots [31], among others, have been extensively researched to analyze the ECG signal [20].

Ziaukas et al. presented a noteworthy study that showcased a proficient mathematical methodology aimed at comprehending the complex relationship between RR interval and JT wave during a bicycle ergometry exercise [18]. Their research findings unveiled that the Perfect Matrices of Lagrange Differences (PMLD) is an effective matrix-based approach to examine the ECG signal and the self-organization of the cardiovascular system during the load and recovery phases of the stress test [18]. Qammar et al. also employed PMLD methodology to examine the RR interval and JT wave to gain insights into the mechanisms underlying the regulation of arterial blood pressure during the stress test [17]. Qammar et al., in another study, analyzed and proved that the PMLD can be employed for the early diagnosis of atrial fibrillation episodes [20]. It is worth mentioning that the utilization of matrix-based techniques has demonstrated efficacy even when working with datasets of limited size [20].
This paper draws inspiration from the works of Ziaukas et al. and Qammar et al., who have convincingly showcased the efficacy of a matrix-based approach for the analysis of ECG signals [17,18,20]. The main aims of this study are as follows:

1. To present and analyze three different architectures of matrices for the early detection of atrial fibrillation episodes. The first two architectures are presented by and adapted from the work of Navickas et al. [32]. In this paper, those matrices are labelled as Matrix Architecture One (MA1) and Matrix Architecture Two (MA2), respectively. It is important to mention that (MA1) and (MA2) are being used for the first time for the analysis of atrial fibrillation dataset. The third architecture is sourced from the work of the Qammar et al. [20], and is referred to as Perfect Matrices of Lagrange Differences (PMLDs).

2. It is noteworthy to mention and acknowledge that the groundwork for exploring the application of the PMLD for atrial fibrillation episodes has already been laid by Qammar et al. [20]. However, ref. [20] focuses on utilizing only three cardiac parameters of the recorded ECG—JT wave, QRS complex, and RR interval [20]. Building upon this foundation, the current paper aims to extend the application of the concept of PMLD matrices to employ a broader array of cardiac parameters. Thus, the multifold objectives of the study are summarized as follows:

   i. To conduct a thorough analysis of the three matrix configurations to determine if the matrix norm or large discriminant yields favorable results for the analysis of ECG signals when utilizing only three cardiac parameters: JT wave, QRS complex, and RR interval.

   ii. Once the most suitable matrix architecture is identified, it will be subjected to additional analysis to evaluate its expandability. The term “expandability” within the context of the matrix architecture denotes its capacity for transformation into higher orders by integrating additional cardiac parameters.

   iii. The expandability analysis will aid in identifying the specific matrix architecture for further analysis.

   iv. Both the matrix sensitivity and variability are then evaluated from the second to the fifth order for the identification and classification of the atrial fibrillation episodes.

   v. Finally, the classification is performed, and the assessment of test applicants is then carried out, which will bring practical implications for furthering the clinical diagnosis.

2. Materials and Methods

   The description of the experimental setup:

   Five cardiac intervals are captured by employing the Kaunas–Load system for fifteen participants during resting ECG: the duration of the JT wave, the duration of the QRS complex, the RR inter-beat interval, the amplitude of the P wave represented as AP, and the duration of the P wave, denoted as DP. The overall workflow diagram is shown in Figure 1.
2.1. Participants

The research cohort is divided into two groups: healthy and unhealthy individuals. The category of healthy people includes those who have not been registered as having atrial fibrillation episodes. On the other hand, individuals within the unhealthy category do have a history of atrial fibrillation events (though they are not undergoing medical treatment). In other words, the typical lengths of the measured cardiac intervals and the characteristic shape of the ECG signal do not change. The healthy cohort consists of eight individuals, whereas the unhealthy cohort consists of six. The average age of the entire cohort is $65.84 \pm 1.4$ years, with a height of $1.75 \pm 0.12$ m, and a weight $79.4 \pm 0.9$ kg respectively. The healthy cohort is labelled as $H_n$ whereas the unhealthy cohort is labelled as $U_n$ where $n$ represents the number of individuals in the corresponding group.

2.2. Ethics Statement

This research met all applicable standards for the ethics of experimentation in accordance with the Declaration of Helsinki as reflected in the prior approval by the Regional Biomedical Research Ethics Committee of the Lithuanian University of Health Sciences (ID No. BE-2-4, 130, 17 March 2016). Participants also provided written informed consent prior to the experiment.

2.3. Description of Matrix Architecture

The basic idea of the matrix-based analysis of the ECG signal originates from the works of Ziaukas et al. and others [17,18,20,33]. For example, refs. [10,11,13] focused on analyzing the algebraic relationships between the ECG parameters for better understanding the complex self-organization of the human cardiovascular system during the load and recovery phases of the stress test. Ref. [24] focused on the development of a decision support system for the early detection of atrial fibrillation episodes. The matrix architectures employed in this study are as follows: (1) Matrix Architecture One (MA1); (2) Matrix Architecture Two (MA2); and (3) Perfect Matrices of Lagrange Differences (PMLD). The following section explains each of the matrix architecture in detail.
2.4. The Architecture of PMLD Matrices

As mentioned previously, the architecture of PMLD matrices were introduced in [18]. This architecture is based on six requirements:

1. Each element within the matrix is distinct.
2. Zero-order differences are positioned along the main diagonal.
3. First-order differences are positioned along the secondary diagonal.
4. The indices \( x \) and \( y \) can assume one of three possible values as explained in [18].
5. The ideal matrix of Lagrange differences maintains lexicographical balance, ensuring that the number of symbols \( x \) and \( y \) in the expressions of all matrix elements is equal.
6. The ideal matrix of Lagrange differences maintains temporal balance, ensuring that the number of indices with subscripts minus delta and plus delta in the expressions of all matrix elements is equal.

