*Article*

# Electrocardiogram Analysis by Means of Empirical Mode Decomposition-Based Methods and Convolutional Neural Networks for Sudden Cardiac Death Detection

Manuel A. Centeno-Bautista [1], Angel H. Rangel-Rodriguez [1], Andrea V. Perez-Sanchez [1], Juan P. Amezquita-Sanchez [1], David Granados-Lieberman [2] and Martin Valtierra-Rodriguez [1,*]

[1] ENAP-Research Group, CA-Sistemas Dinamicos y Control, Universidad Autonoma de Queretaro (UAQ), Campus San Juan del Rio, Rio Moctezuma 249, Col. San Cayetano, San Juan del Rio 76807, Mexico
[2] ENAP-Research Group, Departamento de Ingenieria Electromecanica, Tecnologico Nacional de Mexico, Instituto Tecnologico Superior de Irapuato (ITESI), Carretera Irapuato-Silao km 12.5, Colonia El Copal, Irapuato 36821, Mexico
* Correspondence: martin.valtierra@enap-rg.org

**Abstract:** Sudden cardiac death (SCD) is a global health problem, which represents 15–20% of global deaths. This type of death can be due to different heart conditions, where ventricular fibrillation has been reported as the main one. These cardiac alterations can be seen in an electrocardiogram (ECG) record, where the heart's electrical activity is altered. The present research uses these variations to be able to predict 30 min in advance when the SCD event will occur. In this regard, a methodology based on the complete ensemble empirical mode decomposition (CEEMD) method to decompose the cardiac signal into its intrinsic mode functions (IMFs) and a convolutional neural network (CNN) for automatic diagnosis is proposed. Results for the ensemble empirical mode decomposition (EEMD) method and the empirical mode decomposition (EMD) method are also compared. Results demonstrate that the combination of the CEEMD and the CNN is a potential solution for SCD prediction since 97.5% of accuracy is achieved up to 30 min in advance of the SCD event.

**Keywords:** electrocardiogram signal; sudden cardiac death; empirical mode decomposition; ensemble empirical mode decomposition; complete ensemble empirical mode decomposition; convolutional neural network

## 1. Introduction

Sudden cardiac death (SCD) is characterized by an unexpected death triggered by a cardiac condition in a subject with unknown or known heart disease [1]. SCD is recognized as a significant health problem because it represents about 15–20% of deaths worldwide [2]. In particular, an SCD event is caused by a cardiac arrhythmia [3], which can produce the heart's inability to pump blood to vital organs [4], generating a risk to the person [5]. Particular problems can provoke fatal cardiac arrhythmias in an SCD episode, e.g., bradyarrhythmia, ventricular tachycardia, and ventricular fibrillation (VF), among others [6,7], but several studies indicate that VF is the principal cause of an SCD episode [8–10]. In a VF, the heart signal is unstable because a sequence of wave breaks and self-regenerating reentry are caused in the heart ventricles [11], causing the death of a person a few minutes after the onset signs [12]. Therefore, an early forecast of an SCD episode is transcendental because it will permit people to receive timely treatment before possible SCD, increasing their survival possibilities.

In recent years, investigators around the world have presented diverse methods to forecast an SCD episode employing electrocardiogram (ECG) signals; in this regard, the database untitled "Sudden Cardiac Death Holter" (SCDH) [13] has been used by different researchers to forecast an SCD, proposing methods based on machine learning [12,14–16] and deep learning architectures [17,18]. For example, Acharya et al. (2015) [12] integrate

a discrete wavelet transform, 18 nonlinear features (i.e., Hurst's exponent, approximate entropy, sample entropy, etc.), and a support vector machine to forecast an SCD. The authors reported that their method could predict an SCD 4 min before it occurs with 92.11% of accuracy. Amezquita-Sanchez et al. [14] presented a machine learning scheme based on the integration of wavelet packet transform, homogeneity indicator, and an enhanced probabilistic neural network classifier for predicting an SCD event utilizing ECG signals provided in the database mentioned above. The authors demonstrated that their proposal could predict an SCD 20 min before its onset. Khazaei et al. [15] investigated the forecast of an SCD episode utilizing the heart rate variability (HRV) from the ECG signals employing the Pan–Tompkins algorithm. The obtained HRV signals were analyzed by a recurrence quantification analysis and the increment entropy to discover features in the signal. A decision tree is used to classify these characteristics with a Euclidean distance. They established that their proposal has an accuracy of 95% to forecast an SCD episode 6 min before its onset. In addition, Vargas-Lopez [16] integrated the empirical mode decomposition with fractal dimension (i.e., Katz, Higuchi, and box dimension indices) and entropy (Shannon and permutation entropy indices) in combination with an artificial neural network known as multilayer perceptron classifier to predict an SCD event. The investigators described that their proposal is capable of predicting an SCD 25 min before its onset with 94% of accuracy.

On the other hand, in recent years, the use of deep learning architectures for predicting an SCD event has been introduced. For instance, Kaspal et al. [17] combined a recurrence complex network with a convolutional neural network (CNN) to predict an SCD episode, obtaining 90.60% of accuracy. Saraghi & Isa [18] integrated a wavelet packet transform with a CCN to predict an SCD event 30 min before its commencement. The authors indicated that 95.89% of accuracy in forecasting an SCD is reached. In these works, CNNs are used because they do not need human supervision for the extraction of important features that allow the pattern recognition task.

Although the works mentioned above report promising results to forecast an SCD, there are some opportunities for improvement, e.g., (1) increase the time window to forecast an SCD, (2) reduce the complexity and computational resources of the method, and (3) improve the prediction accuracy. In this regard, the research of novel methods capable of contributing to the solution of the previous points (i.e., time window, computational resources, and accuracy) is essential because it permits patients to go to the health center and take timely medical precautions.

Due to the importance of proposing a new proficient method in forecasting an SCD with high accuracy and enough of a time window for predicting an SCD episode, this work investigates the capabilities of the empirical mode decomposition (EMD)-based methods combined with a CNN for the analysis of ECG signals with the object of forecasting an SCD episode 30 min prior to its onset. To perform this task, the EMD methods are utilized for decomposing the ECG signals in diverse frequency bands, namely intrinsic mode functions (IMFs), according to the frequential information contained in the signals. Next, the obtained frequency bands for each signal are combined to generate images. Finally, the obtained images are employed as inputs in a CNN for automatically forecasting an SCD event. It is essential to mention that diverse EMD-based methods, i.e., EMD, ensemble EMD (EEMD), and complete EEMD (CEEMD), are investigated to determine which one provides the most relevant features to be associated with an SCD episode; also, several CNN configurations (i.e., number of layers, number of filters, batch size, and epochs) are also investigated to obtain the highest accuracy with the least computing resources. The proposed method's effectiveness is tested with the data available on the MIT-BIH databases, which contain information about 20 patients with SCD [13] and 18 patients with a normal cardiac rhythm [19]. The results demonstrate that the features highlighted by EMD methods in combination with the CNN can be employed as a reliable tool for identifying suitable features in the ECG signals for its use in SCD forecast.

## 2. Theoretical Background

### 2.1. ECG Data

2.1.1. Data Used

To validate and test the proposed method to forecast an SCD episode, two open-access databases known or titled SCDH provided by the Massachusetts Institute of Technology (MIT) and normal sinus rhythm (NSR) supplied by Beth Israel Deaconess Medical Center (BIDMC) and MIT are employed. In particular, the SCDH database incorporates the ECG signals acquired experimentally from 23 subjects (15 men and 8 women) with an age range of 49.5 ± 32.5, but only the ECG signals of 20 participants who suffered from an SCD episode provoked by ventricular fibrillation are employed because the other 3 subjects have other alterations (e.g., ventricular tachycardia and hypertrophic cardiomyopathy). Table 1 summarizes the information of the 20 participants, in particular, the exact time when their SCD event started, as well as their gender and age [13]. This time is employed to forecast the occurrence of SCD. Conversely, the NSR database incorporates the ECG signals of 18 subjects (5 men and 13 women) with an age range of 35 ± 15 who presented a regular heart rate; hence, these signals were taken as the control group.

**Table 1.** General information patients provide by SCDH database.

| Patient | Gender | Age | Ventricular Fibrillation Onset Time (Hours:Minutes:Seconds) | Subjacent Cardiac Rhythm |
|---|---|---|---|---|
| 1 | Male | 43 | 07:54:33 | Sinus |
| 2 | Female | 72 | 13:42:24 | Sinus |
| 3 | Unnamed | 62 | 16:45:18 | Sinus with sporadic demand ventricular pacing |
| 4 | Female | 30 | 04:46:19 | Sinus |
| 5 | Male | 34 | 06:35:44 | Sinus |
| 6 | Female | 72 | 24:34:56 | Atrial fibrillation |
| 7 | Male | 75 | 18:59:01 | Atrial fibrillation |
| 8 | Female | 89 | 01:31:13 | Atrial fibrillation |
| 9 | Unnamed | — | 08:01:54 | Sinus |
| 10 | Male | 66 | 04:37:51 | Sinus |
| 11 | Male | – | 02:59:24 | Sinus |
| 12 | Male | 35 | 15:37:11 | Sporadic ventricular pacing |
| 13 | Male | – | 19:38:45 | Sinus |
| 14 | Male | 68 | 18:09:17 | Sinus |
| 15 | Female | – | 03:41:47 | Sinus |
| 16 | Male | 34 | 06:13:01 | Sinus |
| 17 | Male | 80 | 02:29:40 | Sinus |
| 18 | Female | 68 | 11:45:43 | Atrial fibrillation |
| 19 | Female | 67 | 22:58:23 | Sinus with sporadic pacing |
| 20 | Female | 82 | 02:32:40 | Sinus |

