IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 22 February 2024, accepted 11 March 2024, date of publication 18 March 2024, date of current version 22 March 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3378730

RESEARCH ARTICLE

Approach Based Lightweight Custom Convolutional Neural Network and Fine-Tuned MobileNet-V2 for ECG Arrhythmia Signals Classification

HADJER BECHINIA[1], DJAMEL BENMERZOUG[1], AND NAWRES KHLIFA[2]
[1]LIRE Laboratory, Faculty of New Technologies of Information and Communication, University of Abdelhamid Mehri Constantine 2, Constantine 25016, Algeria
[2]Research Laboratory of Biophysics and Medical Technologies, Higher Institute of Medical Technologies of Tunis, University of Tunis El Manar, Tunis 1006, Tunisia

Corresponding author: Hadjer Bechinia (hadjer.bechinia@univ-constantine2.dz)

ABSTRACT Arrhythmia detection in electrocardiogram (ECG) signals is a vital aspect of cardiovascular health monitoring. Current automated methods for arrhythmia classification often struggle to attain satisfactory performance in the detection of various heart conditions, particularly when dealing with imbalanced datasets. This study introduces a novel deep learning approach for the detection and classification of ECG arrhythmia plot images. Our methodology features a Lightweight Custom Convolutional Neural Network model(LC-CNN), comprising just three convolutional layers and a transfer learning model with MobileNet-V2 architecture that leverages pre-trained features to enhance arrhythmia classification. Data preprocessing of the ECG signals involving noise reduction with a Butterworth filter and precise beat segmentation via R-peak detection, ensure high-quality input for our model. Furthermore, a notable contribution for ECG data augmentation, adopting the implementation of an Auxiliary Classifier Generative Adversarial Network (ACGAN), specifically addressing class imbalance in the benchmark MIT-BIH dataset to classify four types of ECG heartbeats. This approach enriches the dataset, enhancing the models' ability to detect underrepresented arrhythmia classes. The proposed system demonstrates an impressive average classification accuracy achieving 99.22% using the LC-CNN model, closely followed by the fine-tuned MobileNet-V2 model with 98.69% accuracy, outperforming other methods and underscoring its effectiveness when faced with diverse irregular heartbeats and arrhythmia.

INDEX TERMS ECG, arrhythmia classification, deep learning, convolution neural network (CNN), transfer learning, MobileNet-V2, data augmentation, auxiliary classifier GAN, computer vision.

I. INTRODUCTION

Cardiovascular disease (CVD) represents a disorder affecting the heart and blood vessels, causing clots that lead to stroke and heart attacks. As per data provided by the American Heart Association in 2019, CVDs have emerged as a predominant global contributor to mortality. In 2016, they were responsible for more than 17.6 million deaths, and projections indicate that this alarming figure is expected to escalate to 23.6 million by the year 2030 [1]. Arrhythmia ranks as the most prevalent type among all CVDs, and it is defined as irregularities in heart rhythm, stemming from abnormal electrical impulses and conduction within the heart. Arrhythmias can manifest as irregular heartbeats (heart tremors), tachycardia (a resting heart rate exceeding 100 beats per minute), or bradycardia (a resting heart rate below 60 beats per minute) [2]. The most commonly employed method for detecting arrhythmias consists of recording an electrocardiogram (ECG), which

The associate editor coordinating the review of this manuscript and approving it for publication was Amin Zehtabian.

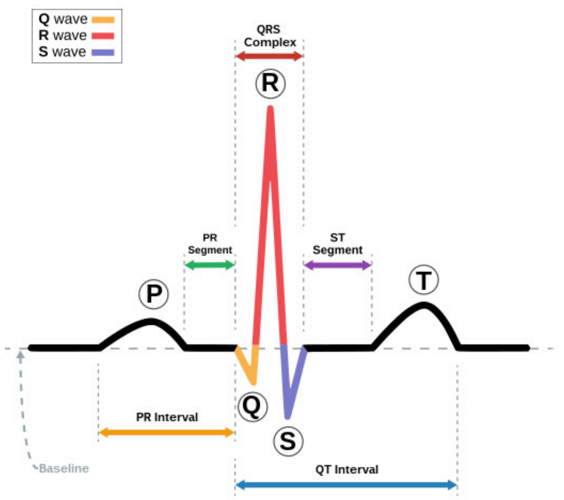

FIGURE 1. The normal ECG Beat.

provides a visual representation of the heart's electrical activity over time through skin-placed electrodes. These ECG leads, capturing the heart's electrical potential from various angles and positions, serve as indicators of disease states by identifying anomalies in waveforms and rhythms. The ECG essentially serves as a comprehensive record of the electrical attributes of the heartbeat [3]. In an ECG signal, each individual heartbeat manifests with a specific waveform. As depicted in Figure 1, we observe an ECG heartbeat featuring its key identifying points(P, Q, R, S, and T), along with the segments and the duration intervals (PR interval, PR segment, QT interval, ST segment). Each of these forms corresponds to a specific phase within the cardiac cycle. Of particular significance, the QRS complex represents the paramount cardiac activity, namely, the ventricular depolarization process. This pivotal occurrence is reflected in the ECG waveform by the most pronounced potential disparities, collectively forming the R-peak. The precise localization of the R-peak holds notable importance, as it functions as a reference point for determining the temporal alignment of a heartbeat [4]. CVDs through the continuous monitoring and analysis of ECG signals has been illuminated. Especially, due to the advent of intelligent enabling technologies, ECG monitoring systems have been developed and have gained extensive utilization within the healthcare domain over the recent decades. These systems harness a diverse array of technologies, encompassing IoT, edge computing, and mobile computing. Furthermore, their functionality has expanded beyond disease diagnosis and management, engaging areas such as monitoring daily activities, enhancing athletic performance, and fulfilling specific mode-related needs [5]. Hence, a large amount of literature research has emerged, presenting various methods and approaches for classifying cardiac arrhythmias utilizing machine learning and deep learning techniques. In particular, Deep learning, as a computer-aided approach known for its robust feature extraction capabilities, has demonstrated remarkable accuracy in the classification of ECG signals. This is achieved through the construction of hierarchical artificial neural networks, wherein the non-linear components in each layer empower deep learning to efficiently process complex non-linear signals like ECG signals. As information traverses through each layer, it becomes progressively more abstract and high-level, which greatly contributes to achieving high classification accuracy. Consequently, when compared to traditional machine learning methods, deep learning excels in its capacity for learning intricate patterns from extensive datasets [1]. Attained that inference, publicly accessible ECG databases encounter challenges related to both data availability and data imbalance [6]. A remarkable difference exists in the prevalence of arrhythmias signals compared to the normal ones in patient records of these datasets [9]. In particular, the MIT-BIH dataset, widely recognized ECG signals database, accessible on the physionet platform [7] and employed in this study, exhibits a significant class imbalance problem when examining its arrhythmia diagnostic categories. Therefore, achieving high performance in identifying the minority cardiac abnormality classes such as the Supraventricular(S) and the Ventricular (V) beats can principally be challenging. On that account, the major contributions of this study can be summarized as follows:

1) We propose two deep learning models based on Convolutional Neural Networks, a lightweight custom CNN model(LC-CNN) with only three convolutional layers and the pre-trained MobileNet-V2 model, fine-tuned to detect and to classify ECG arrhythmias.
2) We perform signal denoising using butterworth Filter and we fulfill the beats segmentation of the ECG signals via R-peak values using the slide window technique.
3) Following that, each segmented beat is plotted and saved as an image in a PNG format.
4) Furthermore, we propose an ACGAN model to tackle the problem of data imbalance and generate synthetic ECG images.
5) We perform intensive experimentation using the MIT-BIH for arrhythmia detection and calculate various performance metrics. The obtained results are also compared with several recently proposed models and approaches for arrhythmia detection for ECG images.

The remaining parts of the paper are organized as follows: In section two, we review the recent related works to our study. In section three, we provide a detailed explanation of the proposed approach, the dataset characteristics, the preprocessing steps, and a detailed explanation of the proposed ACGAN model for data augmentation. Section four presents the experimental results while section five provides a comparison with the related approaches followed by the discussion of these results in section six. Finally, section seven presents the conclusions and suggests possible future works.