There are only eighteen different matrices satisfying these six requirements [18]. Without loss of generality, we will consider PMLD matrix #1 in this paper. Given two synchronously recorded time series \((x_n; n = 0,1,2,3 \ldots)\) and \((y_n; n = 0,1,2,3 \ldots)\), PMLD matrix #1 reads [18]:

\[
PMLD_n = \begin{bmatrix}
 x_n & y_{n+\delta} - x_{n+\delta} \\
 y_{n-\delta} - x_{n-\delta} & y_n \\
 z_{n-\delta} - x_{n-\delta} & z_n - y_{n-\delta} \\
 z_{n+\delta} - x_{n+\delta} & z_{n+\delta} - y_{n+\delta}
\end{bmatrix}
\]

where the index \( n \) represents the current moment and \( \delta \) is used to represent the time lag \((\delta \in \mathbb{N})\). In Equation (1), the elements on the principal diagonal represent the zero-order derivatives. The elements on the secondary diagonal represent the differences between the elements of time series \( x_n \) and \( y_n \). Those differences could be interpreted as first-order cross derivatives.

It has also been demonstrated that PMLD matrices can be extended to higher orders [20]. For example, given three synchronously recorded time series \((x_n; n = 0,1,2,3 \ldots)\), \((y_n; n = 0,1,2,3 \ldots)\), and \((z_n; n = 0,1,2,3 \ldots)\), the third-order (PMLD) matrix #1 reads [20]:

\[
PMLD_n = \begin{bmatrix}
 x_n & y_{n+\delta} - x_{n+\delta} & z_{n+\delta} - x_{n+\delta} \\
 y_{n-\delta} - x_{n-\delta} & y_n & z_n - y_{n-\delta} \\
 z_{n-\delta} - x_{n-\delta} & z_{n-\delta} - y_{n-\delta} & z_n \\
 z_{n+\delta} - x_{n+\delta} & z_{n+\delta} - y_{n+\delta} & z_{n+\delta} - y_{n+\delta}
\end{bmatrix}
\]

Note that this extension to the third order matrix keeps all six requirements in force. In other words, the extended third-order matrix in Equation (2) is also a PMLD matrix.

2.5. The Architecture of MA1 and MA2 Matrices

Navickas et al. [32] proposed two matrix architectures for analyzing the relationships between two time series. Given two time series \((x_n; n = 0,1,2,3 \ldots)\) and \((y_n; n = 0,1,2,3 \ldots)\) the second-order MA1 reads:

\[
A_n = \begin{bmatrix}
 2x_n & x_{n+1} - y_{n+1} \\
 x_{n-1} - y_{n-1} & 2y_n
\end{bmatrix}
\]

From Equation (3), it is evident that all six requirements raised for PMLD are also satisfied for MA1. However, the only difference between Equation (3) and Equation (1) is that the elements on the main diagonal are multiplied by a scalar factor. As detailed in reference [13], the presence of cross derivatives inherently amplifies the subtle variations in the investigated signals. One of the hypotheses which will be investigated in this paper is based on the assumption that the multiplication of the zero-order elements by a scalar factor greater than one does decrease the ability to detect those subtle variations in the investigated signals.

It is clear that the structure of MA1 also allows a straightforward expansion of the dimension of the matrix. For example, three time series \((x_n; n = 0,1,2,3 \ldots)\), \((y_n; n = 0,1,2,3 \ldots)\), and \((z_n; n = 0,1,2,3 \ldots)\) would be mapped into the MA1 architecture in the following way:
2.7. The Degree of Freedom in the Matrix Architecture

The degree of freedom in the matrix architecture encompasses the flexibility to arrange the matrix elements in various configurations while adhering to the six requirements raised for the PMLD matrix. Those requirements ensure that the ordering of elements in the matrix maintains a consistent pattern, which is crucial for the integrity of the local matrix structure. For example, when considering a second-order matrix derived from two synchronously recorded time series (\(x_n\): \(n = 0,1,2,3 \ldots\) and \(y_n\): \(n = 0,1,2,3 \ldots\)), there are four different ways to arrange all the elements of the matrix while preserving the requirements. Each arrangement represents a unique structure of the matrix:

\[
A_1 = \begin{bmatrix}
\frac{x_n}{x_{n-1}} - y_{n-1} & x_{n+1} - y_{n+1} \\
\end{bmatrix}
\]  

(6)
\[ \begin{align*}
A_2 &= \begin{bmatrix}
  x_n & y_{n+1} - x_{n+1} \\
  y_{n-1} - x_{n-1} & y_n
\end{bmatrix} \\
A_3 &= \begin{bmatrix}
  x_n & y_{n+1} - y_{n+1} \\
  y_{n-1} - x_{n-1} & y_n
\end{bmatrix} \\
A_4 &= \begin{bmatrix}
  x_n & y_{n+1} - x_{n+1} \\
  x_{n-1} - y_{n-1} & y_n
\end{bmatrix}
\end{align*} \]

While the MA1 and PMLD matrices allow for this degree of freedom due to their lexicographical balance, the same degree of freedom is not achievable for MA2. Also, in the research conducted by Qammar et al., the specific arrangement depicted in Equation (7) is employed for analytical purposes. In accordance with this precedent, this paper also adopts the same arrangement for subsequent analysis.

