International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

Hybrid Auto Encoder for Dimensionality Reduction in Ct Scan Images of Heart

Jayaram M[1*], J. Lakshmi Narayana[2], A. Pathanjali Sastri[3], Santhi Sri. T[4]

Submitted: 12/11/2022 **Accepted:** 13/02/2023

Abstract: Modern style of food habit system and stress life has completely impacted the human life. Sudden deaths irrespective of age have become most common now days. In order to save human life, predict heart attack system many of the existing systems are working on the textual attributes but in the proposed model, the major focus of the heart attack prediction is based on the CT SCAN, continuous monitoring of which helps the doctors to diagnosis the patients effectively. The proposed model along with the identification of the heart attack, it also analyzes the heart beat rate time to time based on the sensor. This analysis helps the doctor to recognize the part that is blocked and that makes the patient to suffer from cardiac arrest or survival problems in future. Traditional approaches use either principal component analysis or neural networks for identifying those parts but the proposed model extends the de-noising auto encoder that has the capability to prevent similar identity learning features. The model has achieved +3.1% accuracy improvements when compared to the traditional CNN.

***Keywords**: De-noising Encoder, Convolution Neural Network, Sensors, Decoder, Identity Function*

1. Introduction

Heart attack is one of the life impacting medical problems. Doctors can analyze the condition of heart using the CT-SCAN images. Many researches worked on heart attack detection using the traditional neural networks but these systems cannot recognize the blocks in the valves. The proposed model in order to identify the blocked components of the heart performs the feature extraction using the encoders and decoders. Complex images like MRI, CT, UV Scan and other systems need high processing neural network with more number of layers. Most of encoders suffer from a problem known as "Null Identity", when the network designs the layers with more number of nodes than required [1]. The model can focus on the required region by implementing auto encoders.

Auto Encoder:

An artificial neural network called an Auto encoder is used to learn effective codings for unlabeled input (unsupervised learning). By attempting to recreate the inputs from the encoding, the encoding is verified and improved. Auto encoders are mainly for reducing the dimensionality of the neural networks which has recently been popular [2]. This method has high performances in non-linear data compared to linear data. Generally, this method is utilized when the data is compressed represented in the input like image, text, etc. The process of the Auto encoders initialized with acquiring the dataset and defining Auto encoder accordingly, where the dataset is vast the parameters are also high for training so, GPU methods are considered to speed the training [3]. Next, balancing the loss function, the data need to be trained and compiled. The data has to visualize compressed digits from the dataset after auto encoding. Testing is performed on the visualization of latent space. N number of digits is formed where most are separated with similar clusters. But some digits can be overlapped and have differently arranged. The presences of Auto encoders are without the activation functions performed in the hidden layer the representation will be worst and not determined. In this many types of data can be represented with high performances[4]. The two primary components of an Auto encoder are an encoder that converts a signal to code as well as a decoder that extracts the text from the code. Different Auto encoders are defined below Fig.1.

[1] Professor, Department of CSE(Data Science), Sreyas Institute of Engineering & Technology, Hyderabad, India
[2]Professor , Department of ECE, PSCMR College Of Engineering and Technology, Vijayawada, A.P.
[3]Professor, Department of CSE, PSCMR College Of Engineering and Technology, Vijayawada, A.P
[4] Professor, Department of Computer Science and Engineering, Koneru lakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh.
jayaram_258m@yahoo.com,drjln@pscmr.ac.in, akellapatanjali@yahoo.com , sri_santhi2003@yahoo.com
** Corresponding author's Email: jayaram@sreyas.ac.in*