II. LITERATURE REVIEW

In this section, we present a comprehensive review of recent related literature to the detection and classification of ECG arrhythmia, exploring their utilization of deep learning models and various data augmentation techniques employed to tackle the data imbalance problem in ECG signals datasets. In the first study, [2]the authors introduced a novel approach based on Deep Convolutional Generative Adversarial Networks (ECG-DCGAN) to address the balancing of the ECG signals. This method involved the utilization of a 16-layer CNN model for ECG signal classification. The experimental outcomes revealed an accuracy of 98.7%. In [8], the researchers employed multiple models and classifiers for arrhythmia detection. They utilized the GAN model to address class imbalance and among the evaluated models, the proposed GAN-LSTM ensemble model emerged as the top performer, achieving the highest level of accuracy and F1 score at 0.992. Furthermore, for data imbalance, the authors in [9], proposed a novel approach using the Transformer and Convolution-Based Generative Adversarial Network (TCGAN) model, that demonstrated a robust capacity to generate synthetic ECG heartbeats. Additionally, they implemented a CNN-BI-LSTM model for the classification of ECG beats. The results showcased an overall accuracy of 94.69%. The authors in [10], introduced a novel data augmentation algorithm for the ECG signals. These signals, represented as Numpy arrays, were first divided into segments of identical length based on the input, then, rearranged to create new signals. The resulting signals were converted into a.JPEG image format, then used as input for a four-layer CNN for the classification of ECG signals achieving a validation accuracy of 89.87%. Moreover, in [11], a novel approach was introduced, involving a fusion of a 5-layer CNN and a single-block Transformer encoder, aimed at mitigating the imbalance issue associated with the minority classes within ECG signals. This combined model achieved an average accuracy of 97.66% when evaluated on the MIT-BIH dataset. The performance was enhanced as the CNN block was substituted with a pre-trained DAE network, leveraging an additional ECG dataset with distinct attributes, this modification resulted in a higher performance, yielding an average accuracy of 97.93%. in [12]various standard methods and operations were applied to augment the ECG images. Thus, their introduced system fine-tuned the weights of the pre-trained model, Dense-Net to classify 29 types of heartbeats achieving a classification accuracy of 98.92%. In [13], a novel approach called 'Fuzz-Clust-Net' was proposed for arrhythmia detection of ECG signals, The process followed different standard data augmentation techniques to mitigate class imbalance. After that, the CNN model was used to extract the features that next were subjected to a fuzzy clustering algorithm for the classification. The experimental results demonstrated an overall accuracy of 98.66%. The study in [14], introduced a predictive system, employing a GAN model to identify arrhythmia in young martial arts athletes. Notably, the experimental outcomes showed that the proposed model achieved the highest accuracy, reaching an impressive 97%. Moreover, [15], a novel system for ECG classification was introduced, consisting of a two-part methodology. Initially, a deep autoencoder was utilized to extract high-level features. Subsequently, multiple neural networks were employed, following both one-against-all (OAA) and one-against-one (OAO) strategies for multi-class classification. To tackle data imbalance, additional oversampling using SMOTE was incorporated. Notably, the model employing the OAA-MLP approach demonstrated the most promising performance, achieving an accuracy rate of 99.32%. In the research conducted at [16], two deep learning models, CNN and CNN-LSTM, were presented alongside ensemble techniques for the classification of two categories of ECG beats. Remarkably, these models achieved an overall accuracy of 99.9% when tested on different ECG beat datasets. To mitigate class imbalance, the study employed SMOTE and Tomek link resampling techniques, to further enhance the robustness of the classification results. In [17], an innovative data augmentation method utilizing a GAN model was introduced to enhance ECG classification. The research presented two deep learning methodologies: an end-to-end and a two-stage hierarchical approach, both based CNN, highlighting the feature learning capacity of deep CNNs without the need for feature engineering. The outcomes demonstrated that using the GAN-based augmentation substantially improved the performance of the method, resulting in an accuracy rate exceeding 98%. In [18], two deep learning models were proposed for arrhythmia classification, a CNN-LSTM model, which captures local features and temporal dynamics, and the RRHOS-LSTM model, integrating RR intervals and higher-order statistics to highlight abnormal heartbeats. A bagging model was trained on sub-sampled data to address class imbalance, utilizing weighted loss functions. Additionally, a meta-classifier combining these models was validated with the CNN-LSTM model, achieving 95.81% of accuracy. In [19], they introduced a novel CNN model, inspired by the Shuffle-Net architecture. This model was tailored for efficient deployment on resource-constrained wearable mobile devices. The implementation of a variable stride sliding window was used to address the under-represented classes within the dataset. Notably, the utilization of the Focal loss as a loss function contributed to enhanced performance, surpassing traditional CNN models while maintaining significantly fewer parameters, ultimately resulting in a 97% improvement in the F1 score. Last but not least, in [20], they proposed an optimized neural network model design based on Autoencoder and LSTM models, which enhances the accuracy of diagnosing heart diseases through ECG signals while simplifying the preprocessing process and addressing gradient vanishing issues, ultimately, acheiving an accuracy of 98.57%

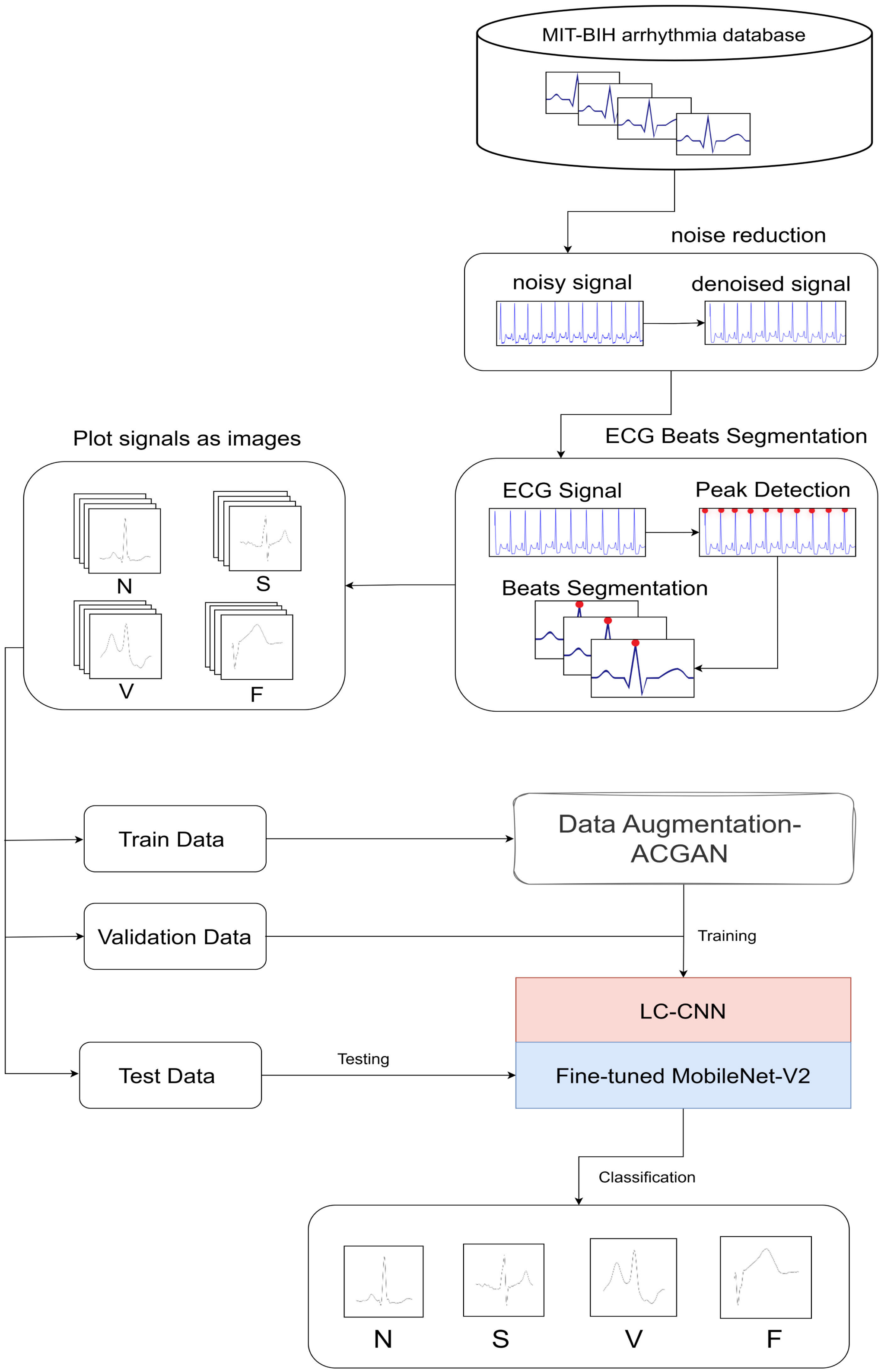

FIGURE 2. The proposed approach.

III. MATERIALS AND METHODS

A. PROPOSED APPROACH

In this section, we present our proposed approach based on deep learning for the detection and the classification of arrhythmia in ECG signals. Figure 2 shows an overall view of the proposed system. In the initial phase of our methodology, using the MIT-BIH dataset, the ECG signals undergo a series of essential preprocessing steps. First, a Butterworth filter is applied to the signals for the purpose of noise reduction and signal enhancement. Subsequently, the detection of the R peaks within the ECG signals is performed, which serves as the foundation for segmenting the signals into beats. Once the ECG beats are fully segmented, we chose to plot and save each ECG beat into a structured format as an image, each with its appropriate labeling. Next, To address class imbalance, particularly in the minority arrhythmia classes within the ECG image dataset, we employ an ACGAN (Auxiliary Classifier GAN) model that generates synthetic ECG images, focusing on the underrepresented arrhythmia classes. This augmentation step enriches our dataset and improves the model's ability to handle the classification of the minority classes effectively. Last but not least, after performing the data augmentation, the augmented dataset is fed into the two proposed CNN architectures, presenting the Lightweight Custom CNN model with only three convolutional layers along with a transfer learning pre-trained MobileNet-V2 model. These two models are designed for the classification task, categorizing ECG signals into four distinct beat classes: Normal (N), Supraventricular (S), Ventricular (V), and Fusion (F). The evaluation of the proposed approach's performance is assessed by several performance metrics such as the accuracy, the F1-score, precision, and recall.