2.8. The Transformation of the Sequence of the Matrices into the Scalar Time Series

The algorithm used to identify the algebraic relationship between two cardiac parameters uses two basic steps [18]. Firstly, two scalar time series (representing two cardiac parameters) are transformed into a sequence of matrices. Then, this sequence of matrices is transformed into a scalar time series by using the mapping \( F: \mathbb{R}^{(m \times n)} \rightarrow \mathbb{R}^{(1 \times n)} \), where \( m \) is the order of the matrix, and \( n \) is the length of the two original sequences.

Various mapping functions \( F \) can be employed for this purpose. For instance, the norm of the matrix is utilized to analyze the algebraic relationships between ECG parameters in [18]. Similarly, the discriminant of the matrix is used to convert the sequence of matrices into the scalar time series in [17].

In this study, two different mapping functions are utilized. The large discriminant of the matrix is used as the mapping function for MA1 and MA2 matrices, whereas the norm of the matrix is used as the mapping function for the transformation of the PMLD matrix into the scalar time series as it is also adopted in [20].

2.9. The Large Discriminant of the Matrix

Let us represent the elements of the matrix MA1 and MA2 as:

\[ MA_{a(n)} = \begin{bmatrix}
  a_{11}^{(n)} & a_{12}^{(n)} & a_{13}^{(n)} \\
  a_{21}^{(n)} & a_{22}^{(n)} & a_{23}^{(n)} \\
  a_{31}^{(n)} & a_{32}^{(n)} & a_{33}^{(n)}
\end{bmatrix} \]

Then, the large discriminant of the matrix is expressed as \( dsk_1(MA_{a}^{(n)}) = \rho_{12} * \rho_{13} \)

where \[ \rho_{12} = \left[ a - \sqrt{(a^2 - 3b)} \right] \left[ a + 2\sqrt{(a^2 - 3b)} \right] - 27c \] \( \rho_{13} = \left[ a + \sqrt{(a^2 - 3b)} \right] \left[ a - 2\sqrt{(a^2 - 3b)} \right] - 27c \), \( a = \text{Inv}_1(MA_{a}(n)) = a_{11}^{(n)} + a_{22}^{(n)} + a_{33}^{(n)} \), \( b = \text{Inv}_2(MA_{a}(n)) = a_{22}^{(n)} \cdot a_{33}^{(n)} - a_{11}^{(n)} \cdot a_{23}^{(n)} - a_{12}^{(n)} \cdot a_{33}^{(n)} + a_{13}^{(n)} \cdot a_{22}^{(n)} - a_{13}^{(n)} \cdot a_{21}^{(n)} - a_{12}^{(n)} \cdot a_{31}^{(n)} - a_{12}^{(n)} \cdot a_{32}^{(n)} + a_{11}^{(n)} \cdot a_{23}^{(n)} - a_{11}^{(n)} \cdot a_{32}^{(n)} \) [32,34].

2.10. The Norm of the Matrix

In [20], the norm of the matrix is used as the mapping function to transform the PMLD into a scalar time series. The same mapping function is also utilized in this paper. Detailed discussion about the norm of matrix is already provided in [18,20].

3. Results and Discussion

The computational analysis of the study is organized into the following segments: (1) the comparison between the matrix norm and the matrix discriminant, (2) matrix expandability analysis and the selection of the best matrix architecture, and (3) the classification of the matrix architecture for different orders. Each of these segments will be discussed in further detail below.
3.1. The Comparison between the Matrix Norm and the Matrix Discriminant

The three cardiac parameters (the JT wave, the duration of QRS complex and RR interval) represented by time series \( (x_n; n = 0,1,2,3...) \), \( (y_n; n = 0,1,2,3...) \), and \( (z_n; n = 0,1,2,3...) \), are preprocessed. A threshold range is applied to ensure that each input value remains within physiologically plausible limits. For instance, the minimum and maximum limits are chosen for each of the time series: \( x_{(\text{min})} = 100 \text{ ms} \); \( x_{(\text{max})} = 400 \text{ ms} \); \( y_{(\text{min})} = 800 \text{ ms} \); \( y_{(\text{max})} = 110 \text{ ms} \); \( z_{(\text{min})} = 600 \text{ ms} \); and \( z_{(\text{max})} = 1200 \text{ ms} \), respectively. Any values falling below the lower bounds or exceeding the upper bounds are adjusted to the respective minimum or maximum values. The adjusted values of the time series \( x_n \), \( y_n \) and \( z_n \) are then normalized to a range between 0 and 1. For instance, the time series \( x_n \) is normalized as \( x_{\text{normalized}} = (x - 100)/(400 - 100) \). Similarly, the time series \( y_n \) and \( z_n \) are processed through the same normalization technique, respectively. The same normalization techniques are also used in [32].

Once the input parameters are preprocessed and normalized, then, without losing the generality, the preprocessed and normalized time series \( x_n \), \( y_n \) and \( z_n \) are fitted to the three matrix architectures as represented by Equation (2), Equation (4) and Equation (5). Moreover, it is crucial to acknowledge that the degree of freedom allows for the choice of any variant specified in Equations (6)–(9). Thus, for this research, the variant mentioned in Equation (7) is chosen due to the reasons explained in [20]. To transform the third-order matrix into a scalar series via the mapping function \( F; \mathbb{R}^{3 \times 3} \rightarrow \mathbb{R}^{1 \times 3} \), twofold mapping techniques are employed. The large discriminant of the matrix is used as the mapping function for MA1 and MA2 architectures. The norm is used as the mapping function for the PMLD architecture following the criteria described in [20]. The internal and external smoothing techniques are applied with the smoothing radius adjusted to the values optimized in [18]. Finally, a statistical metric, such as the variance of the norm or the variance of the large discriminant, is calculated for each produced time sequence from the three original time series \( x_n \), \( y_n \) and \( z_n \). The resulting values are depicted in Tables 1 and 2.