2.1.2. Preparation of the ECG Signals

The ECG signals from the SCDH database were monitored for 24 h utilizing a sampling frequency of 250 Hz [13]; nevertheless, these signals have been resampled to 128 Hz (this procedure is accomplished by convolving the original ECG signal with a digital low-pass filter) in order to maintain consistency with the sampling frequency value of the NSR database. It is worth noticing that from the 24 h monitored for each subject, uniquely the first 30 min before SCD are extracted and evaluated in this work. The selected time window

permits investigating and somehow comparing other works or investigations reported in the literature. The analyzed time window, i.e., 30 min, in this work is separated into 1 min lapses as follows: the first 1 min interval before the SCD onset, the second 1 min interval before the SCD onset, and so on. On the other hand, the control group (CG) is obtained by randomly extracting 1 min intervals from the NSR database. It is worth noticing that the length of the intervals, 1 min, is chosen because this time window enables the detection of reliable characteristics in the ECG signals in order to associate them with different phenomena [14,16,20,21]. For this motive, the 1 min lapse is also investigated in this work in order to forecast until 30 min before an SCD episode. Figure 1 illustrates the monitored ECG signal for a subject with a normal cardiac rhythm (Figure 1b) and the first and second 1 min intervals prior to an SCD episode (Figure 1a), respectively. According to this figure, it is worth noticing that no one difference is visually obtained between both conditions; then, it is essential to provide a new methodology or method with the potential to identify suitable features in ECG signals in order to associate them with an SCD prediction.
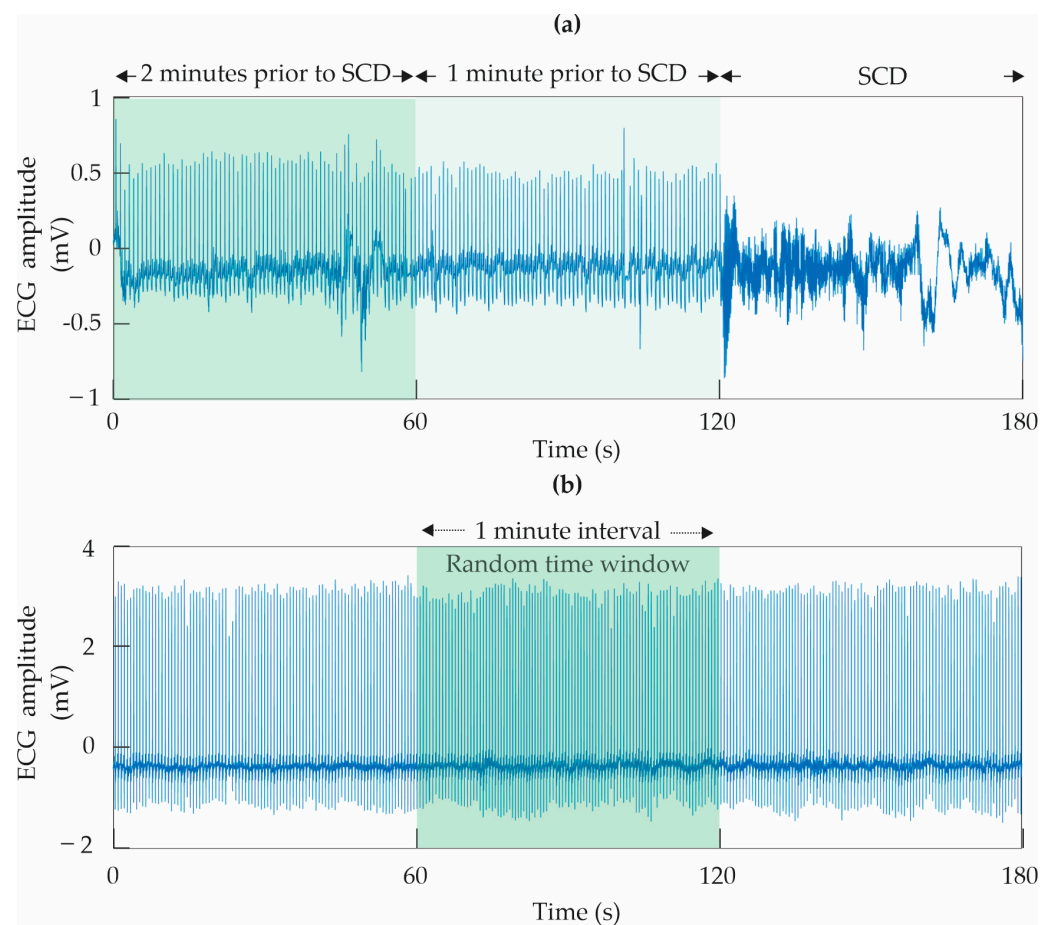


**Figure 1.** Acquired ECG signals for (**a**) first and second 1 min intervals prior to an SCD event and (**b**) normal cardiac rhythm.

## 2.2. EMD-Based Methods

### 2.2.1. EMD Method

Introduced by Huang et al. [22], EMD is an adaptive method capable of analyzing or examining stationary and non-stationary time or 1D signals. In this regard, EMD is characterized by decomposing a 1D signal in diverse intrinsic mode functions (IMFs) or frequency bands according to its frequential information. For obtaining or considering an IMF, it has to fulfill the subsequent conditions: (1) the number of zero crossings as well as extrema in the 1D signal must be either different or equal only by one, and (2) the lower

and upper envelopes have to be symmetrical, indicating that a mean value equal to zero is estimated.

For estimating each IMF, a process called "sifting process" is performed. In general, it is based on five steps defined as follows:

Step 1. Estimate the local maximum and minimum points of the 1D signal ($x(t)$).

Step 2. Connect the local maximum and minimum points identified in step 1 in order to estimate a lower and upper envelope utilizing cubic-splines. The mean of both envelopes, $m_1(t)$, is subtracted from the original 1D signal, $x(t)$, for obtaining a new 1D signal, $h_1(t)$, thus:

$$h_1(t) = x(t) - m_1(t) \tag{1}$$

If the new 1D signal, $h_1(t)$, does not fulfill the conditions (1) and (2), the steps (1 and 2) must be repeated until $h_k(t)$ fulfills both conditions; so, $h_k(t)$ is deemed as the first $IMF_1$ and represented by:

$$c_1(t) = h_k(t) = IMF_1 \tag{2}$$

Step 3. Once the $IMF_1$ is estimated, it is subtracted from the initial 1D signal for calculating a residue signal, $r_1(t)$, denoted by:

$$r_1(t) = x(t) - IMF_1 \tag{3}$$

Step 4. If the obtained residue signal, $r_1(t)$, is deemed a monotonic function, the process is stopped, indicating that no more IMFs are estimated. On the other hand, if it is not considered a monotonic function, the residual signal is now considered the original 1D signal, which is again evaluated by steps (1 to 3) for estimating the other IMFs.

Step 5. The initial 1D signal, $x(t)$, can be reconstructed by summing the obtained IMFs and the final residue signal in this way:

$$x(t) = \sum_{i=1}^{N} IMF_i + r_N(t) \tag{4}$$

Figure 2 shows the flowchart to compute the EMD method which is also the basis for the EEMD and CEEMD methods.

### 2.2.2. EEMD Method

EEMD, an improved version of the EMD method [23], is characterized by being a noise-assisted method employed for decomposing the 1D signals in their fundamental components according to the following steps:

Step 1. Create new 1D signals by combining the initial 1D signal, $x(t)$, with white Gaussian noise series, $n_i(t)$, for $m$ trials as follows:

$$x_i(t) = x(t) + n_i(t) \ for \ i = 1, \ 2, \ 3, \ \ldots, \ m \tag{5}$$

Step 2. Decompose the 1D signals created in step 1 by employing the original EMD method.

Step 3. Estimate a true frequency band or IMF, identified by $j$, as follows:

$$EEMD_j = \frac{1}{m} \sum_{i=1}^{m} c_{ij}(t) \tag{6}$$

where the frequency band or IMF, identified by $j$ for the trial $i$, is denoted by $c_{ij}(t)$.

### 2.2.3. CEEMD Method

Presented by Torres et al. [24], the CEEMD technique is deemed an enhanced version of the EEMD method since it provides a better spectral separation of frequencies or modes

in different IMFs. In this regard, the EEMD method calculates a residue $r_i(t)$ for each trial; but, the CEEMD method calculates a unique first residue, $r_i(t)$, as follows:

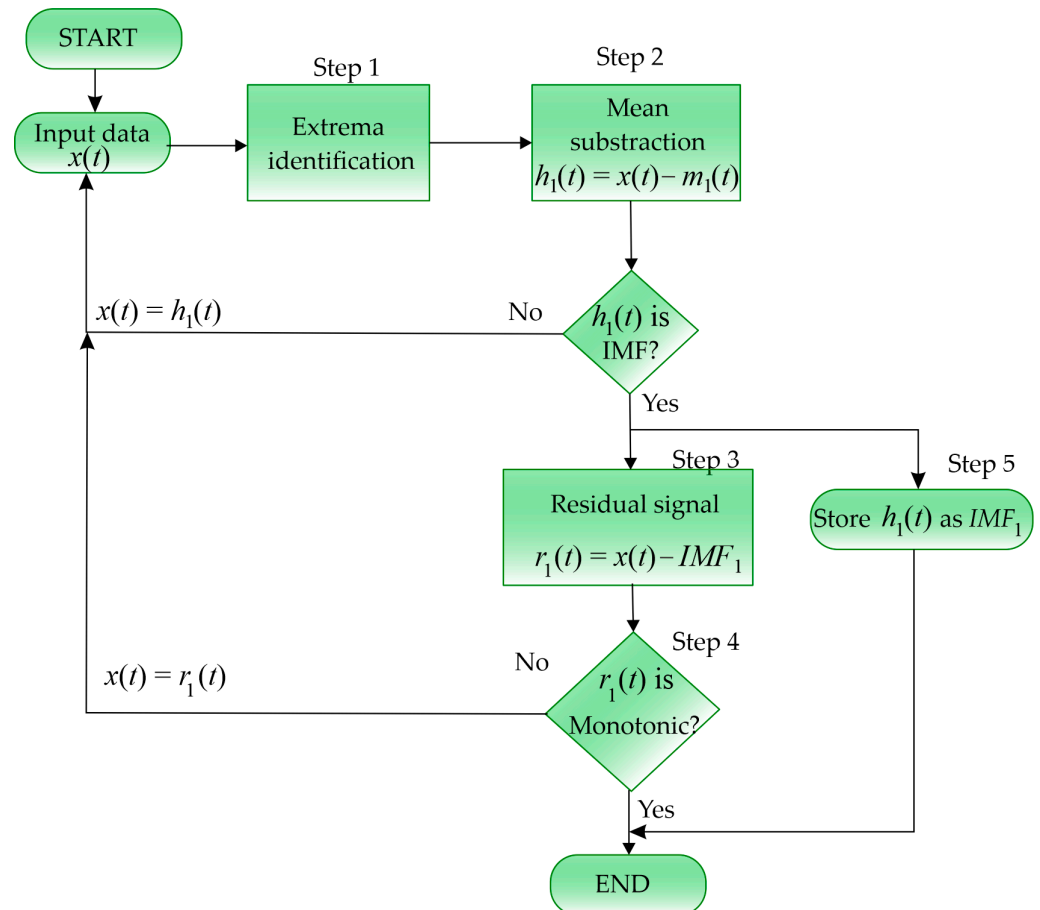$$r_1(t) = x(t) - IMF_1 \qquad (7)$$



**Figure 2.** Flowchart for the EMD method.

where $IMF_1$, a true frequency band or IMF, is estimated according to the EEMD method. Thus, an ensemble of $r_1$ plus different realizations is carried out to calculate the true $IMF_2$ or second frequency band. The near residue is calculated by $r_2(t) = r_i(t) - IMF_1$. This process is repeated until all the true IMFs have been calculated. Although the CEEMD method is considered an improved version of EEMD and EEMD an improved version of EMD, these three methods have to be analyzed since there is no rule/concept that indicates which method is better for a specific application, mainly considering that advantages and disadvantages for each method have to be taken into account.

*2.3. Image Representation*

In this work, a deep learning scheme based on CNNs for automatic pattern recognition is proposed. As these schemes typically work with images, the strategy shown in Figure 3 is implemented. That is, the 1 min segment is processed by any of the EMD-based methods to obtain its IMFs or frequency bands, which contain information related to the phenomenon under study. Then, they are combined to obtain a 3D surface representation, where the top view is selected. These views, treated as images, could provide suitable information to be associated with the SCD event through a CNN method.
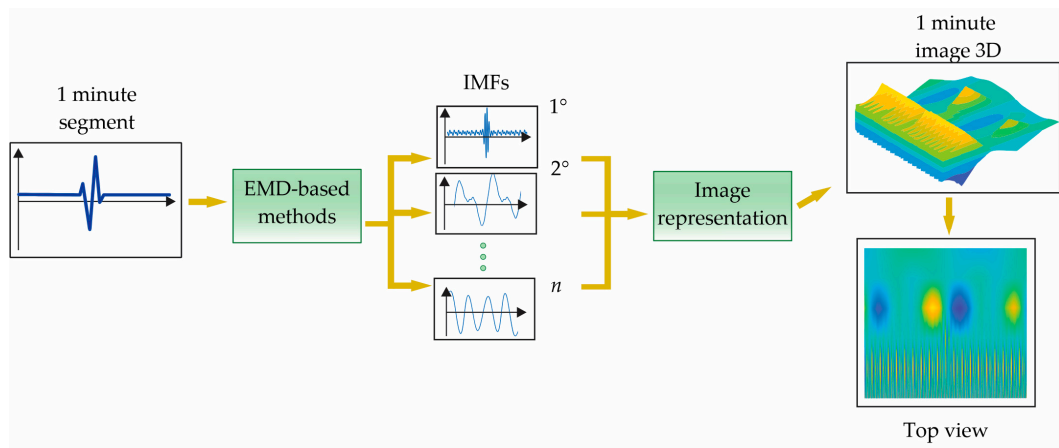
**Figure 3.** Procedure to combine the IMFs estimated by EMD-based methods into an image.

*2.4. CNN*

CNN is characterized by being an innovative deep-learning architecture suitable for pattern recognition in images. In this regard, it employs a single block for identifying and classifying the encountered patterns or features into images in order to associate them automatically with a phenomenon studied, eliminating the hand engineering during the selection and test of features/patterns [25]. For performing this task, it is constituted by four sub-CNNs known as convolutional-, pooling-, fully connected-, and softmax-layer, as displayed in Figure 4.
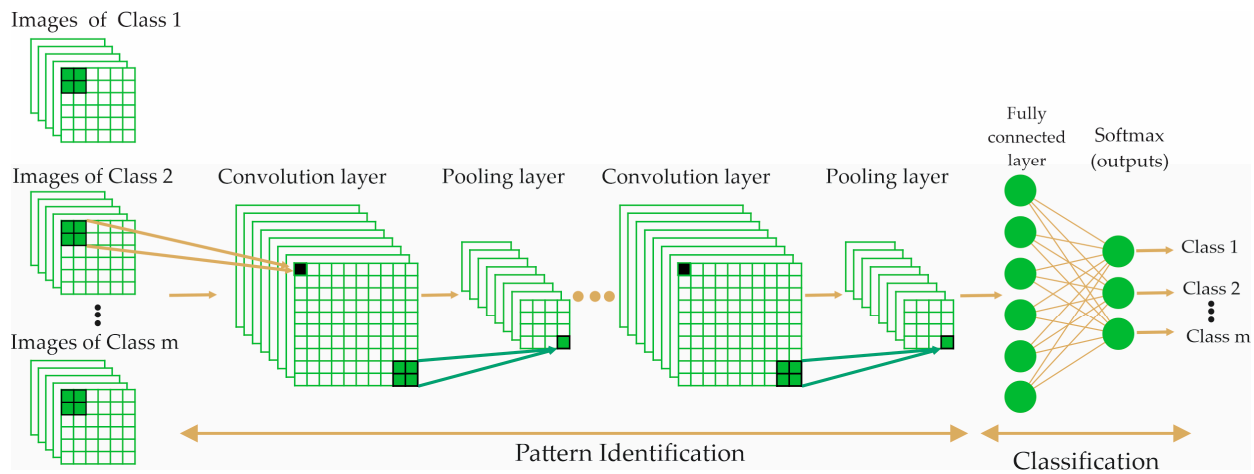


**Figure 4.** CNN architecture.

According to the sequence or structure shown in Figure 3, the analyzed images, $I_m$, (with dimension $h \times \omega$) are convolved, i.e., $*$, with a set of convolutional filters, $F_i$, in order to identify or extract diverse features from the evaluated images as follows [26]:

$$X_i = \sigma\left(\sum F_i * I_m + B_i\right) \tag{8}$$

where $\sigma$ and $B$ imply a nonlinear activation function and a bias term, respectively. To perform this task, each convolutional filter (with a dimension of $k_1 \times k_2$) is convolved with a region or local section of the evaluated image with a stride $s_1$. Hence, the output, $X_i$, is deemed as the maps of patterns estimated by each convolutional filter, which present a dimension of $z_1 \times z_2$ and is determined by [27]:

$$z_1 = \frac{h - k_1 + 2p}{s_1} + 1 \tag{9}$$

$$z_2 = \frac{\omega - k_2 + 2p}{s_1} + 1 \tag{10}$$

where $p$ is a padding parameter, which usually receives a value of 1 with the aim of providing the same spatial resolution for the output and input [27]. The nonlinear activation function recognized as ReLu ($f(y_i) = \max(0, X_i)$), i.e., rectified linear unit, has demonstrated to be the fastest and most suitable function for learning and encountering relevant patterns of each map of patterns in a CNN [28].

Once the feature or characteristic maps are estimated by means of the convolutional layer, they are utilized as inputs to the subsequent layer named pooling layer, which is utilized for subsampling or reducing the dimensionality of the estimated feature maps with the goal of reducing the number or quantity of patterns to be analyzed by next layer or sub-CNN [29]. Particularly, it applies or passes a filter with a size $K_1 \times K_2$ and stride $s2$ on the feature maps for obtaining their maximum (identified as max pooling) of the neighbor values designated by the utilized filter. Thus, the decreased or reduced pattern maps, $Y_i$, with sizes of $Z_1 \times Z_2$, are estimated as follows [28]:

$$Z_1 = \frac{z_1 - K_1}{s_2} + 1 \tag{11}$$

$$Z_2 = \frac{z_2 - K_2}{s_2} + 1 \tag{12}$$

It is worth noting that max pooling has demonstrated to be efficient to enhance the generalization performance and capture invariant patterns [30]. For these reasons, it is employed in this work for identifying the most suitable patterns.

The features estimated by the pooling layer are utilized as input of a traditional multilayer perceptron named fully connected layer, which has the aim of performing the pattern recognition/classification. Finally, a SoftMax layer is utilized for creating the required outputs, in this case, the SCD prediction, through a softmax transfer function. In [27], a more detailed description of CNN can be found.