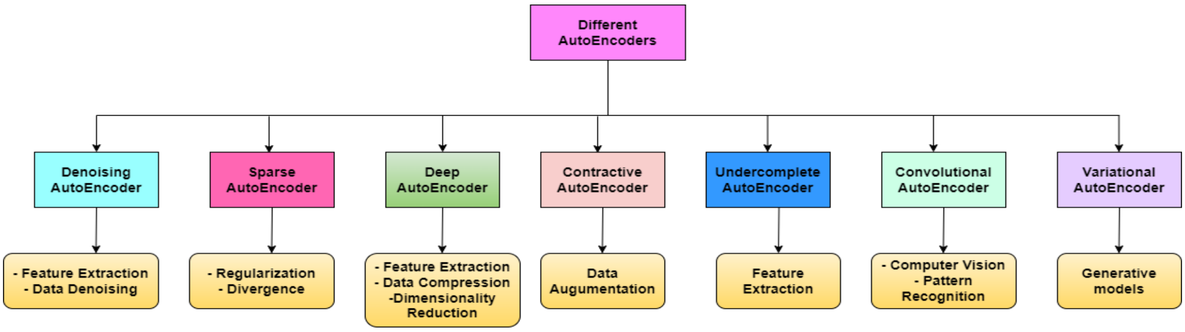

Fig.1: Classification of Encoder and Decoders

Many researchers have implemented **sparse encoder** to store the information about the images because it has the capability to store huge amount of information but the use of bottleneck connections is expressly discouraged by this kind of Auto encoder. As a result, the model becomes more regular suffers overfitting the data. To create a global loss function, this "sparsity penalty" is applied to the reconstruction loss[15].

Another advanced approach is **"Deep Auto Encoder",** with two symmetrical deep belief networks make up deep Auto encoders. The general structure is comparable to this structure. Two constrained Boltzmann machines to serve as decoders and encoders to make up these mirrored components. These systems suffer from the more resource consumption problem because of the server architecture.

The proposed model utilized the "**De-noising Auto encoder" [5].** These auto encoders are designed to efficiently encode noisy data, removing random noise from the code. In doing so, the Auto encoder is intended to produce a de-noised output that is distinct from the input [6]. In the world of medicine, it is very important to understand the data[7]. It is possible only with clear images. De-noising techniques helps the model in handling the random noise very effectively[8].

In this paper, introduction section discusses about the auto encoders which the proposed model wants to fuse. Literature survey section discusses about the techniques implemented by the existing approaches. It also frames the demerits to identify the improvement functionality of the proposed methodology. The proposed methodology presents the fusion model of different encoders to reduce the compression loss. The results and discussion projects the layered architecture along with the comparison between the approaches of both existing and proposed. Finally, conclusion section discusses the drawbacks of PCA and advantages of fusion auto encoders.

2. Literature Review

In [9], Hamada R. H. Al-Absi et Al suggest using deep learning (DL) to separate the CVD cluster from the comparison group. The areas of focus in the retinal pictures that most significantly influenced the choices made by the suggested DL model were highlighted using the gradient class feature map (GradCAM) in the proposed method. To illustrate the role of imaging modalities in identifying CVD within the Qatari population, the authors concentrated on two: DXA scans and retinal fundus pictures. 500 participants total, equally split between the CVD group and the control group, make up the dataset. The median value was used to fill in the missing numbers. The min-max normalization method was then used to normalize the data. The scientists evaluated the hybrid approach on samples that were divided by age and gender, but didn't get any satisfactory outcomes. The results are particular to Qataris and residents of other gulf nations who share their ethnicity and way of life.

In [10], Yucheng Song et al suggest a supervised learning model. A fundamental kind of neural network is called a feed-forward neural network (FFNN). The medium levels are buried, with the input layer at the top and the output layer at the bottom. There is no response in the entire network as the selected sample is linearly from the input to the output layer. It may be challenging to accurately segment (BVS) the cardiac biventricular area due to the significant variability of the right ventricular anatomy and the absence of labeled data. Medical personnel can now execute high-precision CAD smart diagnosis because of endless opportunities for the autonomous segmentation of medical images made possible by the application of deep learning (DL). When paired with data from a specific domain, this method demonstrates the effectiveness of CNN modeling in clinical segmentation. A brand-new loss function built on edge data was also proposed.

In [11], Tulasi Krishna Sajja and Hemantha Kumar Kalluri supported the use of convolutional Neural Networks (CNN) as a potential early disease prediction tool. The input layer is followed by a convolution layer with 16 kernels and an activation function acting as a ReLU in the proposed convolutional architecture. The proposed network handles a significant amount of data. Eight kernels with the previously mentioned parameters were used for the convolutional layer, and a 25% dropout layer was also used. Add an output layer to calculate prediction probability. The proposed network obtained testing and training accuracy of 94.78% and 95.04 percent, respectively. The dataset from the UCI repository for machine learning is used for testing and making predictions about cardiovascular disease (CVD). The same dataset is examined using a variety of machine learning classifiers, including simple neural networks, linear and RBF kernel-based SVM, NB, KNN, and Logistic Regression (LR). All models in comparison fared better than the suggested model.