**Table 1.** The comparison of the variance values for different matrix architectures for healthy candidates.

<table>
<thead>
<tr>
<th>List of Healthy Candidates, ( H_n )</th>
<th>Matrix Architecture One (MA1)</th>
<th>Matrix Architecture Two (MA1)</th>
<th>The Architecture of PMLD Matrices</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Variance of the Large Discriminant</td>
<td>Variance of the Norm</td>
<td>Variance of the Large Discriminant</td>
</tr>
<tr>
<td>( H_1 )</td>
<td>0.0640</td>
<td>0.0016</td>
<td>0.042</td>
</tr>
<tr>
<td>( H_2 )</td>
<td>0.0633</td>
<td>0.0047</td>
<td>0.0313</td>
</tr>
<tr>
<td>( H_3 )</td>
<td>182.4470</td>
<td>0.0027</td>
<td>16.167</td>
</tr>
<tr>
<td>( H_4 )</td>
<td>44.6925</td>
<td>0.0014</td>
<td>1.6402</td>
</tr>
<tr>
<td>( H_5 )</td>
<td>608.0258</td>
<td>0.0076</td>
<td>94.7592</td>
</tr>
<tr>
<td>( H_6 )</td>
<td>12.4591</td>
<td>0.0026</td>
<td>2.9064</td>
</tr>
<tr>
<td>( H_7 )</td>
<td>1618.7723</td>
<td>0.0042</td>
<td>1490.7099</td>
</tr>
<tr>
<td>( H_8 )</td>
<td>62.0865</td>
<td>0.0042</td>
<td>65.921</td>
</tr>
</tbody>
</table>

**Table 2.** The comparison of the variance values for different matrix architectures for unhealthy candidates.

<table>
<thead>
<tr>
<th>List of Unhealthy Candidates, ( U_n )</th>
<th>Matrix Architecture One (MA1)</th>
<th>Matrix Architecture Two (MA1)</th>
<th>The Architecture of PMLD Matrices</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Variance of the Large Discriminant</td>
<td>Variance of the Norm</td>
<td>Variance of the Large Discriminant</td>
</tr>
<tr>
<td>( U_1 )</td>
<td>1144.7393</td>
<td>0.0012</td>
<td>21.2530</td>
</tr>
</tbody>
</table>
In Tables 1 and 2, the variance values for both the large discriminant and the norm of the matrix are presented for the healthy and unhealthy cohort, respectively. It is pertinent to note that while the variance of the norm provides meaningful analytical insights into the matrix architectures, the variance of the large discriminant appears to yield misleading information without a clear analytical significance applicable for classification purposes. It can be observed that the variance of the large discriminant exhibited a wide range of values, with minimum and maximum ranges varying significantly across both healthy and unhealthy cohorts. For instance, in the healthy cohort, the variance of the large discriminant ranges from 0.042 to 1618.7723, whereas in the unhealthy cohort this variance ranged from 0.0006 to 1358.7248. In contrast, the variance of the norm demonstrated more consistency and interpretable trends, with narrower minimum and maximum ranges. For example, the variance of the norm ranges from 0.0010 to 0.0216 in the healthy cohort. The variance of the norm ranges from 0.0004 to 0.0177 in the unhealthy cohort. This observation underscores the importance of employing the norm of the matrices as a useful and relevant statistical metric for further computations for the distribution and classification (seen in Figure 2).

### 3.2. Matrix Expandability Analysis and the Selection of the Best Matrix Architecture

Once it has been decided that the variance of the norm will be used as the metric of choice, further investigation indicates a challenge in selecting the optimum matrix architecture among PMLD, MA1 and MA2. This concern emerges from the closeness in variance values produced by all the architectures for both healthy and unhealthy candidates. For example, when examining the healthy candidate $H_1$, it is observed that the variance of the norm for each architecture are rather close to each other: MA1 = 0.0016, MA2 = 0.0010, and PMLD = 0.0012, respectively. This trend persists consistently across all candidates for both healthy and unhealthy participants.

Thus, a selection criterion is needed to discover the best matrix architecture. The first criterion for the selection of matrix architecture revolves around the expandability of a matrix by focusing on the six requirements [18]. These requirements facilitate the integration of additional data or cardiac parameters into the matrix structure. As mentioned previously, both the MA1 and PMLD architectures can be naturally expanded due to their inherent properties. However, the MA2 architecture fails to fulfill the requirements due to the loss of balance among its elements, rendering it unsuitable for further analysis and leading to its exclusion from the selection criteria.

Moreover, upon a closer examination of the MA1 architecture, it can also be observed that the principle diagonal elements are multiplied by a scalar. As previously discussed, the multiplication of the zero-order elements by a scalar factor greater than one does decrease its ability to detect the subtle variations in the investigated signals [18]. Consequently, the MA1 architecture is also excluded from the selection criteria.

Thus, based upon the selection criteria, the PMLD architecture emerges as the most reliable choice for the subsequent computations, expandability analysis, and the development of the classification techniques. Its balanced structure, coupled with its ability to accommodate new data, positions the PMLD architecture as the optimal matrix architecture for the analysis of complex signals like ECG.

