Journal of Theoretical and Applied Information Technology

15th December 2023. Vol.101. No 23

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

MULTI-CLASS CARDIOVASCULAR DISEASES CLASSIFICATION SYSTEM USING DENSE INCEPTION ATTENTION NEURAL NETWORK

MR. AFROZ PASHA[1], DR.NAGARAJA S R[2]

[1]Research scholar, Department of CSE, Presidency University, Bangalore
[2]Professor, Department of CSE, Presidency University, Bangalore
Email: afroz.nmit@gmail.com[1], nagarajasr@presidencyuniversity.in[2]

ABSTRACT

In the entire world, cardiovascular diseases (CVD) are the main cause of death compared to other disease. Before experiencing a catastrophic heart failure event such as a stroke, heart attack, or myocardial infarction, people with CVDs may not receive a diagnosis. To address the above problems, To address the above problems, this paper proposes a Dense Inception Attention Neural Network (DIAN-Net) for the Classification of Cardiovascular Diseases using ECG Signals, We add residual blocks, and residual convolutional layer pathways are integrated into the atrous spatial pyramid pooling (ASPP) module and Multi-Scale Context Fusion Block (MSCFB). To fuse convolutional feature maps in encoding layers, the ASPP unit used a learnable set of parameters. An efficient architecture for feature extraction during the encoding step is the ASPP unit. We integrated the AD unit with the benefits of the U-Net network for deep and shallow features. The proposed decoder takes advantage of the multi-scale features from the encoder to predict CVD regions. The aforementioned tests demonstrate that the newly created deep learning models may be very helpful in clinical decision-making and nuclear medicine. The experimental results show that the algorithm in this paper has a smaller number of parameters and shorter training time, and outperforms other methods in terms of subjective visualization and objective evaluation metrics on multiple benchmark datasets.
Keywords: *ECG Signal, CNN, Dense Layer, Cardiovascular Diseases.*

1. INTRODUCTION

Heart disease, or cardiovascular disorders, is the world's largest cause of mortality, according to the World Health Organisation. An estimated 17.9 million lives are lost to them annually, making up 32% of all fatalities globally. Myocardial infarctions (MI), commonly referred to as heart attacks, account for over 85% of all heart disease-related deaths [1]. If cardiovascular disease is effectively diagnosed and treated early on, many lives can be spared [1]. The healthcare system uses a variety of methods, including blood testing, computed tomography, cardiac magnetic resonance imaging, echo, ECG, and others, to identify heart disorders [2].[3]. The ECG is a widely used, reasonably priced, and non-invasive method of determining the heart's electrical activity [4]. It is employed to detect cardiovascular disorders linked to the heart [4].

Heart attacks, strokes, and even death can result from cardiovascular diseases (CVDs), which cause plaque to accumulate in the heart's arteries, the brain's arteries, and inside the body's major blood vessels. The primary
cause of mortality in the US is coronary heart disease1. A variety of techniques, including MRI, X-ray, ultrasound, and heart sounds, can be used to diagnose CVDs 2–6.

The standard 12-lead electrocardiogram (ECG), among other techniques, is the clinical instrument most frequently used to diagnose heart disorders since it is straightforward, non-invasive, and reasonably priced. However, ECG interpretation takes a lot of effort and expertise from highly qualified cardiologists. Cardiologists are able to diagnose patients more accurately when they use computed assisted detection and classification of heart problems[5-8].Following the development of Convolutional Neural Networks (CNNs), they have been applied to the field of medical image segmentation. U-Net[1] is one of the most universal networks. However, CNNs are only good at acquiring local features and hard to capture long-distance dependencies effectively due to their restricted perceptual field. Moreover, the fixed size and shape of convolutional kernels cannot effectively adapt to the types of inputs, which limits

the further applicability of CNNs. Most of the time, the medical image segmentation task is not a local issue. Therefore, acquiring high-quality global information is crucial for improving model performance. To make up for the weakness of CNNs, transformers are applied to the area of medical image segmentation. Following the methodology of Vision Transformer (ViT) [2], [3] sequenced images into patches and then utilized transformer modules to get the relationships among the patches, which facilitates subsequent modules to use them to guide feature representations.

Deep learning models including convolutional neural networks (CNNs) gradually become a common method in medical imaging, particularly in artifact correction and noise reduction.

Deep learning methods combine with traditional analytical or iterative schemes to provide greater efficiency and precision [12]. In CVDs classification, neural networks can be trained on large datasets to extract image features and restore missing information from rarefaction angles. For instance, U-Net-based interpolation of sinograms can synthesize missing data in sparse-view sinograms [13]. iCT-net [14] is proposed to incorporate the reconstruction process into the network. The incorporation of prior images into CNN-based methods has also been explored. CycN-Net [15] introduces a novel approach wherein both the degraded and the prior images are encoded and jointly utilized in the decoding step. Recently, there are many processes that combine deep learning with motion compensation schemes. For example, the 3D U-net is used to optimize the image and combine it with motion compensation[16-18]. The remaining sections of the paper is arranged like; Section 2 summarizes the literature review based on heart disease prediction followed with the proposed framework is illustrated in Section 3. Moreover, Section 4 discusses the experimental results and the paper is concluded in Section 5.