In [12], Chen et al create a novel quantitative algorithm to considerably improve the accuracy of estimations of stenosis severity and to overcome the difficult obstacle of quantifying minor stenosis underneath the image resolution. A 3D myocardial regional shortening (RSCT) determined from CT is a superb quantitative classifier that is shown to find regional myocardial wall motion anomalies in cases of cardiac dysfunction. Initially, CT was used to assess coronary artery stenosis. Unsupervised learning, which does not need labeled training data, is the solution. This limits clinical application by requiring specialist viewing software and mechanical processing that could result in reader variability. The unmet demand is therefore for a fully automatic method to carry out these two image-processing operations. A DL framework is employed to automatically and concurrently execute multi-chamber classification and cardiovascular imaging plane forecasting to meet this need.

In [13], Nikolaos Papandrianos et al created a powerful CNN model to identify echocardiographic images by identifying the illuminating elements from an image and correctly classifying them with their help. By looking at whole-body scans, an RGB-based CNN architecture has been suggested to automatically determine if a patient has malignant cells or not. The performance of the novel models is also examined by the authors who perform the necessary parametric and regularisation to address myocardial perfusion imaging detection from SPECT images of patients with ischemic or infarction. At most 85% of such age-predicted maximal heart rate was reached before the Bruce protocol was applied. Data normalization, data shuffling, data augmentation, and data division into training, verification, and testing comprise the pre-processing step. Within this architecture, a deep-layer system is established, consisting of five convolutions, two dense layers, pooling layers, a dropout layer, and a final 2 output layers. Adam Optimizer and the ReLU activation function are employed.

In [14], G. Maragatham et al built a traditional predictive temporal model that is linked to longitudinal time-stepped electronic health records (EHR) and uses recurrent neural networks (RNN) like LSTM. SiLU and tanh are used as training algorithms in the hidden layers of the proposed LSTM model, and Softmax is used as an activation function in the output layer. In the entire network, bridgeout is employed as a regularisation method for weight optimization. Evaluations based on real-time data show that the recommended model is useful and feasible for predicting heart failure risk. To modify the weights of relevant layers at each layer, a gradient descent method is applied. The hidden layer computes the weight derivatives using the back propagation algorithm. In this suggested work, supplier comments are not used. LSTM models are used to account for varying patient counts, with 10 being the highest extreme gradient reasonable value for each epoch. It can be used in various healthcare application domains. Table 1 presents the analysis on the existing approaches

Table 1: Existing System Analysis with Merits & Demerits

Author	Method	Merits	Demerits	Accuracy
Hamada R. H. Al-Absi	Deep Learning	multi-modal approach,	Need to be validated using an exterior clinical setup.	80.3%
Yucheng Song	deep learning model (FFNN)	A new loss function based on edge detection is used.	Not scalable, and only works on sparse datasets.	92%
Tulasi Krishna Sajja	CNN	ReLU is used and can be performed on large datasets.	Complex and high computation cost.	95.04%
Chen	DL	No need for labelled data, and two image	The multi-chamber classification should	93.1%

		processing operations.	be used before, mechanical processing	
Nikolaos Papandrianos	CNN	Bruce protocol is followed.	The age factor is not ranged correctly.	93.47%
G. Maragatham	RNN –LSTM	Scalable works on real-time data	Should also be investigated for other diseases.	79.7%

3. Proposed Methodology

In traditional de-noising neural network to solve the “Null Function” problem, the system identifies the number of nodes that are greater than required and assigns their random weights as “zero”. In some cases, this may give the efficient results but while calculating the loss function; it generates negative values which make the model to suffer from more number of misclassification rates than required. The de-noiser is efficient in handling the noisy data, but the problem with these type of systems is these cannot handle the less amounts of noise in the data. Fig. 2 represents the ways for de-noiser to work with original data.