Table 3 provides results for the norm of the PMLD matrix for the healthy candidates from the second to the fifth orders. A general trend is observed in both the sensitivity and the variability of the produced data. As defined already, sensitivity in the context of this

<table>
<thead>
<tr>
<th></th>
<th>U_2</th>
<th>U_3</th>
<th>U_4</th>
<th>U_5</th>
<th>U_6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1358.7248</td>
<td>15.2796</td>
<td>339.2649</td>
<td>4.8991</td>
<td>84.2605</td>
</tr>
<tr>
<td></td>
<td>0.0006</td>
<td>0.0171</td>
<td>0.0066</td>
<td>0.0023</td>
<td>0.0013</td>
</tr>
<tr>
<td></td>
<td>21.9133</td>
<td>2.5576</td>
<td>1566.6759</td>
<td>0.9349</td>
<td>3.0256</td>
</tr>
<tr>
<td></td>
<td>0.0005</td>
<td>0.0177</td>
<td>0.0111</td>
<td>0.0030</td>
<td>0.0022</td>
</tr>
<tr>
<td></td>
<td>0.0004</td>
<td>0.0082</td>
<td>0.0064</td>
<td>0.0030</td>
<td>0.0014</td>
</tr>
</tbody>
</table>
paper is defined as the ability of the matrix to produce larger variance values as the size of the matrix is expanded. For instance, candidates such as $H_2$, $H_4$, $H_6$, and $H_8$ show higher sensitivity as the order of the matrix is increased. Other candidates like $H_1$, $H_3$, $H_5$ and $H_7$ also show an increase in variance values except for the fourth-order PMLD. For example, the individual $H_7$ has shown an increase in variance values in the second-, third-, and fifth-order matrices, but it displays a lower variance value in the fourth-order as compared to the third-order PMLD. Also, variability is defined as the qualitative measure that reflects the extent of data spread under the Gaussian distribution curves, which is important for the classification of individuals. The variability of the data obtained for a healthy cohort for the third-order PMLD and the fifth-order PMLD is seen in Figure 3a.

Table 3. The values for the variance of the norm of the PMLD matrix (from the second to the fifth orders) for the healthy cohort.

<table>
<thead>
<tr>
<th>List of Healthy Candidates</th>
<th>2nd-Order Matrix</th>
<th>3rd-Order Matrix</th>
<th>4th-Order Matrix</th>
<th>5th-Order Matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>JT, QRS</td>
<td>JT, RR</td>
<td>JT, QRS, RR</td>
<td>JT, QRS, RR, DP</td>
</tr>
<tr>
<td>$H_1$</td>
<td>0.0018</td>
<td>0.0004</td>
<td>0.0012</td>
<td>0.0069</td>
</tr>
<tr>
<td>$H_2$</td>
<td>0.0015</td>
<td>0.0002</td>
<td>0.0029</td>
<td>0.0029</td>
</tr>
<tr>
<td>$H_3$</td>
<td>0.0013</td>
<td>0.0009</td>
<td>0.0023</td>
<td>0.0030</td>
</tr>
<tr>
<td>$H_4$</td>
<td>0.0007</td>
<td>0.0005</td>
<td>0.0014</td>
<td>0.0019</td>
</tr>
<tr>
<td>$H_5$</td>
<td>0.0030</td>
<td>0.0011</td>
<td>0.0040</td>
<td>0.0020</td>
</tr>
<tr>
<td>$H_6$</td>
<td>0.0018</td>
<td>0.0005</td>
<td>0.0025</td>
<td>0.0033</td>
</tr>
<tr>
<td>$H_7$</td>
<td>0.0013</td>
<td>0.0005</td>
<td>0.0031</td>
<td>0.0021</td>
</tr>
<tr>
<td>$H_8$</td>
<td>0.0007</td>
<td>0.0018</td>
<td>0.0018</td>
<td>0.0019</td>
</tr>
</tbody>
</table>

From the data tabulated in Table 4, it can be observed that the candidates $U_1$, $U_3$, $U_6$, and $U_8$ show an increase in sensitivity as the matrix order is increased. Candidate $U_2$ shows a very consistent sensitivity as the order of the matrix is increased. In terms of variability, while there is some degree of variation present within the unhealthy cohort, if a comparison is made between the variability of the healthy and the unhealthy cohort, a higher degree of variability will be observed in the healthy group. This observation suggests that the variability in cardiac parameters among individuals in the healthy group is more pronounced than that among individuals in the unhealthy group. Furthermore, while the fourth-order matrix does yield a greater sensitivity in a majority of the cases, it could not provide a one hundred percent increase in sensitivity in all cases as compared to other matrix orders. Therefore, it is normal to focus solely on the matrix orders that consistently demonstrated a hundred percent increase in sensitivity in all of the matrix dimensions. As a result, the third-order PMLD and the fifth-order PMLD matrices are utilized for classification purposes as outlined in the subsequent section.

Table 4. The values for the variance of the norm of the PMLD matrix (from the second to the fifth orders) for the unhealthy cohort.