2, LITERATURE SURVEY:

Many research works have been carried out to use machine learning and deep learning techniques to automatically forecast cardiovascular illnesses by employing ECG as a digital or picture data representation.

Miao et al. [19] have described an improved random survival forest (iRSF) a multi-hazard approach that determines cardiovascular disease fatality. Depending on 32 indicators, such as biometrics, diagnostic features, analytical relevant data, and prescriptions. This method can identify the inpatient survival of ICU patients experiencing heart failure. It sought to develop predictions that may differentiate participants at varying levels of risk, particularly for authentic parameters and characteristics in smaller numbers. The increased survival function is predicted efficiently. However, in-home monitoring options it is inadequate to regulatory requirements throughout many contexts.

Jinee et al. [20] suggested a hybrid fuzzy-based decision tree algorithm that describes the technique employed in heart disease prognosis. Cardiac sensors, oxygen saturation sensors, and especially hemoglobin duration detectors can be used for illness assessment and therefore are connected to the human chest to determine the existence of the disease. The alerting method incorporates an adjustable message alert, which sends notifications to the appropriate doctor and caregiver. It is efficient, specific, and sensitive to access the suggested technique. Thus, the patient's medical expenses have become more expensive.

Ali et al. [21] have demonstrated a hybrid grid search algorithm (HGSA) to strengthen the diagnosing procedure. The input variables of the two models are combined to form a hybrid grid. The initial parameter of the hyperparameter will produce the ideal points of optimized iterations. The maximum node mostly on the hybrid grid relates towards the perfect extracted features as well as the appropriate deductive approach, both of which should exhibit outstanding results on the ideal extracted features. It is effective in the complexity of time and gives better results within a short period. Moreover, a greater level of professional skills is needed.

Ramesh et al. [22] highlighted an Information Gain based Feature Selection (IGFS) to increase the reliability of cardiovascular disease risk assessment. The two sections of the dataset are testing and training datasets. The dataset retrieved by the feature dataset to complete the process will decrease when each of the functions has been used. The training phase provides a large dataset to teach learners the skills needed for forecasting. It is efficient and improves performance. However, using other models heart diseases are predicted more accurately.

Mienye et al. [23] describe a sparse autoencoder (SAE) approach to accurately forecast cardiac

illness. The architectures with autoencoders recreate their source somewhere at the outcome and are mainly composed of a decoder and encoder. The original data is processed in the input data that are modified in low-level representations, the input is converted into a unique conception. It enhances the model's functionality, efficiency, and stability. Hence, it is complicated to use the features in large datasets.

Gokulnath et al. [24] have determined the support vector machine (SVM) to put forward an optimization function for determining the more important factors to develop heart disease. The important features from the dataset have been determined by using several proper evaluation methods. The noisy and incomplete data with essential functions evaluate the classification features. It is utilized in many systems when there is little communication between variables and in which information with limited dependence is captured. The identification of a cardiac disease with higher efficiency. Hence, the expense of the process is increased.

Gárate-Escamila et al. [25] have evaluated the dimensionality reduction method using a subset of features tool to identify characteristics of heart disease. The original information is evaluated using the six classifiers. To identify a singular and condensed collection of performance rating characteristics associated with the diagnosis of cardiovascular disease and authenticate those using classifiers. This method can be used in a variety of real-world scenarios as well as in other medical diagnoses. Therefore, the limited number of participants makes it impossible to generalize such observations to cardiovascular diseases.

Nagarajan et al. [26] have implemented a genetic-based crow search algorithm (GCSA) a dependable and effective way to identify heart problems. The feature selection is divided into two parts wrapper and filter method. The features are first filtered without first being classified, negating the need for a binary classifier. Every single one of the variables of the database is given model parameters, which are then evaluated for the entire data. It is a powerful and effective approach to detecting heart problems. However, duplicate data may contain noise so the training duration is increased is to be rectified.

3. PROPOSED METHODOLOGY

Tremendous progress in CVDs automatic classification has occurred, primarily due to the advances of deep convolutional neural networks. Those methods have outperformed traditional methods by a large margin in classification performance attributed to the powerful capability for automatic feature extraction. Nevertheless, one of the critical issues is following the strict assumption that the training datasets have an identical distribution to the test datasets. Medical datasets are typically acquired by hospitals with various scanner vendors and scanner protocols, resulting in variations of appearance distribution discrepancies (termed domain shifts). Most of the studies revealed the considerably deteriorated classification performance of well-trained CNNs on the testing set with such distribution discrepancy. Unsupervised domain adaptation , which mitigates the distribution discrepancies across domains by learning domain-invariant features or generative adversarial networks , is one of the typical solutions in solving domain shift issues. On the other hand, existing unsupervised domain adaptation techniques typically require sufficient accessible datasets from the target domain during the training procedure. This prerequisite may not be feasible because of the limited number of target domain and the unpredictable target domain shifts[27], the overview of the developed network, as shown in Fig.1.