B. THE DATASET

As previously indicated in our paper, our study involved conducting the experiments utilizing the MIT-BIH Arrhythmia Database. Comprising of 48 half-hour two-lead ECG recordings sampled at 360 Hz from 47 patients, this dataset offers a diverse set of subjects, with accompanying annotation files detailing beat-by-beat labels, including normal beats and various arrhythmias like premature ventricular contractions (V), supraventricular premature beat (S), among others. In the standard WFDB format, the dataset facilitates reproducibility and collaboration in the field. Table 1 provides an overview of the heartbeat categories defined by the Association for Advancement of Medical Instrumentation (AAMI), which classifies arrhythmia heartbeats into five distinct classes: 1) normal beat (N); 2) supraventricular ectopic beat (S); 3) ventricular ectopic beat (V); 4) fusion beat (F); and 5) unclassifiable beat (Q). In this work, we only focus on identifing the four classes of lead MLII from the MIT-BIH dataset, the N, S, V and F classes.

C. PREPROCESSING

One of the initial and commonly employed steps within the processing of the ECG signals is noise reduction. Noise

TABLE 1. Major types of heartbeats present in the MIT-BIH Arrhythmia Dataset.

class	description	Annotation
N	Normal beat	n
	right bundle branch	l
	Left bundle branch block beat	r
	Atrial escape beat	e
	Nodal (junctional) escape beat	j
S	Supraventricular premature beat	S
	Atrial premature beat	A
	Nodal junctional premature beat	J
	Aberrated atrial premature beat	a
V	Premature ventricular escape beat	V
	Ventricular escape beat	E
F	Fusion of ventricular and normal beat	f
Q	Unclassified beat	q
	Paced beat	p
	Fusion of paced and normal beat	f

in ECG signals can be categorized into two main types: low-frequency noise, which arises from baseline oscillations caused by body movements and respiration, and high-frequency noise, stemming from power line interference and the digitization of analog electrical potentials [4]. In our work, we are using the Butterworth Band Pass Filter, with a low-cut frequency of 0.5 Hz, a high-cut frequency of 50 Hz, a sampling frequency of 250 Hz, and a filter order of 4. This filter applies the bandpass filter both forward and backward within the signal. This dual-pass approach achieves a zero-phase filtered signal, ensuring that the temporal alignment of cardiac data remains intact while effectively eliminating noise and isolating relevant frequency components within the range of 0.5 Hz to 50 Hz.

Next, the beat segmentation process, the detection of peaks plays a pivotal role as key landmarks for segmenting the ECG signal into distinct beats. First, we implement a systematic approach that leverages peak detection and annotation filtering. We employ a 'find-peaks' function to identify significant peaks within the ECG signal while maintaining a reasonable spacing between them. Then, the segmentation step involves creating fixed windows around each peak, with an input size set at 256 samples, equivalent to 128 samples before and after the identified peak. Within these segments, we extract the annotations from the associated annotation file. Importantly, we apply class criteria to include only those segments that contain a single desired annotation, meeting our classification and analysis objectives. Following the segmentation of the ECG beats, the subsequent step entails generating grayscale representations of these individual beats, which are then saved as image files in PNG format. Each image is uniquely named based on its respective position in the signal. The decision to represent ECG beats as images serves the purpose of exploring and developing innovative methodologies in arrhythmia analysis. By harnessing the capabilities of computer vision, there is potential to enhance the accuracy and effectiveness of diagnostic processes. Figure 3 presents the outcomes of the preprocessing procedures.

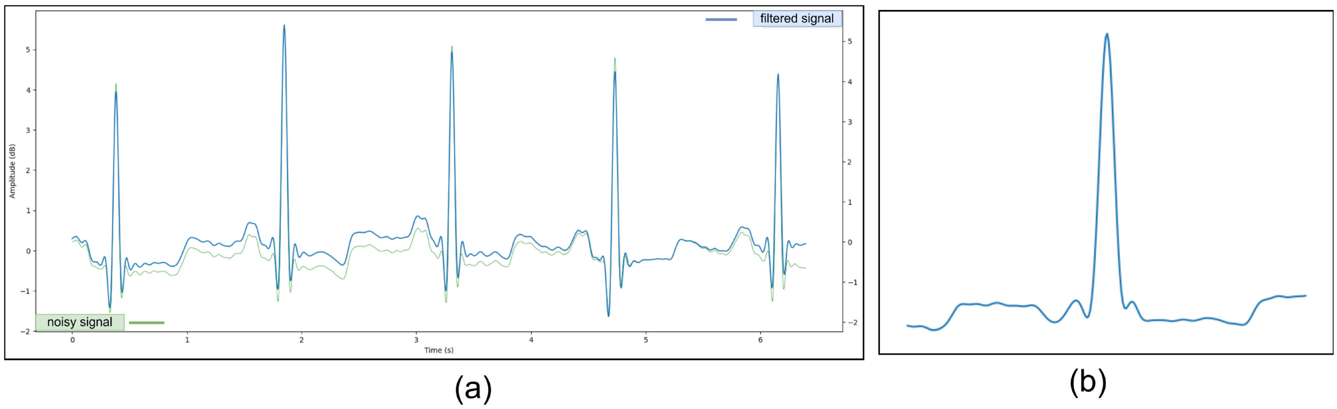

FIGURE 3. The preprocessing steps result.

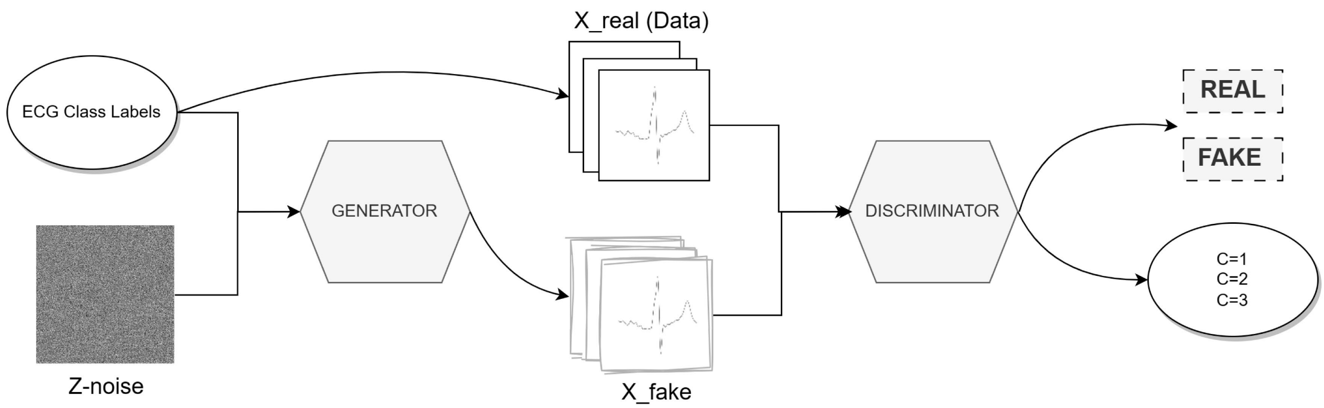

FIGURE 4. The ACGAN architecture.

D. DATA AUGMENTATION

To tackle the issue of data imbalance, we apply data augmentation on the arrhythmias' images representing the minority classes of the MIT-BIH dataset, the S, V and F classes. To do so, we implement an advanced variant of the Generative Adversarial Network (GAN) known as the Auxiliary Classifier GAN, introduced by [21]. The ACGAN model mainly consists of two neural networks, the generator and the discriminator, which are trained together in a competitive manner; the generator's primary role involves creating synthetic ECG heartbeat images from random noise vectors conditioned for specific beat types. In contrast, the discriminator network plays a dual role as it determines whether the input heartbeats are genuine or synthetic and must predict the class label of those synthetic beats. The figure 4 shows the overall architecture of the ACGAN model.

In terms of its structural composition, the ACGAN model doesn't deviate significantly from previous established models. Nevertheless, it yields exceptional outcomes and demonstrates a notable capacity for stabilizing the training process [21]. The discriminator model is specifically designed for binary real or fake image classification, processing images with dimensions (128, 128, 1). The architecture includes convolutional blocks, starting with a `Conv2D` layer with 16 filters, a 3 $\times$ 3 kernel, and 2 $\times$ 2 strides. Subsequent blocks involve `Conv2D` layers with increasing filter sizes (32, 64, 128, 256, and 512), Leaky ReLU activation ($\alpha = 0.2$), batch normalization, and dropout ($p = 0.5$) for regularization. These blocks progressively downsample the input image, generating feature maps with reduced spatial dimensions. The flattened output connects to two `Dense` layers for prediction. The first `Dense` layer, with a single neuron and sigmoid activation, outputs the probability of the image being real or fake (σ). The second `Dense` layer, with neurons corresponding to specified classes (S, V, and F) and softmax activation (`softmax`), provides class probabilities. The model is compiled using the Adam optimizer with a learning rate of 0.0002 and a β_1 value of 0.5. The loss function is a combination of binary cross-entropy and sparse categorical cross-entropy, reflecting the dual objective of the discriminator to classify images into classes and determine their authenticity. Dropout layers contribute to regularization, and batch normalization enhances stability. Conversely, the generator model is designed for conditional image synthesis, taking a latent vector of dimension 100 (z) and a categorical class label (c) as inputs. The label undergoes embedding (`Embedding`) and linear transformation (`Dense`), resulting in a (32, 32, 1) tensor. Simultaneously, the latent vector undergoes processing through a `Dense` layer with ReLU activation, forming the foundation for a 32×32 image. The two streams merge and are upsampled through `Conv2DTranspose` layers with decreasing filter sizes and increasing strides, ultimately generating an image with dimensions (128, 128, 1). Each `Conv2DTranspose` layer is followed by `BatchNormalization` and ReLU activation, and the output layer utilizes tanh activation to ensure pixel values fall within the range $[-1, 1]$. Table 2 and 3 breakdown the structure of the discriminator and the generator models.