<table>
<thead>
<tr>
<th>List of Unhealthy Candidates</th>
<th>2nd-Order Matrix</th>
<th>3rd-Order Matrix</th>
<th>4th-Order Matrix</th>
<th>5th-Order Matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>JT, QRS</td>
<td>JT, RR</td>
<td>JT, QRS, RR</td>
<td>JT, QRS, RR, DP</td>
</tr>
<tr>
<td>$U_1$</td>
<td>0.0005</td>
<td>0.0001</td>
<td>0.0006</td>
<td>0.0031</td>
</tr>
<tr>
<td>$U_2$</td>
<td>0.0002</td>
<td>0.0001</td>
<td>0.0004</td>
<td>0.0004</td>
</tr>
<tr>
<td>$U_3$</td>
<td>0.0073</td>
<td>0.0023</td>
<td>0.0082</td>
<td>0.0229</td>
</tr>
<tr>
<td>$U_4$</td>
<td>0.0049</td>
<td>0.0012</td>
<td>0.0064</td>
<td>0.0039</td>
</tr>
<tr>
<td>$U_5$</td>
<td>0.0014</td>
<td>0.0003</td>
<td>0.0030</td>
<td>0.0100</td>
</tr>
<tr>
<td>$U_6$</td>
<td>0.0007</td>
<td>0.0004</td>
<td>0.0014</td>
<td>0.0036</td>
</tr>
</tbody>
</table>
3.3. The Classification of the Matrix Architecture for the Different Orders

In order to perform the classification, the techniques employed in the previous study are applied here too [20]. To establish the variation interval for classification, the variance values for the third- and the fifth-order PMLD matrices are utilized from Tables 2 and 3. The variance values are fitted to the Gaussian distribution as shown in Figure 2. The variance values for the healthy candidates are denoted by blue circles along the x-axis and the variance values for unhealthy individuals are represented by red circles as depicted in Figure 2. Once the distributions are in place, the normal distribution of the dataset (the variance values) is validated by performing the Anderson–Darling test. The Anderson–Darling test is chosen to validate the normal distribution of the variance values because of its sensitivity to deviations in the tails of the distribution, which is crucial for our dataset and is also utilized in [20]. Also, the 95% confidence intervals for the mean and variance values of both healthy and unhealthy individuals are used. The same confidence intervals are also used in [17]. Upon confirming the normal distribution, we proceed with applying the one sigma rule to the distributed dataset for the third- and the fifth-order PMLD matrices for the generation of the variation interval. To do so, the one sigma rule is applied upon the distributed dataset where the left boundary is defined as mean minus sigma ($\mu_3 - \sigma_3$) and the right boundary is established as mean plus sigma ($\mu_3 + \sigma_3$) for the third-order PMLD matrix as shown in Figure 2a. Similarly, the same statistical technique is adopted to apply the one sigma rule for the fifth-order PMLD matrix where the left boundary is defined as mean minus sigma ($\mu_5 - \sigma_5$) and the right boundary is defined as ($\mu_5 + \sigma_5$). The variation interval is built for classification purposes and is visualized by the double-headed arrow as depicted in Figure 2b. This technique assures both clarity and precision in establishing classification boundaries, allowing for accurate categorization for individuals based on the variance values.

![Figure 2](image)

**Figure 2.** The distribution graphs for the healthy and unhealthy cohort for the third and fifth order PMLD.

To enhance the classification accuracy of test candidates, a machine-trained model was developed. This involved creating a self-trained model to validate the algorithm and
to establish a decision support system for classification purposes. It is crucial for this system to adhere to the rules outlined in Table 5 for both healthy and unhealthy candidates. In Table 5, the variable $Z$ represents the variance value for the test candidate. This variable is tested against the three conditions presented and will classify the individual accordingly for both the third- and the fifth-order PMLD matrices. An indicator variable $IND$ is also introduced, which will determine where the test individual falls inside the classification interval according to the conditions explained as follows.

Table 5. Classification criteria for the development of the decision support system. The variable $Z$ stands for the variance of the new candidate, and $IND$ stands for the probability indicator. The $h$ and $u$ indices denote a healthy individual, whereas $a$ denotes an unhealthy individual.

<table>
<thead>
<tr>
<th>Condition 1</th>
<th>Condition 2</th>
<th>Condition 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Z \leq (\mu_h^3 - \sigma_h^3)$</td>
<td>$Z \geq (\mu_h^3 + \sigma_h^3)$</td>
<td>$(\mu_u^3 + \sigma_u^3) \leq Z \leq (\mu_h^3 - \sigma_h^3)$</td>
</tr>
<tr>
<td>Then, the probability indicator is:</td>
<td>IND = 0</td>
<td>IND = 1</td>
</tr>
<tr>
<td>The candidate is classified as</td>
<td>Healthy</td>
<td>Unhealthy</td>
</tr>
</tbody>
</table>

For the 3rd-Order PMLD

<table>
<thead>
<tr>
<th>Condition 1</th>
<th>Condition 2</th>
<th>Condition 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Z \leq (\mu_h^5 - \sigma_h^5)$</td>
<td>$Z \geq (\mu_h^5 + \sigma_h^5)$</td>
<td>$(\mu_u^5 + \sigma_u^5) \leq Z \leq (\mu_h^5 - \sigma_h^5)$</td>
</tr>
<tr>
<td>Then, the probability indicator is:</td>
<td>IND = 0</td>
<td>IND = 1</td>
</tr>
<tr>
<td>The candidate is classified as</td>
<td>Healthy</td>
<td>Unhealthy</td>
</tr>
</tbody>
</table>

For the 5th-Order PMLD

<table>
<thead>
<tr>
<th>Condition 1</th>
<th>Condition 2</th>
<th>Condition 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Z \leq (\mu_h^5 - \sigma_h^5)$</td>
<td>$Z \geq (\mu_h^5 + \sigma_h^5)$</td>
<td>$(\mu_u^5 + \sigma_u^5) \leq Z \leq (\mu_h^5 - \sigma_h^5)$</td>
</tr>
<tr>
<td>Then, the probability indicator is:</td>
<td>IND = 0</td>
<td>IND = 1</td>
</tr>
<tr>
<td>The candidate is classified as</td>
<td>Healthy</td>
<td>Unhealthy</td>
</tr>
</tbody>
</table>

Condition 1 ($IND = 0$): If $Z$ is less than or equal to the lower boundary of the variance values for healthy individuals ($\mu_h^3 - \sigma_h^3$), the probability indicator is set to zero. This condition signifies a low probability of the candidate belonging to the healthy group.