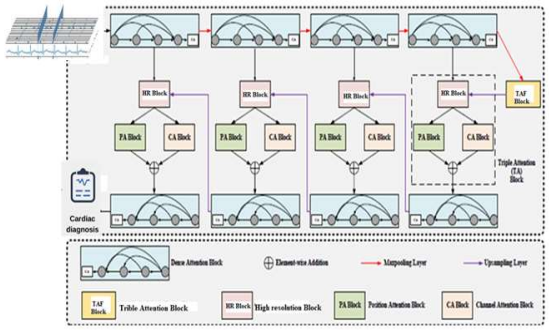

Figure 1 : Overall Flow Diagram Of The Proposed Cvds Classification System

3.1 High-Resolution Block

Currently, the majority of deep learning techniques use serial connections, and the residual network suggestion greatly increases the network's depth[38]. However, this scaling necessitates time-consuming human adjustment, which frequently results in subpar performance. Additionally, the

network's total number of parameters is relatively big, which makes training the model challenging. This issue might be resolved by a high-resolution network (HR-Net). According to Fig. 2, it is a parallel network structure. The parallel network begins with high-resolution feature mapping and eventually connects multi-resolution sub-networks while also introducing high-resolution to low-resolution sub-networks[28]. By repeatedly exchanging data on the parallel multi-resolution sub-network through the exchange unit, HR-Net conducts multi-scale fusion during the entire operation. At the same time, the exchange unit can simultaneously receive data from other parallel sub-networks. Each stage has one more branch than the previous stage. The previous stage's feature maps were all fused together to produce the new branch. The number of channels is double that of the prior level, and the resolution is half that of the previous stage.

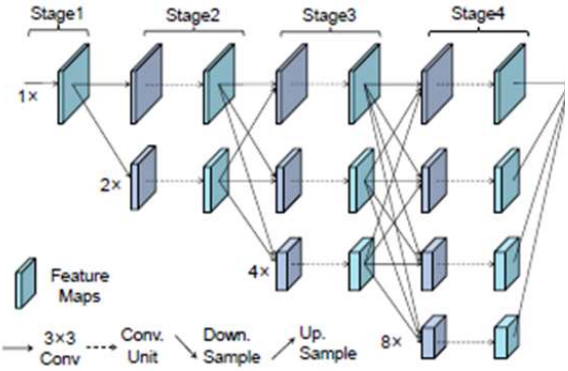

Fig. 2. The Structure Of HR Block With Four Stages.

The {Ix − Is} are the input feature maps, and {Ox − Os} are the output feature maps. Each output can be expressed using Eqn (1):

$$O_k = \sum_{x=1}^{n} a(I_x, k) \quad (1)$$

Equation (1) requires that all of the inputs have the same widths and resolutions. A change in resolution is necessary when the exchange unit between stages requires it, as defined by Eqn (2) for an extra out map.

$$O_{s+1} = a(O_s, s+1) \quad (2)$$

The functional (Oi, k) upsamples or downsamples Ox from resolution x to k, where x, s, and k represent the resolution., if x = k, a(,) is a regular connection.

The suggested framework for high-level feature extraction has four steps, represented by the HR Block in Fig. 2, which is based on HR-Net. A 3X3 convolution kernel, a batch normalization (BN), and four identical convolutional layers make up each convolution unit., and an activation of the rectified linear unit (ReLU). When downsampling, the scale of the original input is decreased using a 3X3 convolution layer with a stride of 2. The nearest neighbor interpolation algorithm is used during upsampling to restore the resolution. After these operations are applied, the outputs of the three branches below are concatenated to form a single input for the ASPP module before being upsampled to adjust their resolutions to match that of the first branch.

3.2 Atrous Spatial Pyramid Pooling

The proposed framework adopts the same ASPP module as DeepLabv3+[29]. The 1X1 convolution branch and three 3X3 atrous convolution branches are part of the first section, as shown in Fig. 3(a). The respective atrous rates are 6, 12, and 18. There are 256 filters and a BN layer in

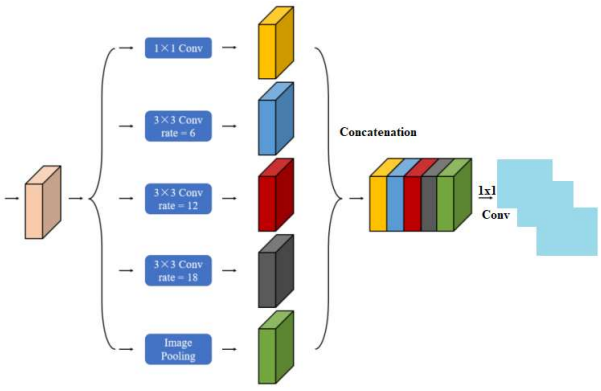

Fig. 3. Details Of The Atrous Spatial Pyramid Pooling Module.

each convolution layer. As seen in Fig. 3 (b), the second section uses global average pooling to extract image-level information and restore resolution by bilinear interpolation. The ASPP module's convolution branches are all combined and put into a 1 X 1 convolution layer with 256 filters at the end. This procedure is used to lower the

parameters of high-dimensional features and enhance the module's capacity for expression. High-level features are the ASPP module's product.