Algorithm 1 Training Process of the ACGAN Model

Input:
- Real samples X_{real}
- Noise samples Z_{gen} from prior $P_g(Z)$
- Data distribution $P_{\text{data}}(X)$
- Number of epochs, n_{epochs}
- Number of steps per epoch, e_{steps}
- Batch size, n_{batch}
- Latent dimension, $latent\_dim$

Output: Trained ACGAN
procedure TrainACGAN
  Initialize discriminator D and generator G parameters.
  for e in 1 to n_{epochs} **do**
    for s in 1 to e_{steps} **do**
      Sample M noise samples Z_k from $P_g(Z)$.
      Sample M examples X_k from $P_{\text{data}}(X)$.
      Update discriminator D by ascending its stochastic gradient.
    end for
    for s in 1 to e_{steps} **do**
      Sample M noise samples Z_k from $P_g(Z)$.
      Update generator G by descending its stochastic gradient.
    end for
  end for
end procedure
Return: Trained ACGAN

Algorithm 1 shows the training process of the ACGAN model, involving alternating steps for updating the discriminator and the generator. The mathematical expressions for updating the discriminator and generator involve binary cross-entropy loss, providing a rigorous framework for the dual objectives of realism and class conditioning in image synthesis. In each training iteration, real samples X_{real} and noise samples Z_{gen} are selected. The discriminator D is then updated by ascending its stochastic gradient, calculated as a combination of the binary cross-entropy loss for real samples and the generated samples. Subsequently, the generator G is updated by descending its stochastic gradient, aiming to minimize the binary cross-entropy loss by fooling the discriminator. This process iterates through multiple epochs and steps, adjusting the model parameters to enhance the generator's ability to produce realistic images while the discriminator learns to distinguish between real and generated samples. The update of the discriminator involves two main loss components: $\mathcal{L}_{\text{real}}$, measuring the dissimilarity between the predicted probabilities and the true labels for real samples, and $\mathcal{L}_{\text{fake}}$, quantifying the discriminator's ability to distinguish between real and generated samples. The discriminator loss $\mathcal{L}_D$ is given by the sum of these components:

$$\mathcal{L}_D = \mathcal{L}_{\text{real}} + \mathcal{L}_{\text{fake}}, \tag{1}$$

where

$$\mathcal{L}_{\text{real}} = -\frac{1}{2}\mathbb{E}_{x,c}[\log D(x|c)], \tag{2}$$

$$\mathcal{L}_{\text{fake}} = -\frac{1}{2}\mathbb{E}_{z,c}[\log(1 - D(G(z|c)))]. \tag{3}$$

TABLE 2. Structure of discriminator in acgan.

Layer	Type	Kernel size	Stride
1	Input	-	-
2	Conv2D	3x3	(2,2)
3	Dropout	-	-
4	Conv2D	3x3	(1,1)
5	BatchNorm	-	-
6	Dropout	-	-
7	Conv2D	3x3	(2,2)
8	BatchNorm	-	-
9	Dropout	-	-
10	Conv2D	3x3	(1,1)
11	BatchNorm	-	-
12	Dropout	-	-
13	Conv2D	3x3	(2,2)
14	BatchNorm	-	-
15	Dropout	-	-
16	Conv2D	3x3	(1,1)
17	BatchNorm	-	-
18	Dropout	-	-
19	Flatten	-	-
20	Dense	-	-
21	Dense	-	-

TABLE 3. Structure of generator in acgan.

Layer	Type	Kernel size	Stride
1	Input	-	-
2	Embedding	-	-
3	Reshape	-	-
4	Input	-	-
5	Dense	-	-
6	Activation	relu	-
7	Reshape	-	-
8	Concatenate	-	-
9	Conv2DTranspose	5x5	(3,3)
10	BatchNorm	-	-
11	Activation	relu	-
12	Conv2DTranspose	5x5	(2,2)
13	BatchNorm	-	-
14	Activation	relu	-
15	Conv2DTranspose	5x5	(2,2)
16	BatchNorm	-	-
17	Activation	relu	-
18	Conv2DTranspose	5x5	(2,2)
19	Activation	tanh	-

TABLE 4. Dataset summary.

Dataset	N	S	V	F	Total
Train	52291	968	4673	550	58482
Augmented Train	52291	52291	52291	52291	209164
Validation	7474	138	667	78	8357
Test	14950	278	1336	158	16722

Similarly, the update of the generator involves two main loss components: $\mathcal{L}_{\text{adv}}$, focusing on generating samples that resemble real ECG signals, and $\mathcal{L}_{\text{aux}}$, encouraging the model to produce synthetic samples with accurate class labels. The generator loss $\mathcal{L}_G$ is given by the sum of these components:

$$\mathcal{L}_G = \mathcal{L}_{\text{adv}} + \mathcal{L}_{\text{aux}}, \tag{4}$$

where

$$\mathcal{L}_{\text{adv}} = -\frac{1}{2}\mathbb{E}_{z,c}[\log D(G(z|c))], \tag{5}$$
$$\mathcal{L}_{\text{aux}} = -\mathbb{E}_{z,c}[\log \text{softmax}(D_{\text{label}}(G(z|c)))]. \tag{6}$$

The algorithm 1 converges over training epochs, optimizing the ACGAN for conditional image generation.

Data augmentation using the ACGAN model was applied on the training dataset. The provided table 4 summarizes the significant impact of this augmentation process on the dataset. Prior to augmentation, training dataset clearly looks imblanace, containing 52291 normal beats (N), 968 supraventricular ectopic beats (S), 4673 ventricular ectopic beats (V), and 550 fusion beats (F), total of 58482 beats. After the augmentation, augmented train dataset was expanded to consist of 52291 beats in each of the four categories (N, S, V, F), resulting in a substantially larger training dataset comprising 209164 beats in total, whereas the validation and test datasets, retained their original composition ensuring unbiased evaluation during model testing and validation. Last, but not least, figure 5 displays a set of samples representing both real and their opposite synthetic heartbeats, each belonging to the various classes sourced from the MIT-BIH dataset. The generated synthetic beats are slightly different but retain the essential characteristics and patterns found in the original heartbeats as we can barely distinguish between the original and the synthetic beats.

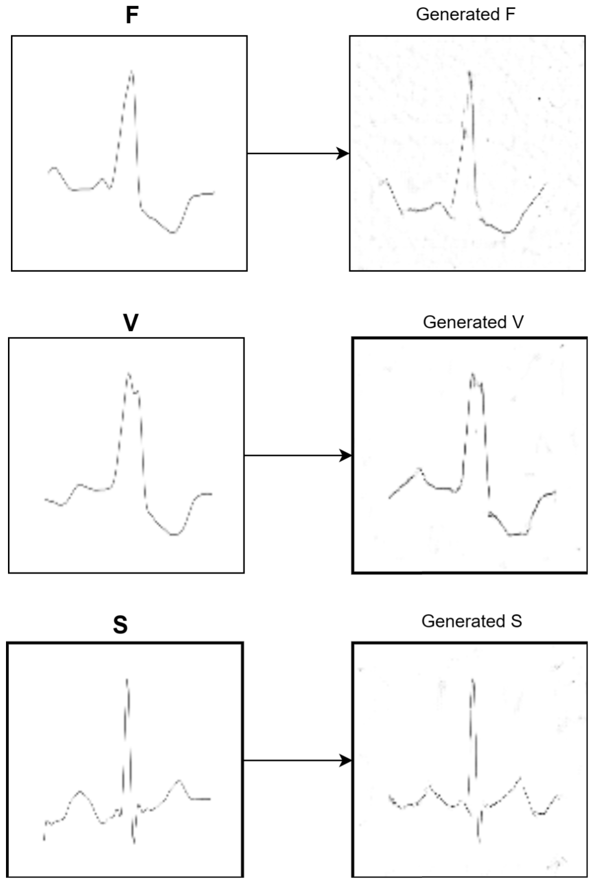

FIGURE 5. The generated synthetic ECG beats.