As there is no specific CNN structure that performs well for any application, mainly considering the computational complexity, in this work, different sizes of input images, different numbers and sizes of filters, as well as different numbers of epochs and sizes of batch are also investigated.

### 3. Methodology

The proposed methodology is shown in Figure 5. In general, it consists of four stages. In the first stage, the signal segmentation in 1 min intervals of both databases is carried out as described in Section 2.1. In the next stage, one EMD-based method, i.e., EMD, EEMD, or CEEMD, is applied to each 1 min segment. Each of these algorithms is set to generate 6 IMFs, since after different tests it was observed that the last IMFs (i.e., the seventh IMF or superior) showed redundant (or nonsignificant) information, mainly in the EMD method due to its inherent operation; in addition, the number of IMFs had to be the same for the three methods in order to keep the same reference/size. At the image representation, the six IMFs are used to make an image through a surface graph; thus, each 1 min segment generates three images, i.e., one per each EMD-based method. Finally, in the fourth stage, the obtained images are used to train and validate the CNN for the prediction of an SCD event. In order to provide a suitable solution in terms of computational complexity and accuracy, different configurations of the CNN by including the input image size, number of filters, size of filters, number of epochs, and batch size were constructed and tested. In order to do so, a simple but complete CNN structure with random parameters is first used for the SCD prediction; then, each parameter was tested individually by using different values. The parameter with the best results is selected and considered for the next tests. This procedure is repeated up to reviewing all the previously mentioned parameters. In the end, the resulting CNN structure becomes an efficient structure in terms of accuracy

and complexity. The equipment where these tests were carried out has the following hardware specifications: CPU @ 2.30 GHz, 16 GB RAM, and 64 bits operating system. The implementation software for the overall methodology is Matlab version 2022a.
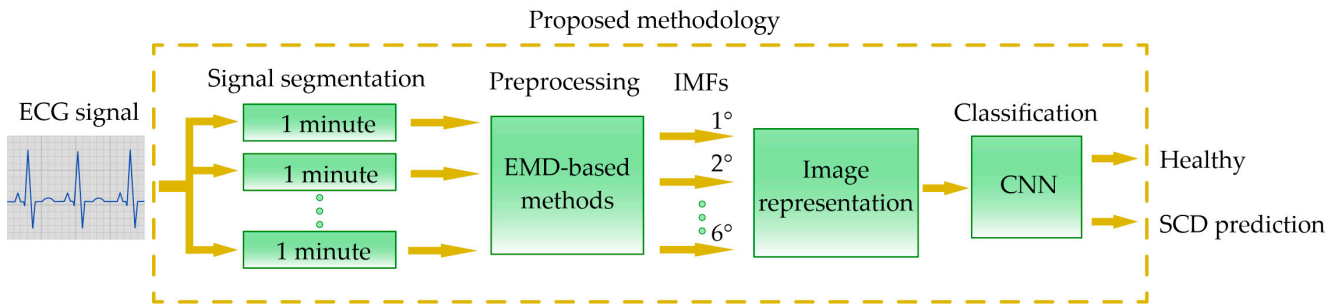


**Figure 5.** Proposed methodology for predicting an SCD event.

## 4. Experimentation and Results

### 4.1. EMD-Based Method

As previously described in the methodology, 1 min time windows of both databases are processed by the EMD, EEMD, and CEEMD algorithms by obtaining 6 IMFs. Figure 5 shows an example of the IMFs obtained for a healthy ECG signal (Figure 6a) and for an ECG signal 1 min before the SCD event (Figure 6b). The three methods are used because they decompose the signal in a slightly different way, allowing obtaining different characteristics in each set of IMFs. In fact, different data can be observed in the results of each method (columns of Figure 6), but no significant changes can be easily determined between both groups per method (rows of Figure 6), requiring the application of an automatic pattern recognition algorithm.
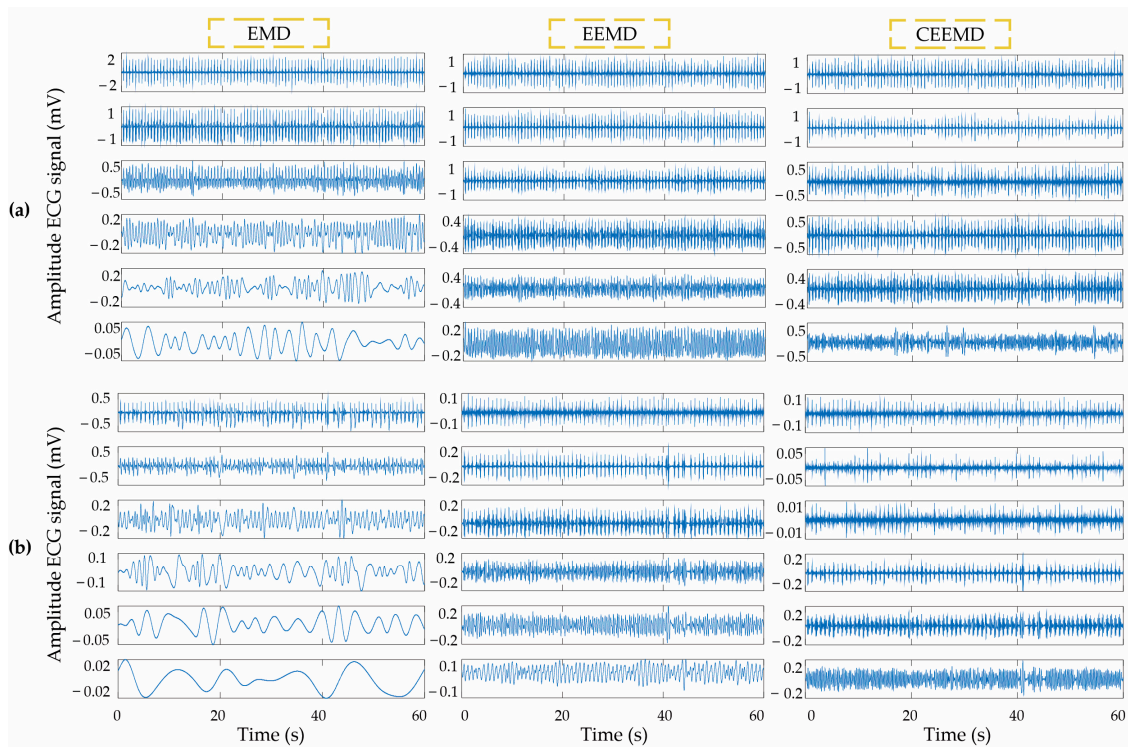


**Figure 6.** IMFs from the different EMD methods for: (**a**) 1 min segment of a healthy patient and (**b**) 1 min before the SCD event.

Following the proposed methodology, the image representation of each group of 6 IMFs is obtained by using the top view of their surface graph. This step makes that the

data generated from the EMD methods can be processed in the CNN and do not lose the spatial correlation between each IMF. Figure 7 shows the image representation for the IMFs depicted in Figure 6, obtaining the same results, i.e., the obtained differences of each method are visually insignificant (columns of Figure 7); therefore, no significant patterns can be easily determined between both groups per method (rows of Figure 7) through a direct visual inspection, requiring the application of an image-based pattern recognition algorithm. In this work, a CNN for automatic classification is proposed.
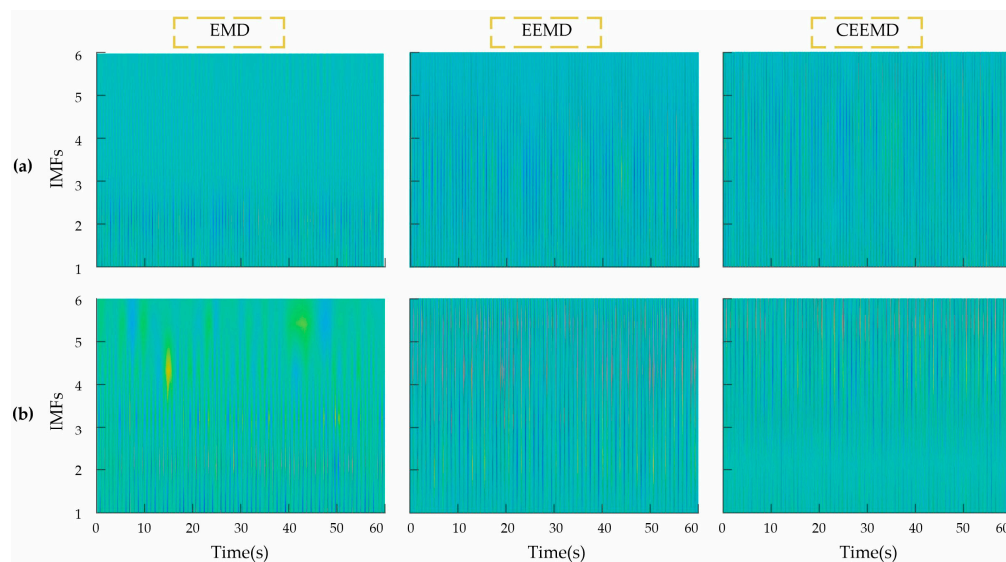


**Figure 7.** Top view of the surface graphics for the set of IMFs of a (**a**) 1 min segment of a healthy patient and (**b**) 1 min before the SCD event.