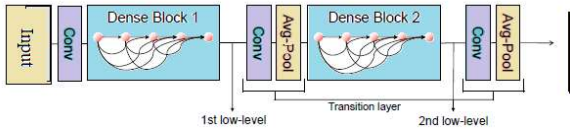

Fig. 4. Low-Level Feature Extraction Module Of The Proposed System.

3.3 Low-level Feature Extraction Module

Texture features play an important role in classification and are one of the popular low-level features in image classifications Experimental research shows that the color-antagonistic receptive field mechanism of the human visual system is of great significance in a variety of visual perception tasks (such as edge detection, target recognition, and boundary detection). Furthermore, a Gabor filter can approximate the receptive field function of the simple cells of higher vertebrates well [30]. A set of Gabor filters with different frequencies and orientations can be used to extract useful visual features (textures and edges) for representation together with particularly appropriate ones. Dense Net[31] proposed an excellent dense connectivity pattern. The dense connectivity can be written as follows:

$$\psi_i = H_i([\psi_0, \psi_1, \ldots\ldots \psi_{i-1}]) \quad (3)$$

Where $[\psi_0, \psi_1, \ldots\ldots \psi_{i-1}]$ to the concatenation of the feature maps produced in layers 0, ..., i-1, The non-Linear transformation, represented by Hi, consists of a BN layer, ReLU activation, and Convolution layer. The output of each Dense BLock is converted by the Transition layer. A BN layer, a ReLU activation, a 1X1 Convolution, and a 2X2 Avg Pooling layer make up a Transition layer. These techniques can shrink the dimensions of channels and lower the size of feature maps.

Considering the lack of spatial information and local geometric invariance of deep features extracted from the full connection layer, intermediate convolutional layers are utilized to extract deep features here. We define the deep feature maps of the convolutional layer as a three-dimensional feature tensor and denoted as Per Eqn(4)

$$P \in \mathbb{R}^{K \times W \times H}: \quad P = \{P_k \in \mathbb{R}^{W \times H}\}, \quad k = [1,2,\ldots,K] \quad (4)$$

where P_k represents the k^{th} feature map, and K is the total number of channels and W, H are the spatial dimensions (width and height, respectively).

To evaluate the deep features, we resize the deep feature maps to pixels. Each feature map is summed by a column to obtain the aggregated deep features $P' \in \mathbb{R}^{K \times 144}$. We utilize the Euclidean distance as the distance metric between the low-level feature $\mathbb{E}$ and deep featuresP', as shown in Eqn(5):

$$D_k(P'_k, \mathbb{E}) = \sqrt{(P'_k - \mathbb{E})^2} \quad (5)$$

$$s.t.\, k = [1,2,\ldots,K]$$

The elements $\{P'_k\}_{k=1}^{K}$ that yield the smaller distance within $\{D_k\}_{k=1}^{K}$ can be considered as more similar to $\mathbb{E}$, and the ranked list is denoted as $Rank\{r\}_{r=1}^{K}$. The deep feature maps with smaller distance values contain more important information.

A pair of serial Dense Blocks is used as the low-level feature extraction module in order to improve the richness of low-level features and conserve computational resources, as illustrated in Fig. 4. The entire network's input is where it is. The original image serves as the module's input, while its output is supplied into the HR Block. In this work, each Dense Block has six 3X3 convolution layers for feature interaction. Following first Dense Block processing of the input image, the output is replicated into two routes. A path's feature maps are transmitted to the transition layer, while low-level features from another path are connected to the decoder's input. The 2X2 average pooling layer cuts the image's height and breadth in half after the transition layer. As a result, as compared to the original image, the low-level feature resolution of the output of the second block is reduced to a fourth.

3.4 Triple Attention Module

The CVDs classification model strives to solve various degradation problems in different scenarios jointly. Due to the hierarchical processing and integration of features at different spatial scales of datas by network models, multi-scale and scale recursive models can be practical solutions to this problem. At the same time, it is necessary to establish an efficient feature extraction program in the process of learning features.