E. THE LC-CNN MODEL

To detect and to classify arrhythmia in the ECG plot images, CNNs are a category of Deep Neural Networks (DNNs) that find extensive application in tasks such as image classification and signal analysis, including their use in disease diagnosis [22], [23]. These networks have the ability to automatically identify complex patterns in data through the use of stacked trainable small filters or kernels. This characteristic reduces the need for extensive preprocessing when compared to manually engineered feature extraction methods. A typical CNN architecture consists of multiple convolutional layers, each followed by essential components like batch normalization, nonlinear activation functions, dropout layers, pooling layers, and a classification layer (typically a fully connected layer). This arrangement enables the network to effectively capture hierarchical features within the data. One of the remarkable advantages of CNNs is their capacity to deliver outstanding performance while maintaining efficient computational processing. This is achieved through shared-weight architectures and parallelization techniques [24]. Based on that, figure 6 illustrates our proposed lightweight custom LC-CNN model which consists of a series of layers designed for effective feature extraction and classification of the ECG plot images. The architecture begins with an input layer, which receives an ECG grayscale image with dimensions of 128 pixels in width, 128 pixels in height, and a single channel to represent grayscale values. This input is then processed through three convolutional layers followed by max-pooling layers for feature extraction. The initial convolutional layer consists of 32 filters with a kernel size of 3×3 and a stride of 1, accompanied by a Rectified Linear Unit (ReLU) activation function. Subsequently, a max-pooling layer with a 2×2 pool size and a stride of 2 is applied to downsample the feature maps. This is followed by a second convolutional layer comprising 64 filters with a kernel size of 3×3 and a stride of 1, also with ReLU activation. Another max-pooling layer with a 2×2 pool size and a stride of 2 is then employed for further downsampling. The third convolutional layer consists

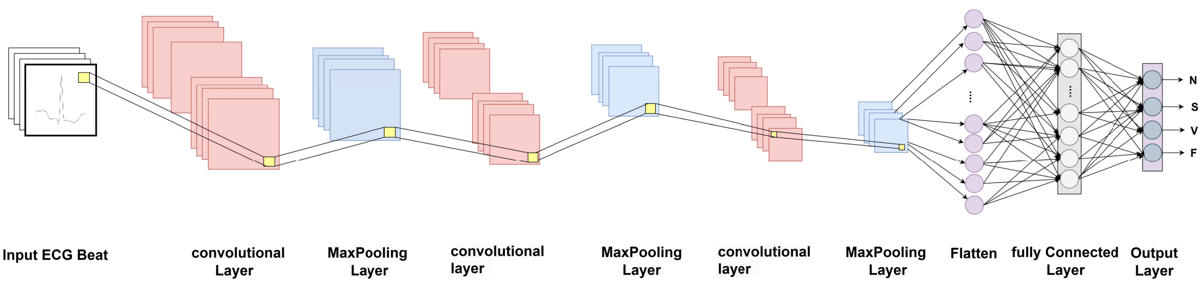

FIGURE 6. The LC-CNN architecture.

TABLE 5. The CNN structure.

Layer Type	Number of Filters	Kernel Size	Stride	Activation	Pooling	Pool Size	Pool Stride
Convolutional	32	3x3	1	ReLU	-	-	-
MaxPooling	-	-	2x2	-	Yes	2x2	2
Convolutional	64	3x3	1	ReLU	-	-	-
MaxPooling	-	-	2x2	-	Yes	2x2	2
Convolutional	128	3x3	1	ReLU	-	-	-
MaxPooling	-	-	2x2	-	Yes	2x2	2
Flatten	-	-	-	-	-	-	-
Dense	128	-	-	ReLU	-	-	-
Dense (Output)	4	-	-	Softmax	-	-	-

of 128 filters with a kernel size of 3×3 and a stride of 1, again with ReLU activation. Once again, a max-pooling layer with a 2×2 pool size and a stride of 2 is added to reduce dimensionality. The flattened output from these convolutional and pooling layers is fed into a dense layer comprising 128 neurons with ReLU activation. Finally, the output layer which utilizes the softmax activation function for class probability estimation, with the output representing the classification of the four ECG arrhythmia classes. For training, the model uses the Adam optimizer with learning rate set to 0.001 and the categorical crossentropy loss function to measure the difference between predicted and actual class probabilities. Table 5 summarizes the proposed LC-CNN model's structure.

F. THE FINE-TUNED MOBILENET-V2

The pre-trained deep learning model, the MobileNet-V2 has been investigated in our paper to detect and to classify ECG arrhythmia. MobileNet-V2, is a neural network architecture designed for efficient and lightweight deep learning on mobile and embedded devices. It aims to strike a balance between model accuracy and computational efficiency. MobileNet-V2 introduces a novel building block called the "Inverted Residual with Linear Bottleneck," which allows for increased representational capacity while keeping the model compact. This architecture includes depth-wise separable convolutions and linear bottlenecks to reduce the number of parameters and computational cost. MobileNet-V2 achieves state-of-the-art results on tasks like image classification and object detection, making it an attractive choice for real-time and resource-constrained applications. It represents a significant advancement in the field of mobile deep learning [25]. By leveraging the MobileNet-V2 model trained on ImageNet, we harness the wealth of knowledge it has acquired during its extensive training. Through fine-tuning, we adapt MobileNet-V2 to our ECG images classification task showcasing the effectiveness of MobileNet-V2 as a versatile and efficient tool for transfer learning, demonstrating its capability to significantly boost the performance of image recognition systems across diverse domains. Figure 5 illustrates the architecture of the proposed fine-tuned MobileNet-V2 model. Our first crucial step was adapting the model to the specific requirements of our grayscale ECG images, which possess a distinct input size of 128×128 pixels and a single channel. To accomplish this, we properly modified the model's input shape and thoughtfully addressed the issue of channel disparity by performing weight averaging on the initial convolutional layer, effectively transforming it from a multi-channel (RGB) layer into a single-channel compatible one. To retain the valuable knowledge ingrained in MobileNet-V2, tailoring it for our ECG classification task, we froze the majority of its layers, excluding only the first convolutional layer and adding our custom classification layers. The latter comprise a Global Average Pooling layer for spatial dimension reduction, a Dense layer for feature extraction with 128 units and ReLU activation, and a final Dense layer with softmax activation for multi-class classification into the four distinct ECG arrhythmia categories. With using the Adam optimizer with a learning rate of 0.0001 and categorical cross-entropy as the chosen loss function, this ensemble of design choice culminates in a tailored model primed for our ECG

TABLE 6. The LC-CNN model evaluation performance metrics.

Run	Accuracy	N			S			V			F		
		Precision	Recall	F1 score	Precision	Recall	F1 score	Precision	Recall	F1 score	Precision	Recall	F1 score
1	99.26%	1.00	1.00	1.00	1.00	0.92	0.96	0.98	0.97	0.97	0.85	0.87	0.86
2	99.25%	1.00	1.00	1.00	0.98	0.96	0.97	0.97	0.97	0.97	0.91	0.82	0.86
3	99.16%	0.99	1.00	1.00	0.98	0.94	0.96	0.99	0.96	0.97	0.84	0.87	0.85
4	99.20%	1.00	1.00	1.00	0.98	0.94	0.96	0.97	0.97	0.97	0.87	0.85	0.86
5	99.25%	0.99	1.0	0.99	0.97	0.95	0.96	0.97	0.97	0.97	0.87	0.82	0.84
Average	**99.22%**	99.6%	100%	99.8%	98.2%	94.2%	96.2%	97.6%	96.8%	97%	86.8%	84.6%	85.4%

TABLE 7. The MobileNet-V2 model evaluation performance metrics.

Run	Accuracy	N			S			V			F		
		precision	Recall	F1 score	Precision	Recall	F1 score	precision	Recall	F1 score	Precision	Recall	F1 score
1	98.74%	0.99	1.00	0.99	0.97	0.85	0.91	0.96	0.94	0.95	0.84	0.78	0.81
2	98.72%	0.99	1.00	0.99	0.95	0.88	0.91	0.96	0.94	0.95	0.80	0.79	0.80
3	98.67%	0.99	1.00	0.99	0.93	0.88	0.91	0.95	0.94	0.95	0.89	0.75	0.81
4	98.71%	0.99	0.99	0.99	0.96	0.88	.092	0.93	0.96	0.95	0.90	0.74	0.81
5	98.53%	0.99	1.00	0.99	0.98	0.82	0.89	0.95	0.93	0.94	0.89	0.69	0.78
average	**98,67%**	99%	99.8%	99%	95.8%	86.2%	90.8%	95%	94.2%	94.8%	88.2%	75%	80.2%

TABLE 8. The LC-CNN and MobileNet-V2 average performance metrics.

Run	Accuracy	N			S			V			F		
		pre	rec	F1	pre	rec	F1	pre	rec	F1	pre	rec	F1
LC-CNN	**99.22%**	99.6%	100%	99.8%	98.2%	94.2%	96.2%	97.6%	96.8%	97%	86.8%	84.6%	85.4%
MobileNet-V2	**98.67%**	99%	99.8%	99%	95.8%	86.2%	90.8%	95%	94.2%	94.8%	88.2%	75%	80.2%

arrhythmia classification, blending the knowledge of pre-trained MobileNet-V2 layers with the specialized capabilities of the custom layers. Figure 7 presents the architecture used during model training.

IV. EXPERIMENTAL RESULTS

This section describes the experimental results obtained from using the proposed deep learning models to detect and to classify ECG arrhythmias.