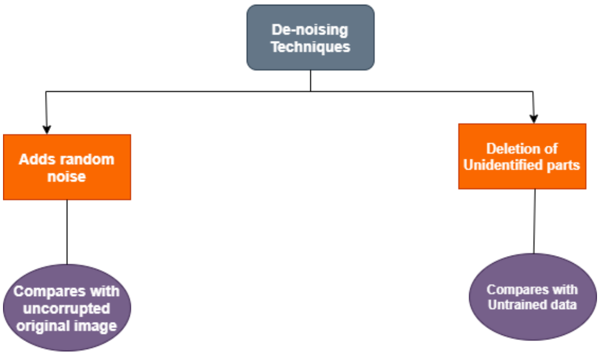

Fig.2: De-noising Induction Techniques

In the proposed research to identify the block passages of values, the model needs to simulate the similar type of original images in different views, which is quite possible with deep auto encoders. The main advantage of these encoders is to store the compressed version of original image with minimal loss. The deep CNN assigns some constraints which help the model to learn the essential features quickly from the latent representation generated. The decoder using these extracted features reconstructs the new image, which is almost similar to the original image. The proposed model constructs the encoder with 2 layers and decoder with 3 layers. In traditional approaches, the number of layers on encoder and decoder are implemented same in size but here in the proposed methodology with the varying dimensionality and using the advantages of deep auto encoder the model implements different sized neural networks. So the proposed model combines the de-noising with deep auto encoders as shown in Fig.3.

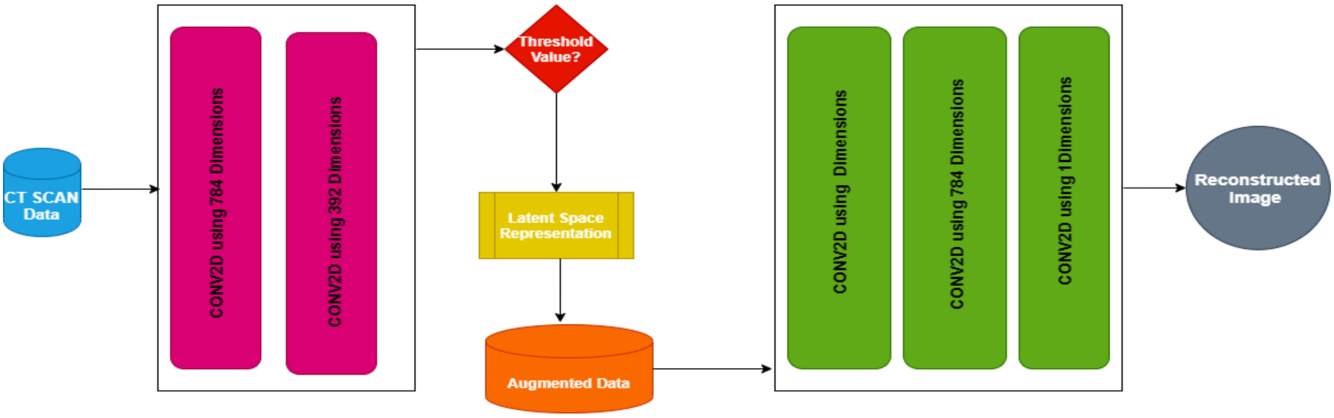

Fig.2: Block Diagram for Extended De-noising Network

3.1. Pseudocode for Extended De-noising Network:

Input: Load the Heart_CT images, HCT

Output: Segmented part of heart

Begin:

```
for i←0 to len(HCT):
noisy_data←clip(HCT[i], random(0,255))
for k←0 to padding_value:
    aug_image[i][k]←α(conv2d(Noisy_data[i][k]
    , kernel=3))
for j← 1 in len(res_block):
    Norm_image[j]←Batch_N(aug_image[j],
    act="LekayReLu",n_features)
for f in n_features:
    upsample_image[f]←MaxPooling(Nom_ima
    ge_inverse)
    print(upsample_image[f])
```

2. Layers in Auto Encoders

Auto encoders are used in neural networks for encoding and decoding data with having 4 layers, 1 code layer, & 4 decoders. Generally, the encoder is present in front of the LS where it encodes the information for compression to reduce dimensions. The info is retrieved from the fully-connected layer in neural networks. The compressed image is contorted from the real picture. Code is for reducing the representation in input that needs to be decoded [9]. The structure of the decoder and encoder are similar. This is responsible for reconstructing the real image send it to the input with the same dimensions through code. The encoding layers have four layers for deriving each element and transferring the code in single layers [7]. The data with a single layer can use some related codes and encode the format. Now the code needs to be decoded and send the output to the related approach. Here they have a fixed layer for the encoding and decoding process.