However, when considering this method, there is a significant flaw: usually, the input feature is spatially decomposed into a pixel through pooling layers to calculate the weight of these channels, which results in a significant loss of spatial information. Therefore, to improve the performance of ARB, Triple Attention Module (TAM) is proposed in the CNMS, which aims to refine the detailed information by concentrating on channel and spatial information and concentrating residual features well on these two types of information. As shown in Fig. 5, given an input tensor $X \in R^{C\times H\times}$, the Z pooling is responsible for reducing the tensor of the channel dimension to 2 dimensions, connecting the average and maximum aggregation features on that dimension. This allows the layer to retain a rich representation of the actual tensor while reducing its depth to make further calculations lighter. It can be expressed as follows:

$$Z\,\text{pooling}(X) = \left[\text{MaxPool}(X), \text{AvgPool}(X)\right]$$

First, the input tensor X is transferred to the three branches of TAM, and the first two branches establish interactions between the H, , and C dimensions, respectively. The concatenated maximum pooling is used to enrich the extracted high-level features. The third branch transforms the feature graph from size (N, C, H,) to size (N, C, 1, 1) through pooling layers, to achieve global context information fusion. Therefore, each unit can learn the contextual and localization information that is concerned with the channel axis and spatial axis separately, in order to accurately restore edges and textures[30-31].

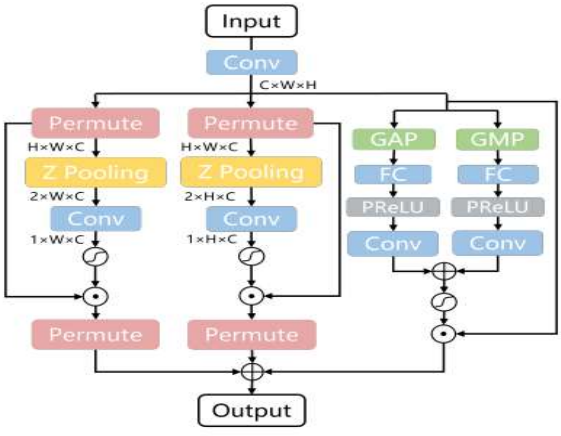

***Fig 5**: Architecture Of Triple Attention Module (TAM).*

The triple attention module composes of three branches to refine features. After going through TAM, the informative features are emphasized, and the less useful features are suppressed in the input features. Generally, the attention mechanism used to calculate channel attention first calculates the weight value and then uses the weight value to uniformly scale the feature map.

3.4 Multi-Scale Context Fusion Block

Continuous convolution and pooling operations can detect feature maps with different scales. Due to the fixed network layers, it has some shortages on multi-scale context extraction. Faced with this issue, a MCF block is proposed for context enhancement to obtain multi-scale spatial context information. Combined with the atrous convolution, it offers a practical way to acquire various receptive fields. Additionally, it produces fewer model parameters than traditional convolution, such as 5*5 convolution, 7*7 convolution, etc. Inspired by this advantage, atrous spatial pyramid pooling (ASPP) block has been proposed to detect multi-scale contexts. However, it exists the grid effect that some image pixels are in the visual blind area, and could not be involved in the information loss to cause the information loss(see Fig.6). Inspired by As the PP block, by stacking four 3*3 atrous convolutions, the MCF block is proposed for CVDs classifications For the MCF block, combined with atrous convolution, four parallel network branches are built for multi-scale feature representation, and the dilation rates for four atrous convolution layers are set as 1, 3, 5, and 7. The specific mathematical description is given in Eq.(6).

$$\begin{aligned} H_{c1} &= F_{d1}(H_a) \\ H_{c2} &= F_{d3}(H_{c1}\Theta H_a) \\ H_{c5} &= F_{d5}(H_{c2}\Theta H_{c1}\Theta H_a) \\ H_{c4} &= F_{d7}(H_{c5}\Theta H_{c2}\Theta H_{c1}\Theta H_a) \end{aligned} \quad (6)$$

The 3*3 atrous convolution of the dilation rate (CODR) of 1 is represented by Fd1, the 3*3 atrous CODR of 3 by Fd3, the 3*3 atrous CODR of 5 by Fd5, and the 3*3 atrous CODR of 7 by Fd7. The four output feature maps from the four parallel network branches are represented by the letters (Hc1, Hc2, Hc3, and Hc4), whereas Ha stands for the input feature maps. Meanwhile, the global pooling layer is also introduced to acquire the average of the global contexts. The final output of the proposed MCF block is defined as per Eqn (7).

$$H_m = G(H_a)\Theta H_{c4}\Theta H_{c3}\Theta H_{c2}\Theta H_{c1}$$
(7)

where G represents the network branch with global pooling.