### 4.2. CNN Performance

To obtain a CNN structure with low computation resources but high accuracy, different configurations were tested. As the first test, the CNN performance for different sizes of images obtained by the 3 methods, i.e., EMD, EEMD, and CEEMD, is compared. The original image size is $218 \times 162$ pixels, where 3 smaller sizes (i.e., $164 \times 122$, $131 \times 98$, and $55 \times 41$ pixels) are analyzed to find the best relationship between accuracy and computational load based on the size of the images. The smaller the image size, the lower the number of mathematical operations required in the CNN. The first row of Figure 8 shows the images depicted in Figure 7. The left side of Figure 8 shows the sizes of the images that were used for the comparison. The image size represents the amount of data that enters a CNN; therefore, a reduction in its size reduces the computational load, but it can also reduce the accuracy of the CNN if most of the information is lost during the compression process. In Figure 8, all the images were enlarged to have the same visual size and observe how the patterns are somehow smoothed due to the compression process.

To perform the comparative equitably, the same CNN structure has to be used. In the beginning, the proposed CNN contains a five-filter convolution layer with a size of $5 \times 5$ and a pooling layer with a $2 \times 2$-size filter as shown in Figure 9. Table 2 contains detailed information on the CNN used. Although the previously mentioned values could be randomly selected, they are established according to the experience of the authors of this work. The results for other values that improve the CNN performance will be shown later.

Once the first CNN was configured, its training and validation were done with each of the methods and each of the image sizes shown in Figure 8. It is worth noting that the dataset consists of 540 healthy images (18 patients $\times$ 30 windows of 1 min) and 600 images with a future SCD condition (20 patients $\times$ 30 windows of 1 min), where 432 images from each condition (i.e., 80% and 72% of data, respectively) are used for training and the remaining images (i.e., 20% and 27% of data, respectively) are used for validation.
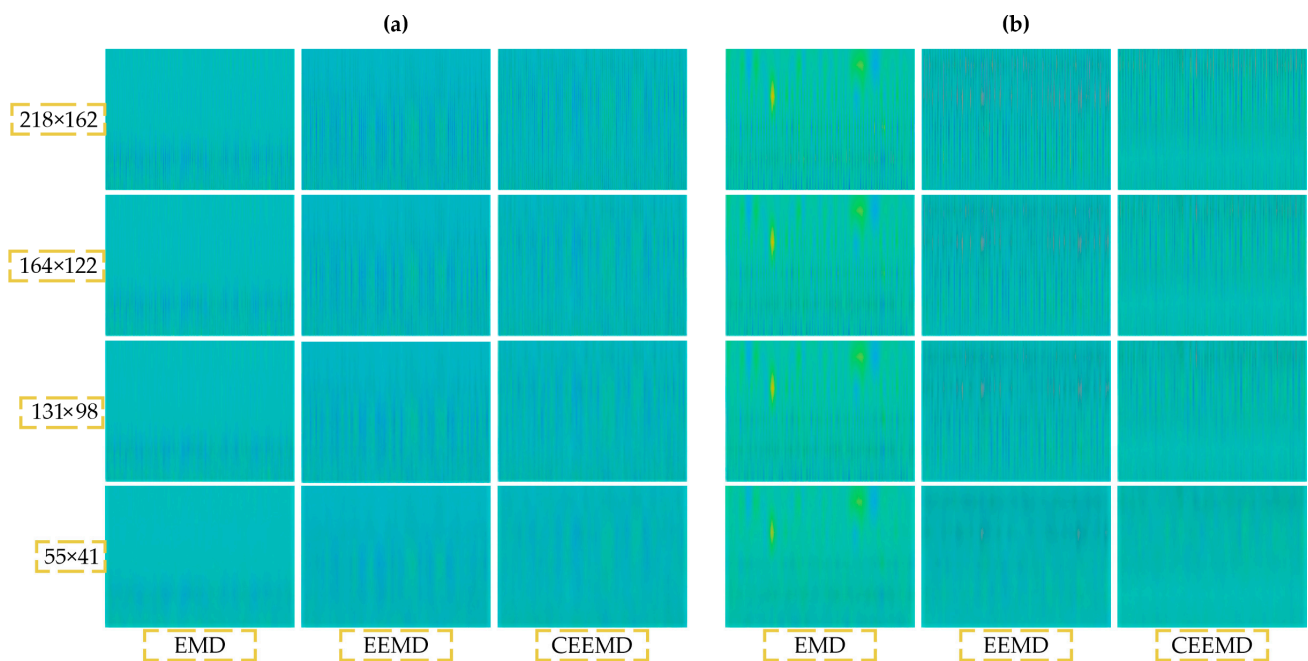
**(a)** **(b)**

218×162

164×122

131×98

55×41

EMD · EEMD · CEEMD · EMD · EEMD · CEEMD

**Figure 8.** Image size comparison for (**a**) images of a healthy patient and (**b**) images 1 min before the SCD.

218×162×3 · 5×5×5 · 2×2×1 · 2 · 2

Image input

Convolution
+
Relu

Max Pooling

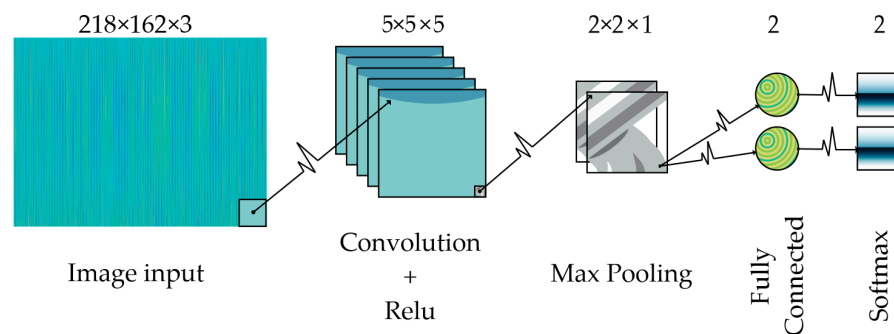Fully
Connected

Softmax

**Figure 9.** Base convolutional neural network (CNN) architecture.

**Table 2.** First CNN configuration.

| Name | Type | Activations | Learnable | Total Learnable |
|------|------|-------------|-----------|-----------------|
| imageinput | Image Input | $162 \times 218 \times 3$ | | 0 |
| conv | Convolution | $162 \times 218 \times 5$ | Weights $5 \times 5 \times 3 \times 5$ Bias $1 \times 1 \times 3$ | 380 |
| batchnorm | Batch Normalization | $162 \times 218 \times 5$ | Offset $1 \times 1 \times 5$ Scale $1 \times 1 \times 5$ | 10 |
| Relu | ReLU | $162 \times 218 \times 5$ | - | 0 |
| maxpool | Max Pooling | $161 \times 217 \times 5$ | - | 0 |
| Fc | Fully Connected | $1 \times 1 \times 2$ | Weights $2 \times 174685$ Bias $2 \times 1$ | 349, 372 |
| softmax | SoftMax | $1 \times 1 \times 2$ | - | 0 |
| classoutput | Classification Output | - | - | 0 |

Figure 10 shows the accuracy obtained from the CNN applied to each of the different image sizes and the preprocessing methods. This figure shows that the image with the largest size presents the best accuracy for the three methods, which is somehow expected because the analyzed images have the greatest amount of data. On the contrary, the worst

results are obtained for a size of 55 × 41 pixels. Despite these results, it is worth noting that the three methods offer results higher than 85% of accuracy. As the goal of testing different CNN configurations is to obtain an efficient method in terms of accuracy and complexity, the second image size is selected, i.e., 164 × 122, because the difference in accuracy is not much, but the input image size is reduced by 25%.
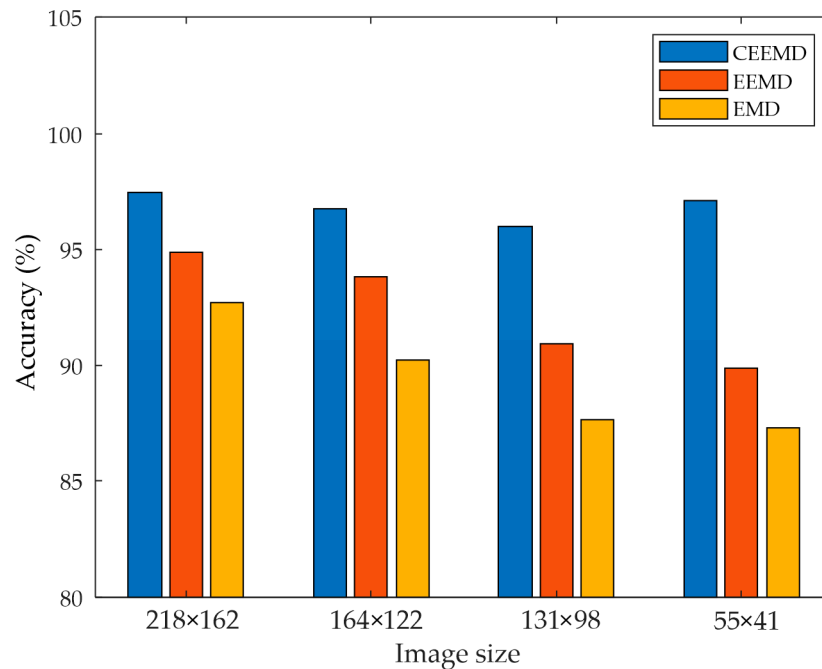