3. 3. Latent Space Representation

Latent space is an idea of spatial representation it can be understood through compressed data related to a particular object. To determine it in a simple way every data can not be remembered so, by encoding observations a figure is chosen and identify those objects which are related. Where in deep learning raw data is considered as an input and related features are represented in output these features are further classified into three different tasks. After retrieving data it should be derived in a low-dimensionality format so, classification, reconstruction, or regression actions can be performed. The performance of this method is high in three areas image, word, & GAN. In the image, extraction supposes the CNN method is considered after the completion of fully-connected layers and ready for high-level extraction in the LS. LS contains similar images related to the raw data and it identifies with the distances vectors. In the word embedding the NLP is used to identify the synonyms of the word and each word can be encoded with low dimensionality. Here in LS, many other approaches can be used for good performances. The GAN is the responsible for underlying distribution in the dataset [8] . This model creates similar images which are related to the database. Here the input is chosen as latent vectors where the output is in low-dimensionality vectors.

3.4. Swish Activation Function

Swish activation is a function that overcomes the issue of inconsistent gains related to ReLU. The ReLU method is highly used in neural networks and has achieved good performance. Swish is a very simple method in which a function of x is equal to the x. sigmoid(x) values. Non-monotonic is the other name for this process which can handle two kinds of domains i.e. classification of image & machine translation this process matches ReLu or outperforms it. Here two types of gating are allowed self and multi where in self as input a scalar is required and in multi, it needs two scalars. In the self gate, it is an activation of the sigmoid by itself. When optimization is tuff & layers are high then SAF will achieve the best performances. This is also proven with multiple irrelevant datasets and has increased their performances accordingly. Even compared with other classifications it has achieved almost 0.9% high accuracy.

3.5. Computation of Loss Function

The loss function is used in many learning approaches for receiving optimized results from the trained datasets. This method is used for error functions which can solve mathematical optimizers & make some decision theories. Here the cost function is also used for mapping values or events related to 1 or more variables which can represent some cost which is associated with the event. To calculate this method two types of values need to be evaluated intercept and slope. To simplify any method without metrics is not declared. If the model is lower than the good model then parameters need to be changed which can minimize loss. There are two kinds of functions one can work on a single trainee i.e. loss function and the other is considered an average loss on entire trained data sets this is the cost function. This process holds six types of functions regression, classification, Auto encoder, GAN, detecting objects, & embedding words. This is the best classifier for binary functions. Compared to categorical the sparse cross-entropy is faster.

4. Results & Discussion

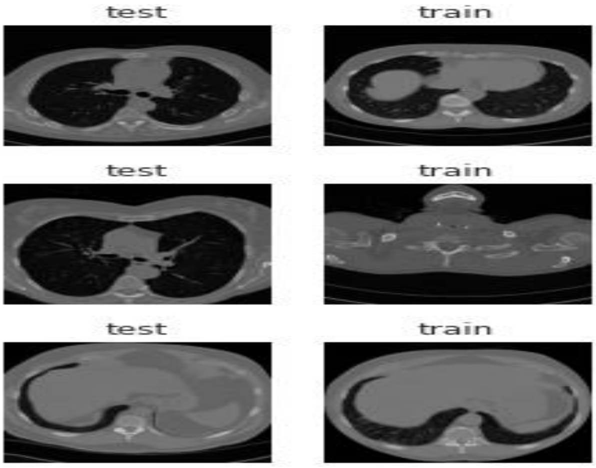

Fig.4: Different Images in dataset

Fig.4 presents different images available in the predefine dataset available in the kaggle repository. All different possibilities of images are available in both the training and test dataset. This makes the dataset as perfect balanced in terms of different augmentations. Implementing the de-noising techniques generates more number of corrupted copies of the original images as shown in Fig.5 by inducing random noise to the original images.

```
<matplotlib.image.AxesImage at 0x7f2316049340>
```

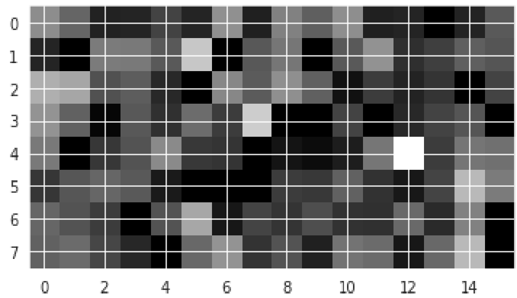

Fig. 5: Noise Induction in CT Scan Images

Fig.6 presents the re-constructed image, which is similar to original image and this representation helps the model to store the data in latent representation, which consumes less space while storing in memory.

```
[372.1247   266.69574  91.713776  96.36696  178.17781   96.698
 379.4988    87.23351 341.7673   251.65524  374.23917   85.88908
  98.85276    0.        93.91368 236.8612    95.41995    0.
 325.6723   320.50967 234.63599  525.9835      0.       232.04749
 310.95993    0.       234.79094 386.57874  123.964294 167.0848
 252.09384  224.69246 462.2809   437.6699   218.29317  246.52614
 106.76603    0.       356.88013 241.10698  378.356    255.02895
```