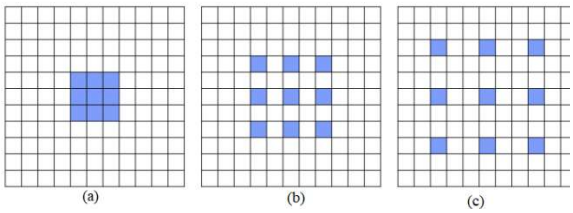

*Fig. 6: Grid Effect On Cascaded Atrous Convolution. (A) 3*3 Atrous Convolution; (B) Two Cascaded 3*3 Atrous Convolution; (C) Three Cascaded 3*3 Atrous Convolution.*

3.6 Loss Function

The loss function employed in the proposed medical image segmentation solution is the Jaccard Loss, which complements the Jaccard Coefficient by evaluating the dissimilarity between two samples. The Jaccard Coefficient is subtracted from 1 to get the Jaccard Loss. The Jaccard Index is computed by dividing the intersection size by the union size. The formula for Jaccard Coefficient has been given in Equation (**8**).

$$Jaccard\ Index = \frac{A \cap B}{A \cup B} = \frac{A \cap B}{A + B - A \cap B}$$
(8)

4. MATERIALS AND EXPERIMENTAL RESULTS

MIT-BIH and PTB-ECG (Goldberger et al., 2000; George Moody and Mark Roger, 2001) are the two datasets that are utilized to obtain ECG samples [21–22]. The MIT-BIH ECG datasets from 2005 and the PTB-ECG datasets from 2004 contain the specifics of these two datasets [23]. Table 1 displays the specifics of the numbers of training sets (70%) and testing sets (30%), as well as the numbers of normal and atypical cases. This table makes it clear that there is an imbalance in the number of normal and atypical instances selected.

Table 1; The Normal And Abnormal Cases Of Heart Disease.

Datasets	No of ECG datasets	Abnormal Cases	Normal cases
MIT-BIH	268	104	164
PTG-ECG	200	54	146

4.2. Evaluation Metrics

Several popular metrics are adopted to evaluate the performances, including the Precision, Specificity, Accuracy. Dice Similarity Coefficient (DSC) , Intersection over Union (IoU) . It is given in the following Eqns(9)

$$\Pr = \frac{TP}{TP + FP}$$
$$Sp = \frac{TN}{TN + FP}$$
$$Acc = \frac{TP + TN}{TP + FP + TN + FN}$$
$$F1 = 2.\frac{Se.\Pr}{Se + \Pr}$$

Where T P, TN, FP, and FN denote true positive, true negative, false positive and false negative, respectively.

$$\text{IoU} = \frac{A \cap B}{A \cup B},$$
$$\text{DSC} = \frac{2\,|A \cap B|}{|A| + |B|},$$
(9)

Where A and B represent pixel sets for the ground truths and their detection results, respectively

4.3 EXPERIMENTAL RESULTS

The experimental results of the proposed DIAN-NET method using in MIT-BIH dataset are given in Table 2and the PTB-ECG dataset is given in Table 3.

Table 2. Comparison Of The Different Cvds Classification Methods In The MIT-BIH Dataset.

Method	IoU	DSC	Acc	Se	Sp	Pr
UNet [13]	0.932	0.9646	0.9903	0.96	0.99	0.965
ResUNet [32]	0.918	0.9562	0.9883	0.96	0.99	0.95
DeepLabV3 [29]]	0.93	0.9666	0.9910	0.97	0.99	0.960
ResUNet++ [30]	0.932	0.9646	0.9675	0.97	0.9545	0.9656
CardioNet [17]	0.907	0.9490	0.9270	0.94	0.9124	0.9434
CycN-Net [15]	0.868	0.9281	0.989	0.97	0.99	0.896
Pranet [33]	0.910	0.9400	0.9534	0.96	0.94	0 .95
Proposed DIAN-NET	**0.947**	**0.9725**	**0.9928**	**0.97**	**0.99**	**0.969**

Comparing the proposed method with other modified ResUNet architectures [22-24,33] in Table 2 shows that the proposed method performs well on the MIT-BIH test dataset. This is because there is a relatively large gap between the quantitative criteria obtained by the architectures designed in [22-24,33]. However, the CardioNet [17] and CycN-Net [15]architectures have better performance than the proposed **DIAN-NET**, as the formers have filters of different sizes that can retrieve spatial information more accurately. On the other hand, as reported in Table 3, the proposed architecture achieves good performance compared to other ResUNet-based architectures due to the use of the transfer learning technique[34-35].

Table 3. Comparison Of The Different Cvds Classification Methods Methods In The PTB-ECG Dataset.

Method	IoU	DSC	Acc	Se	Sp	Pr
UNet [13]	0.91	0.95	0.946	0.93	0.93	0.92
ResUNet [32]	0.87	0.92	0.980	0.90	0.99	0.95
DeepLabV3 [29]]	0.89	0.94	0.983	0.94	0.99	0.94
ResUNet++ [30]	0.92	0.93	0.943	0.93	0.94	0.93
CardioNet [17]	0.88	0.94	0.97	0.98	0.96	0.94
CycN-Net [15]	0.88	0.93	97.21	0.98	0.93	0.94
Pranet [33]	0.92	0.93	0.94	0.97	0.91	0.9567
Proposed DIAN-NET	**0.93**	**0.96**	**0.989**	**0.9**	0.9	0.95

4.4 ABLATION EXPERIMENTS: Similar to this, the ablation study is also carried out on the suggested MFDCN to confirm the performance enhancement function, demonstrating the efficacy of the dense attention block, TA block, and MCF block. The ablation experiments of the suggested DIAN-NET on the PTB-ECG Dataset are shown in Fig. 6.