**Figure 10.** Accuracy results for different image sizes and different EMD methods.

Once the size of the image has been selected, and taking into account that the final accuracy of the CNN can slightly change between different runs of training, this procedure was carried out 30 times in order to have a statistical representation of the accuracy. Figure 11 shows the boxplot for the obtained accuracy of each EMD-based method, being the CEEMD method the one that provides the best results, i.e., a mean value of 97.5% of accuracy, in a consistent way. Therefore, the CEEMD method will be the only one used from now on.
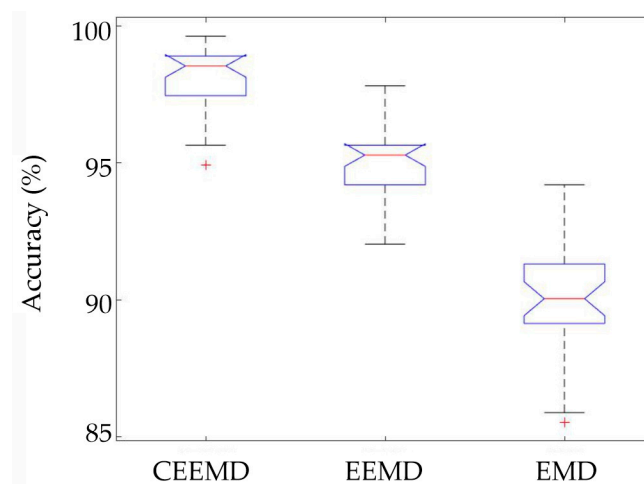


**Figure 11.** Boxplot for the accuracy obtained by different EMD methods.

The next parameters to be tested with different values are the number of filters and their size. Figure 12a shows the obtained results. For the number of filters, 3, 4, 5, 6, and 7

were tested, whereas sizes of $3 \times 3$, $5 \times 5$, $7 \times 7$, and $10 \times 10$ were analyzed. As can be observed, 6 filters of size $10 \times 10$ provide the highest accuracy value; yet, 5 filters of size $3 \times 3$ were chosen because they represent a lower computational load. Finally, the last parameters to be modified are the number of epochs and the batch size which are related with the number of iterations in the training stage. Figure 12b shows the obtained results for 5, 10, 15, . . . , and 40 epochs, as well as for 32, 64, 128, and 256 as batch sizes, showing that the best relationship between the epochs and the batch number is 10 epochs with a batch of 64.
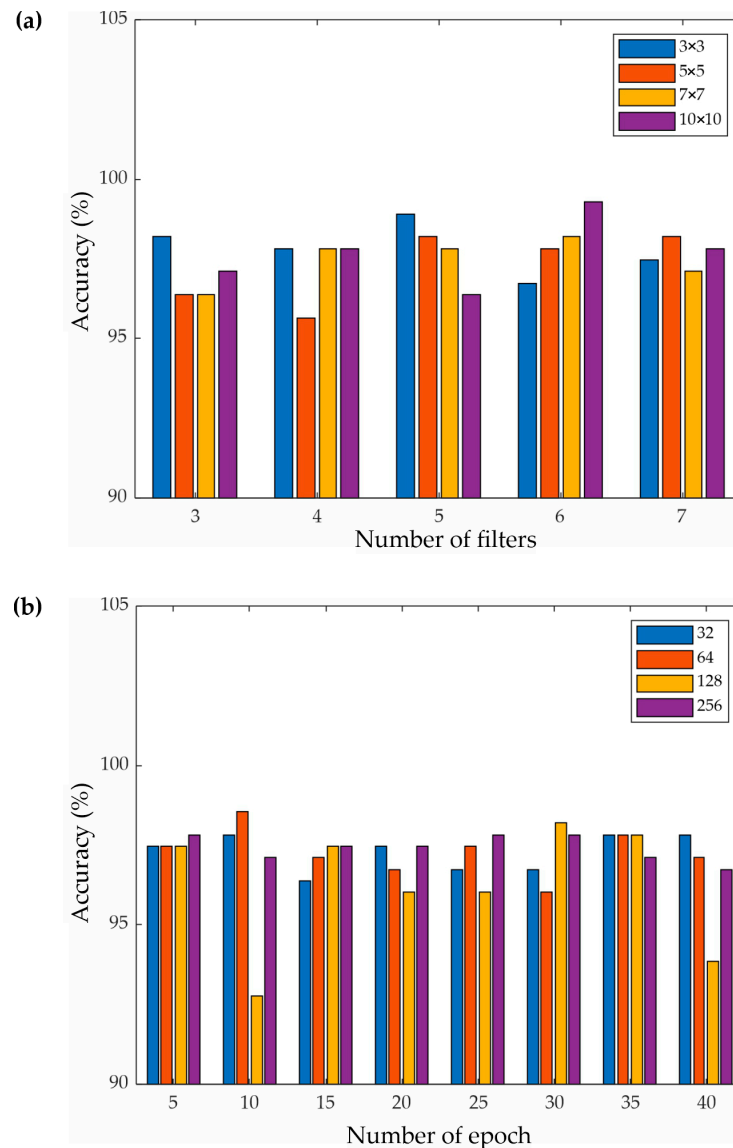
**(a)**

**(b)**

**Figure 12.** CNN accuracy for: (**a**) different filter configurations and (**b**) different iteration values.

After testing different values for the most common parameters of a CNN, the final proposed CNN has the following features: 5 filters of $3 \times 3$ with a batch number of 64 and 10 epochs, where the input image size is $164 \times 122$ and the CEEMD is the method for signal decomposition. Figure 13 shows the final proposed CNN, and Table 3 shows, in a more complete way, its characteristics. One value that can directly show the complexity reduction between the first CNN and the final CNN is the total learnable parameters of the fully connected layer. This value for the first CNN is 349,372 and 197,232 for the final CNN, representing a reduction of 43.55%. It is worth noting that this important reduction represented the main challenge of this work since many tests had to be carried out in order

to determine better characteristics for the CNN structure, always taking into account the computational load and the accuracy with respect to previous publications. This was also possible due to the more reliable results provided by the CEEMD method.
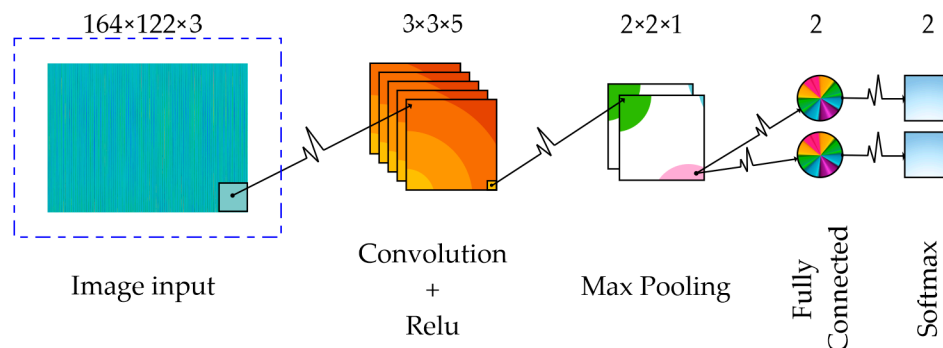


**Figure 13.** Final convolutional neural network (CNN) architecture.

After obtaining the final structure of the CNN, the training and validation were completed. In this regard, Figure 14 shows the obtained confusion matrix for the validation cases: 168 images for an SCD event and 108 for a healthy patient. As can be observed, 97.5%, 97.2%, and 98.2% are obtained for accuracy, precision, and recall, respectively, indicating the reliability of the proposal.



**Figure 14.** Confusion Matrix. The correct predictions are in the green fields. The incorrect predictions are in the red fields. The numbers in green are the positive rates for each row and column. The numbers in red are the negative rates for each row and column.

In order to assess the proposal effectiveness throughout the prediction time (i.e., 30 min), the accuracy of each minute was obtained. Figure 15 shows the accuracy obtained minute by minute, where it is observed that the values range from 100% to 97%. The dotted line indicates the average accuracy, and the black lines represent the variability obtained.