Condition 2 ($IND = 1$): If $Z$ is greater than or equal to the upper boundary of the variance values for unhealthy individuals ($\mu_u^3 + \sigma_u^3$), the probability indicator is set to one. This condition suggests a high probability of the candidate belonging to the unhealthy group.

Condition 3 ($0 \leq IND \leq 1$): If $Z$ falls within the range between the upper boundary of the variance values for unhealthy individuals and the lower boundary of the variance values for healthy individuals, the probability indicator is calculated using a linear interpolation formula as follows:

$$IND = \frac{Z - (\mu_u^3 + \sigma_u^3)}{(\mu_h^3 - \sigma_h^3) - (\mu_u^3 + \sigma_u^3)}$$  \hspace{1cm} (11)

The condition presented in Equation (11) reflects a probability between 0 and 1, indicating the candidate’s likelihood of belonging to either the healthy or the unhealthy group based on the degree to which their variance value deviates from the established boundaries.

Testing Candidates

The cardiac parameters are captured from the resting ECG for the three test candidates. The recorded cardiac parameters are the duration of the JT wave, the duration of the QRS complex, the RR inter-beat interval, the AP wave, and the DP wave, respectively. Those cardiac parameters are transformed into the time series $(x_n: n = 0,1,2,3\ldots)$, $(y_n: n = 0,1,2,3\ldots)$, $(z_n: n = 0,1,2,3\ldots)$, $(u_n: n = 0,1,2,3\ldots)$, and $(v_n: n = 0,1,2,3\ldots)$, respectively. It is important to note that there is pre-knowledge about the three
candidates regarding their health status (having or not having had atrial fibrillation occurrences before). The first two candidates have no recorded history of atrial fibrillation whereas the third test candidate has a recorded history of atrial fibrillation occurrences. With this prior knowledge, the aim is to test the accuracy of the proposed matrix architectures (the third- and the fifth-order PMLD) as well as the proposed classification techniques.

First, the preprocessing techniques are applied to the time series as mentioned already in the methods section. The preprocessed parameters are then fitted into the third- and the fifth-order PMLD matrix architectures, and the variance of the norm is produced for the three test candidates.

For the first candidate, it is observed that the individual’s variance values have placed the person within the variation interval, positioning them towards the left end of the distribution for the third-order PMLD as seen in Figure 3a. This can also be observed as a visual representation by the inclusion of an asterisk symbol, indicating the individual’s placement within the classification interval in Figure 3a. The interpolation techniques are performed using the developed decision support system using the criteria listed in Table 5. The decision support system shows that the individual satisfies condition 1, which is an indication of a healthy state (cardiovascular health-wise). Furthermore, the probability distribution chart illustrates a probability of less than 0.1 for the individual to be unhealthy, signifying their placement within the healthy region as seen in Figure 3b.

![Figure 3](image.png)

**Figure 3.** First test candidate under the classification criteria for the third and fifth order PMLD matrices.

Finally, a semi-gauge indicator tool is used to visualize an individual’s health status using three indicated regions (green, yellow, and red). When the needle of the gauge points to green, it indicates that the person is in the health-wise (cardiovascular health-wise) safety zone. If the needle of the gauge is in the center yellow zone, it indicates that the individual is recommended to pursue clinical attention for follow ups on his cardiovascular health. Finally, if the needle of the gauge is in the red zone, it indicates that the person is unhealthy (cardiovascular health-wise) and requires serious attention from
specialists. The semi-gauge indication tool in Figure 3c produces a reading in the green zone, meaning that the test candidate is health-wise, in the safe zone.

The same computational techniques are employed for the classification of the same individual using the fifth-order PMLD matrix architecture. It can be observed that the variance value produced for this individual places him into the healthy candidates category as seen in Figure 3d. The decision support system performs the interpolation based upon the rules established in Table 5 and shows a zero probability for the individual to be classified as unhealthy as seen in Figure 3e. Finally, the semi-gauge indication tool also categorizes the individual under the green region as seen in Figure 3f.

Both the third and fifth order PMLD matrix configurations identify the individual as a healthy candidate. It is important to emphasize that the individual is known to be healthy, as there have been no previously recorded symptoms of atrial fibrillation episodes. The purpose of this analysis is to add the individuals’ cardiac parameters into the entire algorithm and classification technique and to determine if the system is capable of classifying the person as healthy for both the third and fifth order PMLD matrix architectures.