To ensure the robustness and reliability of our results, the training process for both, LC-CNN and MobileNet-V2 models on the MITBIH dataset were rigorously repeated for five independent runs, allowing us to capture the variability inherent in the training process and ensure the stability and generalization of the models. Subsequently, the models' performances was comprehensively evaluated, and average metrics were calculated across the multiple runs. The use of averaged performance metrics provides a robust and representative assessment of the models' capabilities, mitigating potential biases introduced by a single training instance. Thus, saving the weights associated with the highest validation accuracy during each repetition, serves as an important safeguard against potential model overfitting and ensures that we retained the most optimal models for evaluation. Throughout these repetitions, we incorporated data generators and callbacks' checkpoints to efficiently handle the augmented ECG images and capture the optimal state of the models during training. Given the critical nature of arrhythmia diagnosis, our focus was on achieving the highest accuracy in ECG signal classification, choosing the single best-performing model based on the highest accuracy achieved on the validation set. Notably, with each repetition, we observed that the models consistently tended to reach their optimal accuracy and stabilize at that level. The dataset was split into three parts: 70% for training, 10% for validation, and 20% for testing. We implemented early stopping callbacks based on the validation loss and accuracy which allowed us to end the training process if there was no improvement in validation metrics over a specified number of epochs, preventing overfitting and ensuring optimal model generalization. In the context of the first proposed model, the LC-CNN, we noticed that after just ten training epochs, the model demonstrated remarkable progress with a training accuracy of 99,96% and a validation accuracy of 99,22%. Notably, the loss function consistently decreased during training until it eventually stabilized. In contrast, the second proposed model, which involved fine-tuning the MobileNet-V2 architecture, achieved a training accuracy of 99,8%. However, the validation accuracy exhibited a slight decline, settling at 98,82%. This is quite different from the LC-CNN model, where the disparities between its training and validation accuracies remained relatively small. Additionally, it's noteworthy that the LC-CNN model reached its peak accuracy in significantly less time compared to the MobileNet-V2 model, which required twice the training duration achieving approximate results. Figures 8 and 9 depict the training and validation accuracy and loss curves for the proposed LC-CNN and MobileNet-V2 models, respectively.

In the evaluation, both LC-CNN and MobileNet-V2 models' showed outstanding performance with high accuracy and robust precision, recall and F1 scores across all four

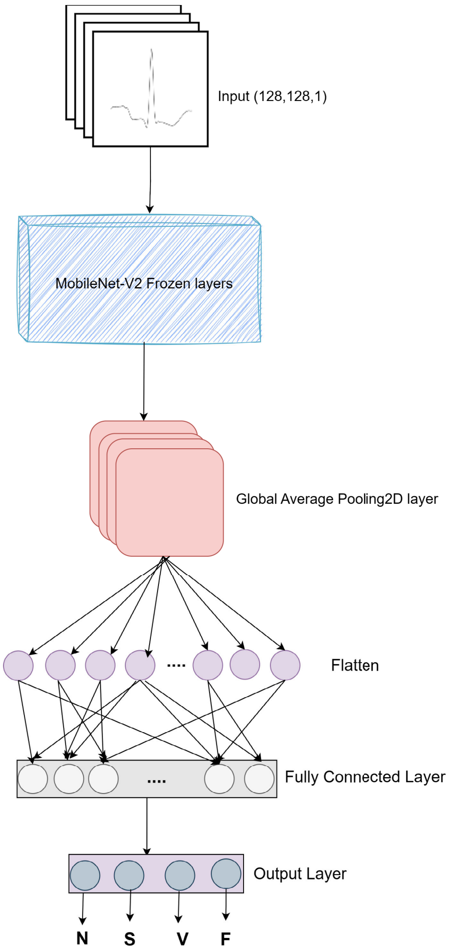

FIGURE 7. The fine-tuned MobileNet-V2.

classes (F, N, S, V). The LC-CNN model displays remarkable consistency and high accuracy in classifying ECG signals into the N, S, V, and F classes across the five repetitions of training. As shown in table 6, the accuracy is consistently high, ranging from 99.16% to 99.26%, demonstrating exceptional performance with an average accuracy of 99.22%. Precision, recall, and F1 score also reveal robust capabilities in distinguishing normal (N) from abnormal (S, V, F) heartbeats. Notably, the model exhibits near-perfect precision and recall for normal beats, showcasing its efficiency in this category. Although slightly lower, precision and recall for the S and V beats remain impressive, while F beats yielded relatively lower scores. The model's overall stability across runs is evident, and its high accuracy, precision, recall, and F1 scores underscore its efficacy in classifying ECG images. Additionally, the MobileNet-V2 model also showcases strong overall performance with an average accuracy across runs of 98.69% as depicted in table 7. The model excels in distinguishing N beats, with precision, recall, and F1 scores consistently exceeding 99%. For the S and V beats, the model maintains high precision and recall, indicating its effectiveness in identifying abnormal heartbeats. However, like the LC-CNN model, the MobileNet-V2 showed slightly low performance

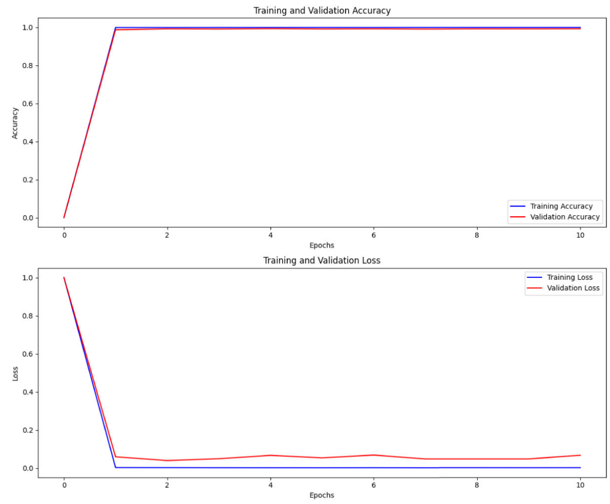

FIGURE 8. Train and Validation accuracy and loss of LC-CNN.

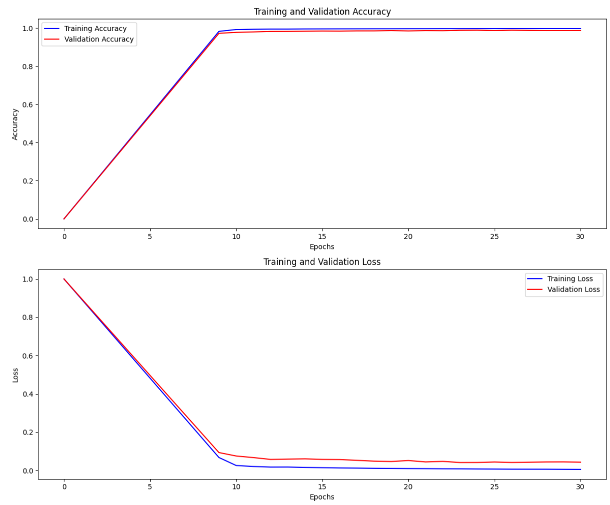

FIGURE 9. Train and validation accuracy and loss of MobileNet-V2.

classifying the F beats, reflected in lower precision, recall, and F1 scores. The model's overall demonstrated a high level of stability, with small variations in metrics between different runs. Overall, the MobileNet-V2 model demonstrates strong potential for ECG signal classification. Table 8 resumes the average performance metrics, encompassing accuracy, precision, F1 score, and recall, for each proposed model.

Furthermore, Figures 10 and 11 illustrate the confusion matrices of the best evaluation of the second run model, offering insights into the prediction outcomes of proposed models. The consistently high precision, recall, and F1 scores observed across most classes indicate that the models prove to be highly dependable and accurate in their classification capabilities, making them valuable choices for our arrhythmia ECG beats classification approach.

To conclude these experiments, a paired t-test with a significance level of 0.05 was conducted to compare the performance of the two deep learning models, LC-CNN and MobileNet-V2, using the accuracy values obtained from the multiple runs. The purpose of this statistical analysis was

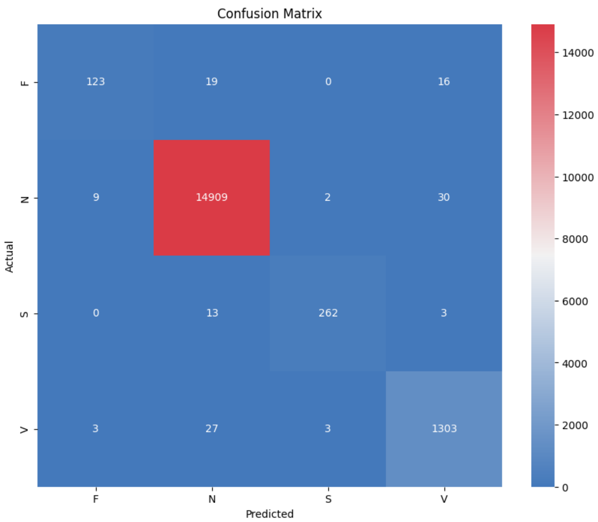

FIGURE 10. Confusion matrix of LC-CNN.