Fig. 6. Ablation results on propose method on PTB-ECG Dataset.

5. CONCLUSION

Heart problems, also referred to as cardiovascular diseases, are the leading cause of death worldwide. The sooner they are detected and classified, the more lives can be saved. An extensively used, inexpensive, noninvasive technique for evaluating the electrical activity of the heart and detecting cardiovascular diseases is the electrocardiogram, or ECG. In this work a Multi-Level Features Based Deep Convolutional Neural Network for Classification of Cardiovascular Diseases using ECG Signals The proposed DIAN-NET is a modified ResUNet architecture consisting of a feature extractor based on the pre-trained Xception architecture and an additional decoder called the middle decoder. The results of the experiments showed the desirable performance of the proposed architecture in comparison with other studies. By using the pre-trained Xception architecture, not only the evaluation criteria were improved, but also the generalizability capability of the proposed architecture on the unseen dataset was increased. Besides, with the help of the middle decoder, the affected area was gradually refined. Future work on the task of CVDs classification can provide a more diverse dataset or can also extend this study to other classification tasks.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: MIT-BIH ECG datasets - https://physionet.org/content/mitdb/1.0.0/ and PTB-ECG datasets - https://www.physionet.org/content/ptbdb/1.0.0/.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

REFERENCES

[1] World Health Organization (WHO), "Cardiovascular diseases," Jun. 11, 2021. Accessed: Dec. 27, 2021. [Online]. Available: https://www.who. int/health-topics/cardiovascular-diseases

[2] Government of Westren Australia, Department of Health, "Common medical tests to diagnose heart conditions," Accessed: Dec. 29,2021. [Online]. Available: ttps://www.healthywa.wa.gov.au/Articles/A_E/Common-medical-tests-to-diagnose-heart-conditions

[3] M. Swathy and K. Saruladha, "A comparative study of classification and prediction of cardio-vascular diseases (CVD) using machine learning and deep learning techniques," ICT Exp., to be published, 2021. [Online].

[4] R. R. Lopes et al., "Improving electrocardiogram-based detection of rare genetic heart disease using transfer learning: An application to phospholamban p.Arg14del mutation carriers," Comput. Biol. Med., vol. 131, 2021, Art. no. 104262.

[5] R. J. Martis, U. R. Acharya, and H. Adeli, "Current methods in electrocardiogram characterization," Comput. Biol. Med., vol. 48, pp. 133–149, 2014.

[6] Abdalla Fakheraldin, Wu Longwen, Ullah Hikmat, Ren Guanghui, Noor Alam, Zhao Yaqin. ECG arrhythmia classification using artificial intelligence and nonlinear and nonstationary decomposition. Signal Image Video Process 2019:13.

[7] Jayachandran, A & Dhanasekaran, R, "Severity Analysis of Brain Tumor in MRI Images using Modified Multi-Texton Structure Descriptor and Kernel- SVM, The Arabian Journal of science and engineering October 2014, Volume 39, Issue 10, pp 7073-7086,(2014).

[8] Fan Xiaomao, Yao Qihang, Cai Yunpeng, Miao Fen, Sun FM, Li Ye. Multi-scaled fusion of deep convolutional neural networks for screening atrial fibrillation from single lead short ECG recordings. IEEE J Biomed Health Inform 2018:1.

[10]. Jayachandran, A and R.Dhanasekaran ,(2017) 'Multi Class Brain Tumor Classification of MRI Images using Hybrid Structure Descriptor and Fuzzy Logic Based RBF Kernel SVM' , Iranian Journal of Fuzzy system , Volume 14, Issue 3, pp 41-54 , 2017.

[11]. Nozadi S. H., Kadoury S. Classification of Alzheimer's and MCI patients from semantically parcelled PET images: a comparison between AV45 and FDG-PET. International Journal of Biomedical Imaging. 2018;2018:13.

[12].Jiang, Z., Chen, Y., Zhang, Y., Ge, Y., Yin, F.F., Ren, L., 2019. Augmentation of cbct reconstructed from under-sampled projections using deep learning. IEEE Transactions on Medical Imaging 38, 2705–2715.

[13] Lee, H., Lee, J., Kim, H., Cho, B., Cho, S., 2018. Deep-neural-networkbased sinogram synthesis for sparse-view ct image reconstruction. IEEE Transactions on Radiation and Plasma Medical Sciences 3, 109–119.