Figure 16 shows the obtained results for accuracy and loss, where it is observed that ~97% of accuracy is obtained during the fifth epoch.
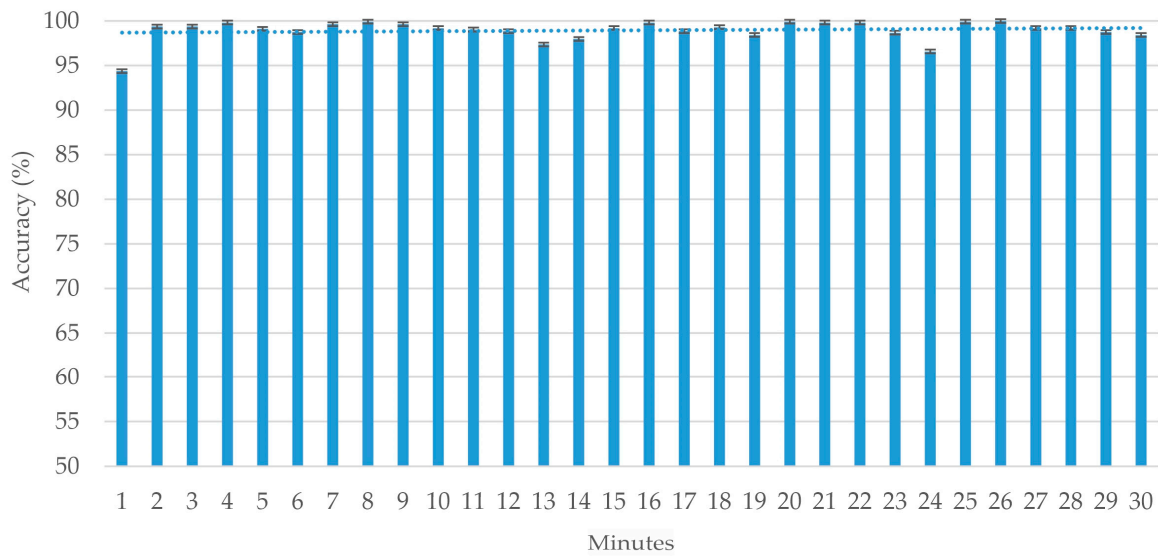
**Figure 15.** Obtained accuracy with the final CNN configuration for different minutes before the SCD. The dotted line indicates the mean value.
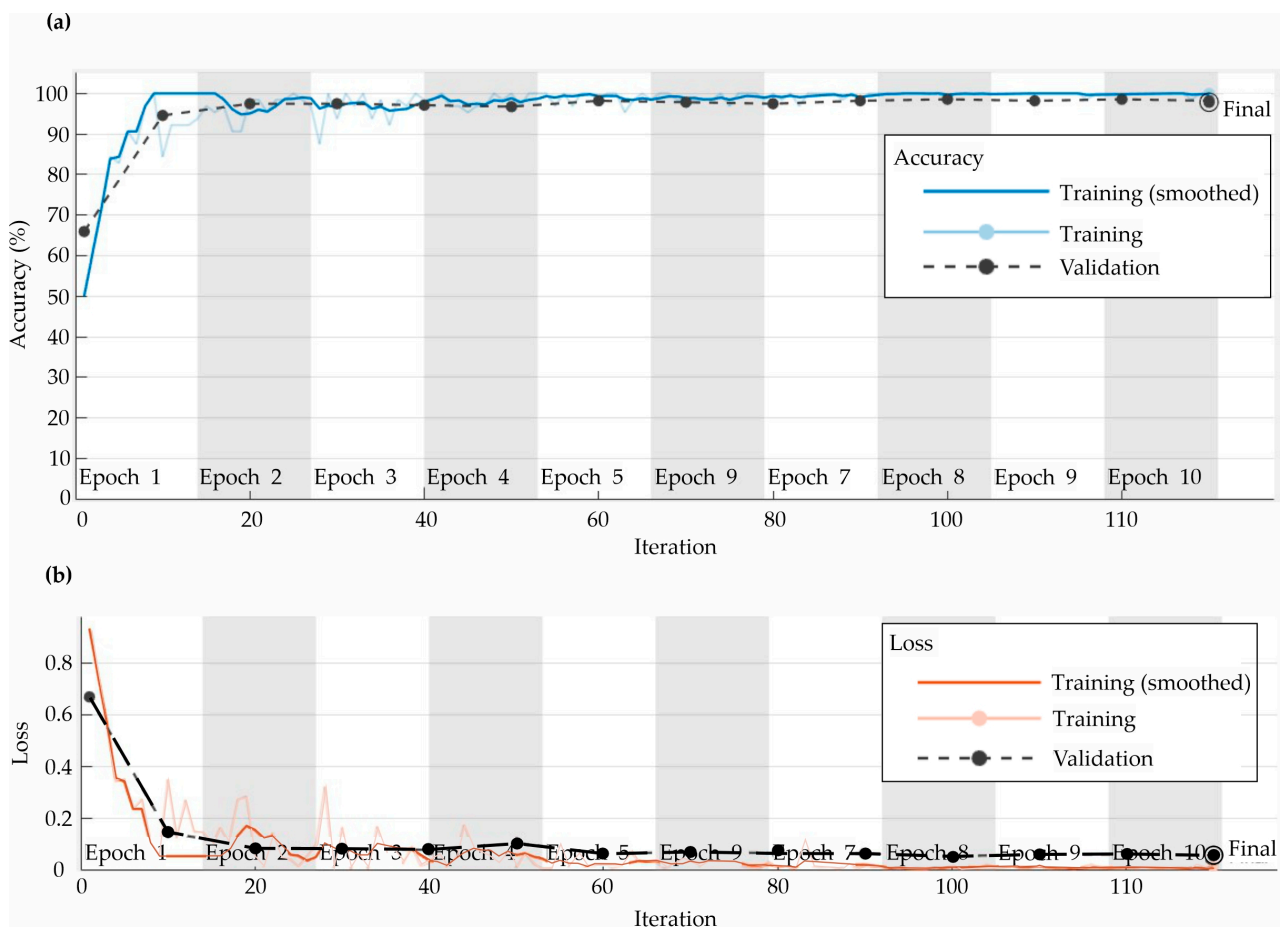
**Figure 16.** CNN training and validation: (**a**) Accuracy and (**b**) Loss.

**Table 3.** Final CNN configuration.

| Name | Type | Activations | Learnable | | Total Learnable |
|------|------|-------------|-----------|--|----------------|
| imageinput | Image Input | $122 \times 164 \times 3$ | - | | 0 |
| conv | Convolution | $122 \times 164 \times 5$ | Weights<br>Bias | $3 \times 3 \times 3 \times 5$<br>$1 \times 1 \times 3$ | 140 |
| batchnorm | Batch Normalization | $122 \times 164 \times 5$ | Offset<br>Scale | $1 \times 1 \times 5$<br>$1 \times 1 \times 5$ | 10 |
| relu | ReLU | $122 \times 164 \times 5$ | - | | 0 |
| maxpool | Max Pooling | $121 \times 163 \times 5$ | - | | 0 |
| fc | Fully Connected | $1 \times 1 \times 2$ | Weights<br>Bias | $2 \times 98{,}615$<br>$2 \times 1$ | 197,232 |
| softmax | SoftMax | $1 \times 1 \times 2$ | - | | 0 |
| classoutput | Classification Output | - | - | | 0 |

## 5. Discussion

Accuracy and prediction time have been the most important issues during the development of methods for SCD prediction. Table 4 shows a comparison between the proposed work and other works reported in the literature that use the same database. As qualitative features, the steps used in the method are discussed, while the prediction time and accuracy are discussed as quantitative parameters. It is worth noting that for the method description, the steps have been divided into two bullets; in general, the first one represents the signal processing and feature extraction, whereas the second one describes the algorithm used for automatic pattern recognition.

As can be observed in Table 4, most of the works use the raw ECG signal since its conversion to HRV adds an extra computational cost, e.g., in [15] the Pan–Tompkins algorithm is used to obtain the HRV signal. In the works that used ECG signals, all of them have higher values than 90% of accuracy; however, the prediction time changes greatly; for instance, the works [12,15] report values shorter than 10 min, whereas the remaining works report values longer than 20 min, excepting the work [17] that does not report the prediction time. It should be noticed that the works that provide longer times consider signal processing techniques such as wavelet transform-based methods or EMD-based methods, indicating that the use of these techniques contributes to highlighting features or patterns associated with an SCD condition. Going back to the prediction times, it is evident that the longer the prediction time, the better the medical treatment that a patient can receive. In this regard, the proposed work and the work presented in [18] become the best solutions because 30 min of prediction time are provided. Although similar results are obtained, some advantages of the proposed method can be highlighted: (1) the accuracy is slightly higher, (2) during the design stage, several parameters in the wavelet transform (such as the mother wavelet and the decomposition level) have to be tested and selected to decompose the signal in order to obtain more suitable results, unlike the CEEMD that performs an empirical decomposition without the need of tunning parameters, (3) the proposed CNN is less complex, e.g., the proposal has one convolutional layer whereas the CNN presented in [18] has four convolutional layers, impacting seriously in the computational cost. In addition, in the proposal, different settings in the CCN structure are tested in order to select which ones provide more suitable results by establishing a balance between computational cost and accuracy. As a result, 97.5% of accuracy until 30 min before the event was obtained, where a computational time of 3.2 s is required. This computational time allows the real time since the analyzed time windows are of 1 min.

Although promising results are obtained, they are preliminary due to the small database. In this regard, further research with larger databases that even include other rhythm abnormalities must be carried out.

**Table 4.** A quantitative and qualitative comparison between the proposal and similar works.

| Work | Signal | Methods | Prediction Time/ Accuracy |
|---|---|---|---|
| Acharya et al. (2015) [12] | ECG | • 18 nonlinear features<br>• Support vector machine | 92.11%/4 min |
| Khazaei et al. (2018) [15] | HRV | • Pan-Tompkins and two nonlinear features<br>• Decision tree | 95%/6 min |
| Amezquita-Sanchez et al. (2018) [14] | ECG | • Wavelet packet transform and homogeneity index<br>• Enhanced probabilistic neural network. | 95.8%/20 min |
| Olivia-Vargas et al. (2020) [16] | ECG | • Empirical mode decomposition and Higuchi fractal with permutation entropy<br>• Multilayer perceptron network | 94%/25 min |
| Kaspal et al. (2021) [17] | ECG | • Recurrence complex network<br>• CNN | 90.6%/− |
| Saragih et al. (2022) [18] | ECG | • Wavelet packet transform<br>• CNN | 95.89%/30 min |
| Proposed work | ECG | • CEEMD<br>• CNN | 97.1%/30 min |

− Not reported.

## 6. Conclusions

At first sight, ECG records of a healthy person and a person with a risk of suffering an SCD event appear identical. Hence, it is not possible to detect significant differences between both registries with simple and direct comparisons. In this regard, it is imperative to develop a technique that can detect reliable changes in the ECG signal in order to associate them with a person with normal cardiac rhythm and a person at risk of presenting an SCD event. In this research, a methodology based on EMD algorithms and CNNs was designed to predict 30 min prior to an SCD event by using the ECG signals provided by the MIT/BIH-SCDH and MIT/BIH-NSR databases. The obtained results show that the combination of the CEEMD algorithm with a CNN provides the best predictions of an SCD event, reaching 97.5% of accuracy until 30 min before the event. Although the settings of the proposed methods are not strictly optimized since an optimization algorithm is not used, different parameters were tested in order to provide a suitable solution in terms of computational complexity and accuracy. Thus, the first six IMFs in the signal decomposition stage, a size of $164 \times 122$ pixels for their image representation, and a CNN with the following structure: one convolutional and max pooling layers with five filters of

$3 \times 3$ and training with a batch number of 64 and 10 epochs, were enough to obtain the previously mentioned results.