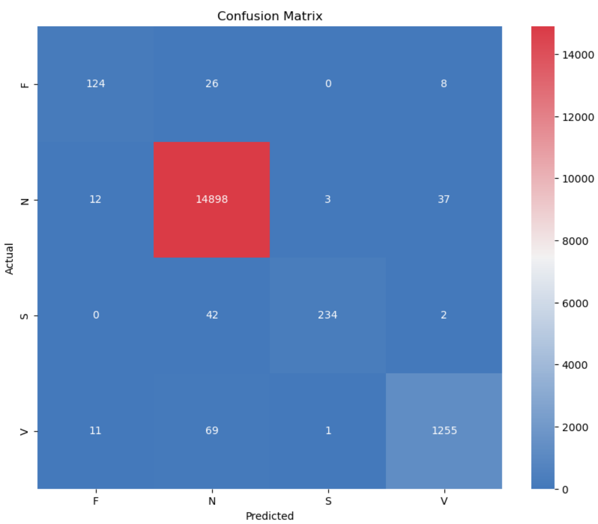

FIGURE 11. Confusion matrix of MobileNet-V2.

to assess whether there is a significant difference in the mean accuracies of the two models. The null hypothesis (H_0) proposed that there is no significant difference in the mean accuracies between MobileNet-V2 and CNN, while the alternative hypothesis (H_1) suggested the presence of a significant difference. The computed results of the paired t-test were a t-statistic of -12.71867547673 and a p-value of 0.000220137589. The t-statistic represents the strength of the evidence against the null hypothesis, and the p-value indicates the probability of observing such results if the null hypothesis were true. In our case, the remarkably low p-value (less than the common significance level of 0.05) led to the rejection of the null hypothesis. Further exploration of the variation in accuracy was identified, revealing a variance of 0.00713 for MobileNetV2 and 0.00183 for CNN since the mean accuracy for MobileNet-V2 was determined to be 98.674%, while CNN exhibited a slightly higher mean accuracy of 99.224%. These statistics clarifies not only the central tendency of the models' performances but also the degree of variability. The conclusion is that there is a significant difference in the performance of the MobileNet-V2 and LC-CNN models as already observed based on the accuracy metrics. The negative t-statistic suggests that on average, the CNN model outperformed MobileNet-V2 for our MiT-BiH ECG image dataset.

V. PERFORMANCE COMPARISON

We present a comprehensive performance comparison between our proposed approach using the LC-CNN and MobileNet-V2 models, along with several existing deep learning studies aimed at detecting and classifying ECG signals. As demonstrated in Table V, our research prominently distinguishes itself by achieving the highest accuracy rate of 99.22% through the utilization of the LC-CNN model. Notably, Works [8] and [17] incorporated GAN models for data augmentation, while the study of [14] employed the GAN for ECG signal classification. Conversely, the remaining studies relied on conventional techniques such as cropping, resizing, shifting, flipping, and resampling to balance the dataset. However, none of these methods managed to surpass the accuracy achieved in our paper. Therefore, the adoption of ACGAN model for data augmentation totally emerges as a pivotal innovation, substantially enhancing our models' proficiency in accurately classifying ECG beats. This underscores the effectiveness of our selected data augmentation strategy in elevating the performance of deep learning models for ECG beat classification. Additionally, the MobileNet-V2 model, with its average accuracy of 98.69%, surpasses four out of the eight compared works. It's essential to highlight that these studies have employed diverse deep learning methodologies for ECG arrhythmia classification, encompassing various CNN model architectures and combinations with other deep learning models. This reaffirms the exceptional effectiveness of our approach, establishing its ability to outperform a multitude of methods documented in the existing literature.

VI. ABLATION STUDY

In this section, we conduct an ablation study to analyze the impact of using the different components within our proposed approach on the achieved results. We examine the effects of denoising and augmentation on our proposed approach for both proposed models. We assessed three different outcomes for the following configurations: without denoising and augmentation, without augmentation, without denoising, compared to the proposed full approach. The provided table 10 outlines the performance metrics of the various models and configurations. Notably, the LC-CNN without denoise demonstrates impressive accuracy of 99.15% and balanced precision, recall, and F1 score, emphasizing its robustness even without denoising. The slight drop in accuracy for LC-CNN without augmentation suggests the importance of data augmentation for improved generalization. In contrast, MobileNet-V2 without denoise and augmentation exhibits

TABLE 9. Comparison of our work with other results.

Ref	Database	Data Augmentation	Deep Learning classifier	No of classes	Accuracy
[8]	MIT-BIH +PTB	GAN	Ensemble GAN-LSTM	2	99.2%
[10]	PTB-XL	DA algorithm	CNN	12-leads	89.87%
[12]	MIT-BIH + PTB	Standard operations and techniques	Dence-Net	29	98.92%
[13]	MIT-BIH + PTB	Standard techniques	CNN + fuzzy clustering	5	98.66%
[14]	MIT-BIH	/	GAN	2	97.02%
[20]	MT-BIH	/	LSTM, Autoencoder	5	98.57%
[17]	MIT-BIH	GAN	CNN TWO STAGE CNN	15	98.30 % 98 %
This study	MIT-BIH	AC-GAN	LC-CNN	4	**99.22%**
			MobileNet-V2		**98.69%**

TABLE 10. Ablation study for our proposed approach to study the effect of denoising and augmentation.

Method	Accuracy	Precision	Recall	F1 score
LC-CNN without denoise	99.15%	95%	94%	95%
LC-CNN without augmentation	99.10%	95%	91%	93%
MobileNet-V2 without denoise	98.70%	95%	88%	91%
MobileNet-V2 without augmentation	98.58%	95%	88%	91%
LC-CNN without denoise and augmentation	99.07%	97%	91%	93%
MobileNet-V2 without denoise and augmentation	98.22%	97%	81%	87%
Proposed full LC-CNN	99.22%	98.2%	94.2%	96.2%
Proposed full MobileNet-V2	98.69%	95.8%	86.2%	90.8%

lower performance, highlighting the impact of architectural differences and the necessity of denoising and augmentation. The proposed LC-CNN model stands out with the highest accuracy, precision, recall, and F1 score, indicating the success of the proposed components, especially exhibited superior F1 scores for S and V beats. The proposed MobileNet-V2, while slightly falling behind the LC-CNN, still outperforms individual configurations, emphasizing its effectiveness. This ablation study underlines the critical role of denoising and augmentation, with models lacking these components experiencing significant performance drops. The proposed models, incorporating all the components, showcase the potential for further improvements, demonstrating the positive impact of a comprehensive approach in achieving superior classification results.

VII. DISCUSSION

in this study, we propose contributions to the domain of arrhythmia detection using ECG signals and deep learning models. One innovative aspect of our work involves the conversion of ECG signals into PNG images through plotting. This process aims to explore novel approaches in arrhythmia analysis, employing computer vision to potentially enhance diagnostic accuracy. Two Convolutional Neural Network models trained and evaluated on the MIT-BIH dataset were proposed. These models consist of the lightweight custom CNN (LC-CCN) model and the pre-trained MobileNet-V2. Opting for the LC-CNN model with only three convolutional layers highlights the effectiveness of simplicity in deep learning architectures. By doing so, we demonstrate that high accuracy can be achieved without the need for excessive model complexity. Furthermore, incorporating transfer learning, especially with MobileNet-V2, involves leveraging pre-trained features and transfer learning knowledge to improve the accuracy of arrhythmia detection. MobileNet-V2 is known for its lightweight design which makes it suitable and effective to be investigated along with the proposed LC-CNN to significantly boost the performance of ECG images classification. Additionally, the preprocessing steps played a crucial role in improving classification accuracy by ensuring high-quality, well-segmented signals. Furthermore, one of our main contributions lies in addressing the challenge of data imbalance, a common issue in the MIT-BIH dataset. We introduce an ACGAN model to generate synthetic ECG signals, with a specific focus on the underrepresented arrhythmia classes. This innovative data augmentation technique significantly bolsters our dataset, thereby improving the proposed models' capability to handle the minority classes effectively. The rationale behind opting for the ACGAN model was that in essence, we employed a conditional Generative Adversarial Network (CGAN) [26], to address the issue of data imbalance. Our objective was to utilize the conditional feature within the GAN model to mitigate class imbalance issues specifically within the minority classes. However, despite its architectural framework similarity with the ACGAN, the performance of the CGAN remained unsatisfactory even after a prolonged training period of approximately 14,000 epochs. The CGAN was capable of generating synthetic data, but it failed to effectively eliminate the inherent noise in the generated data, also, it was unable to handle the considerable diversity present within the ECG beats belonging to a single class. Whereas, the ACGAN displayed superior performance, producing high-quality synthetic ECG images with relative ease. remarkably, the ACGAN exhibited promising results within the initial epochs of training. This improved performance can be attributed to its capacity for additional class label prediction, making it a more favorable choice for the generation of ECG images in our particular application. To evaluate the impact of denoising and augmentation, we performed an ablation study. Results indicate that the proposed approach's performance significantly benefits from these enhancements. Furthermore, we conducted a comprehensive analysis of average performance metrics to assess the overall effectiveness of our proposed models. Metrics such as accuracy, precision, recall, and F1 score were computed. The inclusion of denoising and augmentation consistently elevated these metrics, underscoring the positive impact of these techniques on the models' overall performance. Hence, to rigorously validate our findings, we performed statistical tests comparing the LC-CNN and MobileNet-V2 models. The results of the tests confirm that the LC-CNN model outperforms MobileNet-V2 in terms of arrhythmia detection accuracy. This supports that simplicity, coupled with effective design choices, can lead to superior model performance. Eventually, while the proposed approach has shown promising results, there are certain limitations that require attention and further investigation. To begin with, in our study, we opted to transform ECG signals into structured plots as images PNG format, representing sequences of ECG data points. Nevertheless, it's noteworthy to acknowledge that ECG signals inherently belong to the category of time series data, given their continuous recording over time. Future research efforts would benefit from generating ECG images that encapsulate both temporal and frequency/scale information, employing techniques like spectrogram generation or wavelet transforms. Moreover, our approach was implemented using relatively simple deep learning models for the detection and classification of ECG signals. After yielding excellent results, we believe that employing more intricate and deeper deep learning architectures has the potential to further enhance the classification performance of ECG signals.