Fig.6: Vector Representation of Reconstructed Image

Fig.7 presents the summary of the proposed de-noising architecture. The proposed model on the decoder side contains three layers. The output layer presents the shape of the image after the layers in the network performs the necessary operations. From the results it is evident that the image is reconstructed with the same dimensionality and high number of non-trainable parameters.

```
Model: "model_2"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 input_3 (InputLayer)        [(None, 224, 224, 1)]     0

 model (Functional)          (None, 7, 7, 1)           492897

 model_1 (Functional)        (None, 224, 224, 1)       62711

=================================================================
Total params: 555,608
Trainable params: 553,702
Non-trainable params: 1,906
_________________________________________________________________
```

Fig.7: Summary of the Proposed Model

```
_________________________________________________________________
Model: "autoencoder"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 img (InputLayer)            [(None, 28, 28, 1)]       0

 flatten (Flatten)           (None, 784)               0

 dense (Dense)               (None, 128)               100480

 dense_1 (Dense)             (None, 128)               16512

 dense_2 (Dense)             (None, 784)               101136

 reshape (Reshape)           (None, 28, 28, 1)         0

=================================================================
Total params: 218,128
Trainable params: 218,128
```

Fig.8: Result Analysis Based on Traditional Encoder

In the traditional, the model has considered less input size which is not efficient to analyze the ROI segments. To achieve this, the model has implemented flatten layer with more number of dimensions and then extended the model with three layers of dense modules. The model to reduce the loss of information at last it has implemented reshapes operation. But from Fig.8 it is clear that the model is producing huge of loss in terms of both training and validation datasets.

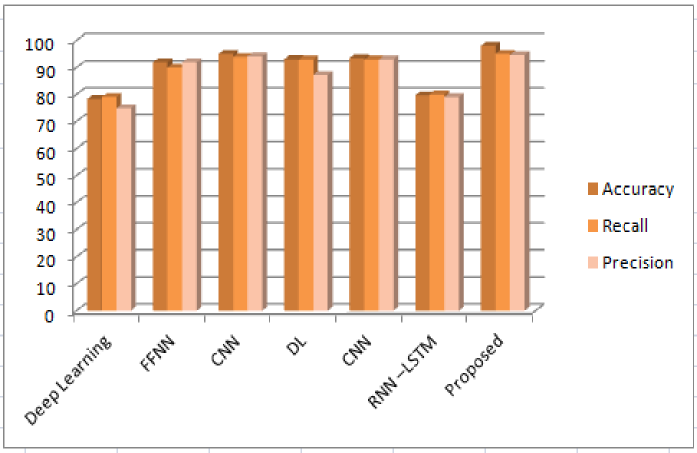

Fig. 9: Metrics on Different Datasets

Fig.9 represents the most curial parameters of the learning model i.e, accuracy, recall, precision on the existing approaches along with the proposed one. From the figure it is observed that proposed model is efficient in terms of all the evaluations. The model has acquired +3.1% more than the traditional CNN.

5. Conclusion

The famous approach for any segmentation is using the PCA technique but the auto encoders are good in creating a fusion technique in which the model doesn't omit the attributes instead it checks all the possibilities of combinations and fuses the one with high informative data. In all the existing works, most of the systems have failed due to capturing of complex blocks from the scanned images. With the proposed system, De-noising process on the segmentation of the infection regions by analyzing the complex structure and it performs background subtraction operation to detect the blocks easily. In future works, the model can focus on the multi classification of heart attacks using pre-trained models.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Author Contributions

Authors' contributions All authors contributed to the problem design, data collection, analysis and implementation. The first draft of the manuscript was written by [Jayaram M] and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript

References

[1] Diniz, J. O. B., Ferreira, J. L., Cortes, O. A. C., Silva, A. C., & de Paiva, A. C. (2022). *An automatic approach for heart segmentation in CT scans through image processing techniques and Concat-U-Net*. In Expert Systems with Applications (Vol. 196, p. 116632). Elsevier BV. https://doi.org/10.1016/j.eswa.2022.116632