As the obtained results are considered preliminary due to the limited databases, in future works, the proposal must be tested and calibrated with larger databases that include more patients and other rhythm abnormalities. In addition, it is proposed to carry out a hardware implementation based on a field-programmable gate array (FPGA) in order to exploit its parallelism to perform the IMFs' extraction and the CNN-based prediction, thus providing a portable solution that contributes to the prediction of SCD from a technological point of view.

**Author Contributions:** Conceptualization, M.A.C.-B., J.P.A.-S and M.V.-R.; methodology, M.A.C.-B., A.H.R.-R. and M.V.-R.; software, M.A.C.-B. and A.H.R.-R.; formal analysis, resources, and data curation, M.A.C.-B., A.H.R.-R. and A.V.P.-S.; writing—review and editing, all authors; supervision, project administration, and funding acquisition, J.P.A.-S., D.G.-L. and M.V.-R. All authors have read and agreed to the published version of the manuscript.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are not publicly available due to privacy issues.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

| | |
|---|---|
| BIDMC | Beth Israel Deaconess Medical Center |
| CEEMD | Complete Ensemble empirical mode decomposition |
| CNN | Convolutional neural network |
| ECG | Electrocardiogram |
| EEMD | Ensemble empirical mode decomposition |
| EMD | Empirical mode decomposition |
| FPGA | Field programmable gate array |
| HRV | Heart rate variability |
| IMF | Intrinsic mode function |
| MIT | Massachusetts Institute of Technology |
| NSR | Normal sinus rhythm |
| SCD | Sudden cardiac death |
| SCDH | Sudden cardiac death holter |
| VF | Ventricular fibrillation |

## References

1. Kelly, K.L.; Lin, P.T.; Basso, C.; Bois, M.; Buja, L.M.; Cohle, S.D.; d'Amati, G.; Duncanson, E.; Fallon, J.T.; Firchau, D.; et al. Sudden cardiac death in the young: A consensus statement on recommended practices for cardiac examination by pathologists from the Society for Cardiovascular Pathology. *Cardiovasc. Pathol.* **2023**, *63*, 107497. [CrossRef] [PubMed]
2. Srinivasan, N.T.; Schilling, R.J. Sudden Cardiac Death and Arrhythmias. *Arrhythm. Electrophysiol. Rev.* **2018**, *7*, 111–117. [CrossRef] [PubMed]
3. Khan, A.H.; Hussain, M.; Malik, M.K. Arrhythmia Classification Techniques Using Deep Neural Network. *Complexity* **2021**, *2021*, 9919588. [CrossRef]
4. Shilla, W.; Wang, X. Wavelet Transform and Convolutional Neural Network Based Techniques in Combating Sudden Cardiac Death. *Emit. Int. J. Eng. Technol.* **2021**, *9*, 377–389. [CrossRef]
5. Pagidipati, N.J.; Gaziano, T.A. Estimating deaths from cardiovascular disease: A review of global methodologies of mortality measurement. *Circulation* **2013**, *127*, 749–756. [CrossRef]

6. Moore, B.; Semsarian, C.; Chan, K.H.; Sy, R.W. Sudden Cardiac Death and Ventricular Arrhythmias in Hypertrophic Cardiomyopathy. *Heart Lung Circ.* **2019**, *28*, 146–154. [CrossRef]

7. Jazayeri, M.A.; Emert, M.P. Sudden Cardiac Death: Who Is at Risk? *Med. Clin.* **2019**, *103*, 913–930.

8. Myerburg, R.J. Cardiac arrest and sudden cardiac death. In *Heart Disease, a Textbook of Cardiovascular Medicine*; W.B. Saunders: Philadelphia, PA, USA, 1992; pp. 756–789.

9. Tseng, L.M.; Tseng, V.S. Predicting Ventricular Fibrillation through Deep Learning. *IEEE Access* **2020**, *8*, 221886–221896. [CrossRef]

10. Sumner, G.L.; Kuriachan, V.P.; Mitchell, L.B. Sudden Cardiac Death. *Encycl. Cardiovasc. Res. Med.* **2018**, *8*, 511–520.

11. Nash, M.P.; Mourad, A.; Clayton, R.H.; Sutton, P.M.; Bradley, C.P.; Hayward, M.; Paterson, D.J.; Taggart, P. Evidence for multiple mechanisms in human ventricular fibrillation. *Circulation* **2006**, *114*, 536–542. [CrossRef]

12. Acharya, U.R.; Fujita, H.; Sudarshan, V.K.; Sree, V.S.; Eugene, L.W.J.; Ghista, D.N.; Tan, R.S. An integrated index for detection of Sudden Cardiac Death using Discrete Wavelet Transform and nonlinear features. *Knowl. Based Syst.* **2015**, *83*, 149–158. [CrossRef]

13. MIT/BIH-SCDH. Available online: https://physionet.org/physiobank/database/sddb/#clinical-information/databased (accessed on 14 February 2023).

14. Amezquita-Sanchez, J.P.; Valtierra-Rodriguez, M.; Adeli, H.; Perez-Ramirez, C.A. A Novel Wavelet Transform-Homogeneity Model for Sudden Cardiac Death Prediction Using ECG Signals. *J. Med. Syst.* **2018**, *42*, 176. [CrossRef] [PubMed]

15. Khazaei, M.; Raeisi, K.; Goshvarpour, A.; Ahmadzadeh, M. Early detection of sudden cardiac death using nonlinear analysis of heart rate variability. *Biocybern. Biomed. Eng.* **2018**, *38*, 931–940. [CrossRef]

16. Vargas-Lopez, O.; Amezquita-Sanchez, J.P.; De-Santiago-Perez, J.J.; Rivera-Guillen, J.R.; Valtierra-Rodriguez, M.; Toledano-Ayala, M.; Perez-Ramirez, C.A. A new methodology based on EMD and nonlinear measurements for sudden cardiac death detection. *Sensors* **2020**, *20*, 9. [CrossRef]

17. Kaspal, R.; Alsadoon, A.; Prasad, P.W.C.; Al-Saiyd, N.A.; Nguyen, T.Q.V.; Pham, D.T.H. A novel approach for early prediction of sudden cardiac death (SCD) using hybrid deep learning. *Multimed. Tools Appl.* **2021**, *80*, 8063–8090. [CrossRef]

18. Saragih, Y.V.; Isa, S.M. CNN Performance Improvement Using Wavelet Packet Transform for SCA Prediction. *J. Theor. Appl. Inf. Technol.* **2022**, *100*, 5458–5468.

19. MIT/BIH-NSR. Database. Available online: https://www.physionet.org/physiobank/database/nsrdb/ (accessed on 14 February 2023).

20. Chinara, S. Automatic classification methods for detecting drowsiness using wavelet packet transform extracted time-domain features from single-channel EEG signal. *J. Neurosci. Methods* **2020**, *347*, 108927.

21. Perez-Sanchez, A.V.; Valtierra-Rodriguez, M.; Perez-Ramirez, C.A.; De-Santiago-Perez, J.J.; Amezquita-Sanchez, J.P. Epileptic seizure prediction using Wavelet Transform, Fractal Dimension, Support Vector Machine, and EEG signals. *Fractals* **2022**, *30*, 2250154. [CrossRef]

22. Huang, N.E.; Shen, Z.; Long, S.R.; Wu, M.C.; Shih, H.H.; Zheng, Q.; Liu, H.H. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. Lond.* **1998**, *454*, 903–995. [CrossRef]

23. Wu, Z.; Huang, N.E. Ensemble empirical mode decomposition: A noise-assisted data analysis method. *Adv. Adapt. Data Anal.* **2009**, *1*, 1–41. [CrossRef]

24. Torres, M.E.; Colominas, M.A.; Schlotthauer, G.; Flandrin, P. A Complete Ensemble Empirical Mode Decomposition with Adaptive Noise. In Proceedings of the 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Prague, Czech Republic, 22–27 May 2011.

25. Liu, T.; Xu, H.; Ragulskis, M.; Cao, M.; Ostachowicz, W. A Data-Driven Damage Identification Framework Based on Transmissibility Function Datasets and One-Dimensional Convolutional Neural Networks: Verification on a Structural Health Monitoring Benchmark Structure. *Sensors* **2020**, *20*, 1059. [CrossRef] [PubMed]

26. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet Classification with Deep Convolutional Neural Networks. *Commun. ACM* **2017**, *60*, 84–90. [CrossRef]

27. Ieracitano, C.; Mammone, N.; Bramanti, A.; Hussain, A.; Morabito, F.C. A Convolutional Neural Network Approach for Classification of Dementia Stages Based on 2D-Spectral Representation of EEG Recordings. *Neurocomputing* **2019**, *323*, 96–107. [CrossRef]

28. Mammone, N.; Ieracitano, C.; Morabito, F.C. A Deep CNN Approach to Decode Motor Preparation of Upper Limbs from Time–Frequency Maps of EEG Signals at Source Level. *Neural Netw.* **2020**, *124*, 357–372. [CrossRef]

29. Wang, L.H.; Zhao, X.P.; Wu, J.X.; Xie, Y.Y.; Zhang, Y.H. Motor Fault Diagnosis Based on Short-Time Fourier Transform and Convolutional Neural Network. *Chin. J. Mech. Eng. Engl. Ed.* **2017**, *30*, 1357–1368. [CrossRef]

30. Scherer, D.; Müller, A.; Behnke, S. Evaluation of Pooling Operations in Convolutional Architectures for Object Recognition. In Proceedings of the Artificial Neural Networks–ICANN 2010, Thessaloniki, Greece, 15–18 September 2010; pp. 92–101.