[14]. Li, Y., Li, K., Zhang, C., Montoya, J., Chen, G.H., 2019. Learning to reconstruct computed tomography images directly from sinogram data under a variety of data acquisition conditions. IEEE transactions on medical imaging 38, 2469–2481.

[15] Zhi, S., Kachelrieß, M., Pan, F., Mou, X., 2021. Cycn-net: A convolutional neural network specialized for 4d cbct images refinement. IEEE Transactions on Medical Imaging 40, 3054–3064.

[16] R. Bharti, A. Khamparia, M. Shabaz, G. Dhiman, S. Pande, and P. Singh, "Prediction of heart disease using a combination of machine learning and deep learning," Comput. Intell. Neurosci., vol. 2021, 2021.

[17] A. Pal, R. Srivastva, and Y. N. Singh, "CardioNet: An efficient ECG arrhythmia classification system using transfer learning," Big Data Res., vol. 26, 2021.

[18] R. Avanzato and F. Beritelli, "Automatic ECG diagnosis using convolutional neural network," Electronics, vol. 9, no. 6, 2020.

[19] Miao, F., Cai, Y.P., Zhang, Y.X., Fan, X.M. and Li, Y., 2018. Predictive modeling of hospital mortality for patients with heart failure by using an improved random survival forest. *IEEE Access*, *6*, pp.7244-7253.

[20] Jinny, S.V. and Mate, Y.V., 2021. Early prediction model for coronary heart disease using genetic algorithms, hyper-parameter optimization and machine learning techniques. *Health and Technology*, *11*(1), pp.63-73.

[21] Ali, L., Niamat, A., Khan, J.A., Golilarz, N.A., Xingzhong, X., Noor, A., Nour, R. and Bukhari, S.A.C., 2019. An optimized stacked support vector machines based expert system

for the effective prediction of heart failure. *IEEE Access*, *7*, pp.54007-54014.

[22] Ramesh, G., Madhavi, K., Reddy, P.D.K., Somasekar, J. and Tan, J., 2021. Improving the accuracy of heart attack risk prediction based on information gain feature selection technique. *Materials Today: Proceedings*.

[23] Mienye, I.D., Sun, Y. and Wang, Z., 2020. Improved sparse autoencoder based artificial neural network approach for prediction of heart disease. *Informatics in Medicine Unlocked*, *18*, p.100307.

[24] Gokulnath, C.B. and Shantharajah, S.P., 2019. An optimized feature selection based on genetic approach and support vector machine for heart disease. *Cluster Computing*, *22*(6), pp.14777-14787.

[25] Gárate-Escamila, A.K., El Hassani, A.H. and Andrès, E., 2020. Classification models for heart disease prediction using feature selection and PCA. *Informatics in Medicine Unlocked*, *19*, p.100330.

[26] Nagarajan, S.M., Muthukumaran, V., Murugesan, R., Joseph, R.B., Meram, M. and Prathik, A., 2021. Innovative feature selection and classification model for heart disease prediction. *Journal of Reliable Intelligent Environments*, pp.1-11.

[27] K. Sun, B. Xiao, D. Liu, J. Wang, Deep high-resolution representation learning for human pose estimation, in: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 5686–5696. doi:10.1109/CVPR.2019.00584.

[28] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, IEEE.

[29] B. Quan, B. Liu, D. Fu, H. Chen, X. Liu, Improved deeplabv3 for better road segmentation in remote sensing images, in: 2021 International Conference on Computer Engineering and Artificial Intelligence (ICCEAI), 2021, pp. 331–334. doi:10.1109/ICCEAI52939.2021.00066.

[30].Jha D, Smedsrud PH, Riegler MA, Johansen D, De Lange T, Halvorsen P, et al., editors. Resunet++: An advanced architecture for medical image segmentation. 2019 IEEE International Symposium on Multimedia (ISM); 2019: IEEE.

[31] J. Xu, O. Chutatape, P. Chew, Automated optic disk boundary detection by modified active contour model, IEEE Trans. Biomed. Eng. 54 (3) (2007) 473–482. doi:10.1109/TBME.2006.888831.

[32] Y. Zheng, X. Zhang, X. Xu, Z. Tian, S. Du, Deep level set method for optic disc and cup segmentation on fundus images, Biomed. Opt. Express 12 (11) (2021) 6969–6983.

[33].Fan D-P, Ji G-P, Zhou T, Chen G, Fu H, Shen J, et al., editors. Pranet: Parallel reverse attention network for polyp segmentation. International conference on medical image computing and computer-assisted intervention; 2020: Springer.

[34].Mahiba C, A Jayachandran,"Severity analysis of diabetic retinopathy in retinal images using hybrid structure descriptor and modified CNNs",Measurement,Vol 135,PP 762-767,2019.

[35]. Jesu Prabhu A and Jayachandran, A, "Mixture Model Segmentation System for Parasagittal Meningioma Brain Tumor Classification based on Hybrid Feature Vector" Journal of Medical System, vol 42, issues 12, 2018.