VIII. CONCLUSION

In this paper, we presented a novel deep learning approach for the diagnosis and classification of cardiac arrhythmias using ECG signals plot images. Our approach incorporated the design of a Lightweight Custom Convolutional Neural Networks (LC-CNN) and fine-tuning the pre-trained MobileNet-V2 model. Both models, when trained with our proposed architectures, exhibited remarkable performance, achieving average accuracy rates of 99.22% and 98.69%, respectively. We emphasize the significance of data preprocessing, encompassing noise reduction and beat segmentation in ensuring high-quality input for our model. Not to mention one of the highlights of this work is our novel approach to class imbalance through the utilization of an ACGAN model for data augmentation. This addresses a common challenge in medical datasets and enhances the model's capability to detect underrepresented arrhythmia classes. Our comprehensive evaluation, including precision, recall, and F1-score demonstrates the robustness of our approach beyond just high accuracy. After achieving excellent results, in the future studies, our focus will extend to conduct comprehensive generalization testing on diverse ECG datasets. Furthermore, more complex deep learning models would be considered to enhance the classification performance of the ECG signals as images using techniques like spectrogram generation or wavelet transforms. Lastly, we would delve into the implementation of real-time ECG arrhythmia monitoring applications based on wearable IoT technologies.

REFERENCES

[1] X. Liu, H. Wang, Z. Li, and L. Qin, "Deep learning in ECG diagnosis: A review," *Knowl.-Based Syst.*, vol. 227, Sep. 2021, Art. no. 107187.
[2] S. Ma, J. Cui, C.-L. Chen, X. Chen, and Y. Ma, "An effective data enhancement method for classification of ECG arrhythmia," *Measurement*, vol. 203, Nov. 2022, Art. no. 111978.
[3] F. Murat, O. Yildirim, M. Talo, U. B. Baloglu, Y. Demir, and U. R. Acharya, "Application of deep learning techniques for heartbeats detection using ECG signals-analysis and review," *Comput. Biol. Med.*, vol. 120, May 2020, Art. no. 103726.
[4] E. Merdjanovska and A. Rashkovska, "Comprehensive survey of computational ECG analysis: Databases, methods and applications," *Exp. Syst. Appl.*, vol. 203, Oct. 2022, Art. no. 117206.
[5] M. A. Serhani, H. T. El Kassabi, H. Ismail, and A. Nujum Navaz, "ECG monitoring systems: Review, architecture, processes, and key challenges," *Sensors*, vol. 20, no. 6, p. 1796, Mar. 2020.
[6] L. Berger, M. Haberbusch, and F. Moscato, "Generative adversarial networks in electrocardiogram synthesis: Recent developments and challenges," *Artif. Intell. Med.*, vol. 143, Sep. 2023, Art. no. 102632.
[7] (2009). *PhysioNet*. [Online]. Available: https://physionet.org/content/mitdb/1.0.0/
[8] A. Rath, D. Mishra, G. Panda, and S. C. Satapathy, "Heart disease detection using deep learning methods from imbalanced ECG samples," *Biomed. Signal Process. Control*, vol. 68, Jul. 2021, Art. no. 102820.
[9] Y. Xia, Y. Xu, P. Chen, J. Zhang, and Y. Zhang, "Generative adversarial network with transformer generator for boosting ECG classification," *Biomed. Signal Process. Control*, vol. 80, Feb. 2023, Art. no. 104276.
[10] M. F. Safdar, P. Palka, R. M. Nowak, and A. A. Faresi, "A novel data augmentation approach for enhancement of ECG signal classification," *Biomed. Signal Process. Control*, vol. 86, Sep. 2023, Art. no. 105114.
[11] Y. Xia, Y. Xiong, and K. Wang, "A transformer model blended with CNN and denoising autoencoder for inter-patient ECG arrhythmia classification," *Biomed. Signal Process. Control*, vol. 86, Sep. 2023, Art. no. 105271.
[12] A. Pal, R. Srivastva, and Y. N. Singh, "CardioNet: An efficient ECG arrhythmia classification system using transfer learning," *Big Data Res.*, vol. 26, Nov. 2021, Art. no. 100271.
[13] S. Kumar, A. Mallik, A. Kumar, J. D. Ser, and G. Yang, "Fuzz-ClustNet: Coupled fuzzy clustering and deep neural networks for arrhythmia detection from ECG signals," *Comput. Biol. Med.*, vol. 153, Feb. 2023, Art. no. 106511.
[14] J. Zhuang, J. Sun, and G. Yuan, "Arrhythmia diagnosis of young martial arts athletes based on deep learning for smart medical care," *Neural Comput. Appl.*, vol. 35, no. 20, pp. 14641–14652, Jul. 2023.
[15] R. Siouda, M. Nemissi, and H. Seridi, "ECG beat classification using neural classifier based on deep autoencoder and decomposition techniques," *Prog. Artif. Intell.*, vol. 10, no. 3, pp. 333–347, Sep. 2021.
[16] H. M. Rai and K. Chatterjee, "Hybrid CNN-LSTM deep learning model and ensemble technique for automatic detection of myocardial infarction using big ECG data," *Int. J. Speech Technol.*, vol. 52, no. 5, pp. 5366–5384, Mar. 2022.

[17] A. M. Shaker, M. Tantawi, H. A. Shedeed, and M. F. Tolba, "Generalization of convolutional neural networks for ECG classification using generative adversarial networks," *IEEE Access*, vol. 8, pp. 35592–35605, 2020.
[18] E. Essa and X. Xie, "An ensemble of deep learning-based multi-model for ECG heartbeats arrhythmia classification," *IEEE Access*, vol. 9, pp. 103452–103464, 2021.
[19] H. Tesfai, H. Saleh, M. Al-Qutayri, M. B. Mohammad, T. Tekeste, A. Khandoker, and B. Mohammad, "Lightweight shufflenet based CNN for arrhythmia classification," *IEEE Access*, vol. 10, pp. 111842–111854, 2022.
[20] P. Liu, X. Sun, Y. Han, Z. He, W. Zhang, and C. Wu, "Arrhythmia classification of LSTM autoencoder based on time series anomaly detection," *Biomed. Signal Process. Control*, vol. 71, Jan. 2022, Art. no. 103228.
[21] A. Odena, C. Olah, and J. Shlens, "Conditional image synthesis with auxiliary classifier GANs," in *Proc. Int. Conf. Mach. Learn.*, 2017, pp. 2642–2651.
[22] O. Moussa, H. Khachnaoui, R. Guetari, and N. Khlifa, "Thyroid nodules classification and diagnosis in ultrasound images using fine-tuning deep convolutional neural network," *Int. J. Imag. Syst. Technol.*, vol. 30, no. 1, pp. 185–195, Mar. 2020.
[23] R. Mastouri, N. Khlifa, H. Neji, and S. Hantous-Zannad, "A bilinear convolutional neural network for lung nodules classification on CT images," *Int. J. Comput. Assist. Radiol. Surgery*, vol. 16, no. 1, pp. 91–101, Jan. 2021.
[24] S. Hong, Y. Zhou, J. Shang, C. Xiao, and J. Sun, "Opportunities and challenges of deep learning methods for electrocardiogram data: A systematic review," *Comput. Biol. Med.*, vol. 122, Jul. 2020, Art. no. 103801.
[25] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 4510–4520.
[26] M. Mirza and S. Osindero, "Conditional generative adversarial nets," 2014, *arXiv:1411.1784*.

HADJER BECHINIA received the bachelor's degree and the M.S. degree in information systems from the Faculty of New Technologies of Information and Communication, University of Abdelhamid Mehri Constantine 2, in 2016 and 2018, respectively, where she is currently pursuing the Ph.D. degree with the LIRE Laboratory, Department of TLSI, with a focus on heart monitoring systems using deep learning. Her research interests include medical image processing and deep learning in the IoT applications.

DJAMEL BENMERZOUG received the Ph.D. degree in computer science from Pierre and Marie Curie University, Paris, France. He is currently a Full Professor with the Department of TLSI, Faculty of New Technologies of Information and Communication, University Constantine 2, Algeria. He has published many papers in many international conferences and journals. He supervises many master's and Ph.D. students. His current research interests include the Internet of Things, cloud computing, advanced enterprises systems, multiagent systems, service oriented computing, and business processes modeling and verification.

NAWRES KHLIFA received the Ph.D. degree from the National School of Engineers of Tunis (ENIT). She is currently a Professor with the Higher Institute of Medical Technologies of Tunis, Tunis El Manar University. She is also an Engineer. She coordinates the TIMEd Team: Medical Image Processing, BTM Laboratory. Her research interests include artificial intelligence and CAD design in medical imaging, emotion recognition, and gaze tracking.