[2] Sharobeem, S., Le Breton, H., Lalys, F. *et al. Validation of a Whole Heart Segmentation from Computed Tomography Imaging Using a Deep-Learning Approach*. *J. of Cardiovasc. Trans. Res.* **15**, 427–437 (2022). https://doi.org/10.1007/s12265-021-10166-0

[3] Li, F., Sun, L., Lam, K.-Y., Zhang, S., Sun, Z., Peng, B., Xu, H., & Zhang, L. (2022). *Segmentation of human aorta using 3D nnU-net-oriented deep learning*. In Review of Scientific Instruments (Vol. 93, Issue 11, p. 114103). AIP Publishing. https://doi.org/10.1063/5.0084433

[4] Nainamalai, V., Lippert, M., Brun, H., Elle, O. J., & Kumar, R. P. (2022). *Local integration of deep learning for advanced visualization in congenital heart disease surgical planning*. In Intelligence-Based Medicine (Vol. 6, p. 100055). Elsevier BV.

[5] Pieszko, K., Shanbhag, A., Killekar, A., Miller, R. J. H., Lemley, M., Otaki, Y., Singh, A., Kwiecinski, J., Gransar, H., Van Kriekinge, S. D., Kavanagh, P. B., Miller, E. J., Bateman, T., Liang, J. X., Berman, D. S., Dey, D., & Slomka, P. J. (2022). *Deep Learning of Coronary Calcium Scores From PET/CT Attenuation Maps Accurately Predicts Adverse Cardiovascular Events*. In JACC: Cardiovascular Imaging. Elsevier BV.

[6] Reddy, A. M., Reddy, K. S., Jayaram, M., Venkata Maha Lakshmi, N., Aluvalu, R., Mahesh, T. R., Kumar, V. V., & Stalin Alex, D. (2022). *An Efficient Multilevel Thresholding Scheme for Heart Image Segmentation Using a Hybrid Generalized Adversarial Network*. In S. Bhattacharya (Ed.), Journal of Sensors (Vol. 2022, pp. 1–11). Hindawi Limited.

[7] Sri Silpa Padmanabhuni and Pradeepini Gera, "*Synthetic Data Augmentation of Tomato Plant Leaf using Meta Intelligent Generative Adversarial Network: Milgan*" International Journal of Advanced Computer Science and Applications(IJACSA), 13(6), 2022.

[8] SRI SILPAPADMANABHUNI SRINIVASA REDDY K, P. VENKATESWARA RAO, A.MALLIKARJUNA REDDY, K SUDHEER REDDY, DR. J. LAKSHMI NARAYANA "*NEURAL NETWORK AIDED OPTIMIZED AUTO ENCODER AND DECODER FOR DETECTION OF COVID-19 AND PNEUMONIA USING CT-SCAN*" Journal of Theoretical and Applied Information Technology,15th November 2022. Vol.100. No 21,pp.no.6346-6360, 2022.

[9] Al-Absi, Hamada R. H., Mohammad Tariqul Islam, Mahmoud Ahmed Refaee, Muhammad E. H. Chowdhury, and Tanvir Alam. 2022. "*Cardiovascular Disease Diagnosis from DXA Scan and Retinal Images Using Deep Learning*" *Sensors* 22, no. 12: 4310.

[10] Song, Y., Ren, S., Lu, Y., Fu, X., & Wong, K. K. (2022). *Deep learning-based automatic segmentation of images in cardiac radiography: A promising challenge. Computer Methods and Programs in Biomedicine*, *220*, 106821.

[11] Sajja, T.K., Kalluri, H.K. (2020). *A deep learning method for prediction of cardiovascular disease using convolutional neural network*. Revue d'Intelligence Artificielle, Vol. 34, No. 5, pp. 601-606.

[12] Chen, Zhennong. "*Analysis of Cardiovascular Disease using Cardiac Computed Tomography and Deep Learning*." Ph.D. diss., UC San Diego, 2022.

[13] Papandrianos, Nikolaos, and Elpiniki Papageorgiou. 2021. "*Automatic Diagnosis of Coronary Artery Disease in SPECT Myocardial Perfusion Imaging Employing Deep Learning*" *Applied Sciences* 11, no. 14: 6362. https://doi.org/10.3390/app11146362

[14] Maragatham, G., & Devi, S. (2019). *LSTM Model for Prediction of Heart Failure in Big Data. Journal of Medical Systems, 43(5).*