Figure 4 shows the ECG parameters of another test candidate. The decision support system for third-order PMLD matrices implies that the person is healthy, as shown in Figure 4a. The probability of the individual being unhealthy is close to zero, as observed in Figure 3b. Furthermore, the semi-gauge indicator tool depicts the individual’s health state under the green color (Figure 4c). Again, the same computational routine is performed for the ECG parameters for the fifth-order PMLD matrix. It can be noted that the classification criteria classify the individual under the category of healthy persons (Figure 4d). The decision support system shows a zero probability for the individual being unhealthy as seen in Figure 4e. Similarly, the semi-gauge indication tool also depicts the health status of the person under the green zone (Figure 4f). The increased sensitivity attained with the fifth-order PMLD is obvious when the needle is positioned inside the green zone. Whether using the third or fifth order matrices, the decision support system continually provides engaging conclusions, situating the person within the green zone, which indicates a positive health state.

Figure 4. Second test candidate under the classification criteria for the third and fifth order PMLD matrices.
The ECG parameters of the third test candidate undergo evaluation within the classification criteria of both the third- and fifth-order PMLD matrices. As depicted in Figure 5a–c, the individual falls within the classification interval of unhealthy candidates for the third-order PMLD matrix. The developed decision support system shows a large probability for the individual being unhealthy. This classification also aligns with the indication provided by the semi-gauge indication tool, which points towards the red zone.

Moreover, the individual’s ECG parameters are also tested and subjected to evaluation under the fifth-order PMLD matrix. Again, the developed decision support system classifies the individual as an unhealthy candidate, as evidenced by the semi-gauge indication tool (Figure 5d–f).

The utilization of the PMLD matrix architecture emerges as a robust and flexible mathematical tool for effectively analyzing the relationships between the parameters of the ECG signal. Previous studies [12–14] have already underscored the efficacy of PMLD as a mathematical technique for analyzing ECG signals and their parameters.

Various computational aspects are taken into account in order to analyze the ECG parameters for the detection of atrial fibrillation episodes for a specific group of individuals in this paper. This study elucidates various factors contributing to the functionality of PMLD matrices. For example, the arrangement of ECG parameters within the matrix structure, the degree of freedom available for the balancing of its elements, and the order of the matrix play crucial roles in the proposed classification technique.

In general, increasing the matrix order has been shown to enhance the PMLD’s sensitivity. However, an important observation is made about the sensitivity of the fourth-order PMLD. In few of the cases, it is observed that the sensitivity of the fourth-order matrix was slightly lower than that of the third order. It other words, the second, third and fifth order of the PMLD have shown a reliable increase in the sensitivity whereas the same increase cannot be achieved by the fourth-order PMLD matrix. Again, as previously mentioned, there are several aspects that influence the intended results. For instance, many contemporary methods for the detection of atrial fibrillation are focused on finding
the impact of certain diseases or medication on atrial fibrillation. On the other hand, major techniques used for the detection of atrial fibrillation are based on the analysis of relatively long ECG records (sometimes persons are required to wear ECG monitors for more than 24 h). Our approach is completely different. Firstly, we are designing an algorithm capable of detecting atrial fibrillation episodes from short ECG records (no longer than 07 min). Secondly, the participants in the unhealthy group do not take any medications, so the impact of diseases and medications completely falls out of the scope of this paper. Therefore, a straightforward comparison with other techniques and methods is simply not possible.

Nonetheless, the PMLD’s overall effectiveness, sensitivity, and variability in ECG signal analysis are clear and unquestionable. It is important to mention that the proposed method is computationally efficient, making it suitable for real-time applications. The time complexity of the PMLD matrix approach is manageable, even with higher-order matrices as shown already.

Furthermore, the PMLD matrix approach demonstrates its efficacy even with limited datasets. While this serves as a strong point to its robustness, it also poses a limitation to the study. Specifically, the self-training nature of the decision support system implies that its efficiency and accuracy are contingent upon the size of the data fed into it. Hence, the study highlights the critical interplay between dataset size and the system’s predictive capabilities.

In essence, the study underscores the multifaceted nature of PMLD in ECG signal analysis, emphasizing its significance over the MA1 and MA2 matrix architectures.

4. Conclusions

Three matrix architectures have been presented for the analysis of ECG parameters, for various orders, for the detection of atrial fibrillation episodes. A group of 15 individuals was tested with the three matrix architectures. It has been found that the matrix norm has produced convincingly better results as compared to the large discriminant of the matrix. Moreover, the spatial expansion of each of the matrix structures was analyzed in detail, and it has been shown that the PMLD is a better fit for the expandability of analysis as compared to the MA1 and MA2 matrices. Therefore, the third- and fifth-order PMLD architectures were used for classification techniques, and it is shown that the sensitivity of the matrix structure is increased as the order of the matrix is increased. These outcomes are affirmed after testing a few candidates under the developed classification criterion. Without doubts, the proposed study could be useful for the clinical diagnostics of atrial fibrillation.

Author Contributions: Conceptualization, N.W.Q. and M.R.; methodology, N.W.Q. and M.R.; software, N.W.Q.; validation, M.R., Z.N.; formal analysis, N.W.Q.; investigation, N.W.Q.; resources, A.V. and G.J.; data curation, G.J.; writing—original draft preparation, N.W.Q.; writing—review and editing, M.R.; visualization, N.W.Q. and M.R.; supervision, M.R.; project administration, A.V.; funding acquisition, G.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: This research was conducted with the approval of the LUHS Bioethics Committee (permission number BE-2-4 15 March 2016).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study. Written informed consent was obtained from the patients to publish this paper.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.
Abbreviations

ECG  Electrocardiogram
FMLD  Perfect Matrices of Lagrange Differences
MA1  Matrix Architecture 1
MA2  Matrix Architecture 2
AF  Atrial Fibrillation
CNN  Convolution Neural Networks
RNN  Recurrrent Neural Networks

References


**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.