Journal of Theoretical and Applied Information Technology

31st October 2023. Vol.101. No 20

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

LOCALIZED NETWORK MODEL FOR IMAGE DENOISING AND PREDICTION USING LEARNING APPROACHES

A.KARTHIKRAM[1], M.SARAVANAN[2,*]

[1]Department of Computer Science and Engineering, SRM Institute of Science and Technology, Kattankulathur, Chennai, India

[2]Department of Networking and Communications, Faculty of Engineering and Technology, SRM Institute of Science and Technology, Kattankulathur, Chennai, India

Email: [1]karthikram86@gmail.com, [2]saravanm7@srmist.edu.in

ABSTRACT

Computer Tomography (CT) imaging provides a promising solution for various health-based evaluations and diagnoses. Certain parametric mappings of cerebral parenchyma are performed with continuous CT scans. It is highly solicited to diminish the CT dosage for constant application owing to the higher radiation exposure due to continuous scans. Thus, there is a need of novel denoising and classification technique. Here, image denoising is essential to attain a reliable diagnosis. This research concentrates on modelling a novel deep learning approach with a Localized convolutional image denoising auto-encoder (L-CNM) for CT image denoising, which avoids the higher-dose referral images during the training process. The proposed network model is trained by mapping the image frames captured from CT and evaluating the adjacent frames. The noise over the CT source is independent, and the proposed model intends to eradicate the noise. The anticipated model can be easily adapted to the various real-time analyses as the model deals effectually with the high-dose training images. The proposed method is validated using the online available public dataset and simulated in MATLAB 2020a environment. The model attains improved image quality compared to various existing denoising approaches.

Keywords: *Computer Tomography, Denoising, Deep Learning, Autoencoder, Training Model*

1. INTRODUCTION

Two common medical imaging approaches that use pathological diagnosis tools, computed tomography (CT) and perfusion imaging, enable doctors to communicate with the body's organs more effectively. While photon starvation effects immediately affect the mixed noise with convoluted distribution in CT perfusion images, random noises formed during acquisition are more likely to degrade the quality of MR images [1]. With higher noise levels, the image quality can deteriorate, which can impact following analyses like classification, segmentation, and registration [2] are three examples. Therefore, it is crucial to establish an efficient way to eliminate MR and CT picture noise before processing. Numerous image-denoising techniques have been researched recently, including the Wiener filter based on spatial filtering, the Gaussian filter, wavelet-based filters, and nonlocal self-similarity models. Some academics pushed the idea of incorporating these methods into the medical picture, denoising problems [3].

Martin Fernandez et al. used MR images to reduce noise. They proposed an approach that determined the shrinking of wavelet coefficients based on the conditional likelihood that something is noise or detail [4]. To determine the de-noising process's parameters, which need not consider noise variance's impact; they employed the expectation maximization operation. For MR image denoising, A procedure based on a nonlocal means filter and a Laplacian of Gaussian filter is recommended in [5]. The denoising procedure will be accelerated to highlight the edge information, and they only chose a small number of comparable patches based on edge proximity. It can effectively maintain the image's edges and structural elements. An enhanced higher-order singular value decomposition (HOSVD) approach for denoising MR images was proposed in [6]. They employed the conventional HOSVD approach and then added a recursive process. The recursive operation repeated the

weighted summation of the noisy and clean picture filtering procedure. They validated acceptable weights for denoising studies by employing a variety of image formats and noise levels. However, these traditional methods can take a long time to process and could be more effective at addressing the convoluted distribution and Rician or low-dose perfusion noise [7]. Recently, algorithms based on deep learning have outperformed more conventional methods at denoising challenges. Deep convolutional neural network with residual learning CNN (D-CNN) [8] employs denoising and has attained cutting-edge performance.

Yet another method of reducing noise in CT scans is based on deep learning was published by Yang Q et al. [9]. To enhance the denoising effectiveness, they used the generative adversarial network (GAN) to utilize the two ideals that can statistically migrate the noise distribution and partially suppress noise using perceptual loss and Wasserstein distance approaches. A cascaded convolutional neural network (CNN) was created by Wu D et al. [10] to eliminate the artefacts that typical machine learning denoising techniques leave behind. The image's intricate residual noise can be further reduced and image quality improved using CNN. On datasets from low-dose CT, they compared the performance of single CNN structures with that of cascaded network architectures and discovered that the latter performed better. A technique based on CNN was recommended for low-dose CT image denoising by Chen H et al. [11]. Their approach successfully reduced noise and improved clarity by creating a mechanism for converting patches in low-dose CT images. However, the last convolution layer's supervisory information from earlier CNN-based denoising techniques needs to be improved, and they overlook the detailed data from earlier levels. Additionally, the denoising procedure will blur the image's fine details [12] – [15].

We suggest a network model (L-CNM) that eliminates noise using an auto-encoding model; subsequently, in low-dose CT images, noise can be decreased via the proposed network model. With the guidance of the last three convolution layers, L-CNM is built first to achieve precise noise predictions using the auto-encoding model. The L-CNM has several render blocks for understanding complex characteristics, a block for making hierarchical sounds, a block for fusing different latent sounds to create the final output noise, and a block for learning complex features. A refining network is introduced in the second stage to recover the image's lost information after L-CNM denoises it. The following are the contributions of this work:

1) We suggest merging an auto-encoding and classification network to form L-CNM. The proposed L-CNM restricts multi-supervision data to improve denoising effectiveness. When L-CNM denoises data, the proposed network can recover lost features.

2) We use a progressive training technique in L-CNM is trained cooperatively with a refinement network to create better noise estimations that can improve network performance. First, the independent prediction of the initial noise is taught to L-CNM.

3) The resilience of our model is demonstrated by L-CNM success in situations for both specified and blind noise levels. The proposed model works well in identifying the noise over the provided image effectually compared to other existing approaches with better prediction outcomes.

The work is organized as follows: section 2 gives a wider analysis of various prevailing approaches. In section 3, the proposed L-CNM model is elaborated, combining auto-encoding and prediction. The numerical outcomes are discussed in section 4, and the conclusion is given in section 5.

2. RELATED WORKS

Several current techniques for CNN picture denoising will be covered in this section. We distinguish between two methods for CNN picture denoising: There are two types of CNN denoising: (1) CNN for broad-based imagery and CNN for specialized images. In the first method, general images are denoised using CNN architectures, whereas, in the second method, specific images are denoised using CNN [16]. Compared to the second technique, the first is more frequently utilized in CNN denoising applications. Pictures that serve a general function rather than a specific one are called "generic images" [17]. Images purposely made with a special or specific kind are referred to as specific images. Examples of specific images include remote sensing, infrared technology, and medical imaging. To update readers on the most recent CNN architecture about the image, CNN denoising is broken down by image category.

The reference suggested the attention-guided CNN (ADNet) for picture denoising [18]. The feature improvement block, the sparse block, the attention block (FEB), and the reconstruction block are the four blocks that make up the 17 layers of the

ADNet (RB). The SB was utilized to increase performance and efficacy and to shallow the denoising framework because sparsity has been shown to work effectively when applied to images [19]. Conv+BN+ReLU and dilated Conv+BN+ReLU are the two types that make up the twelve layers of the SB. Comparatively, the FEB contains four layers of three. The AB only has one layer of convolution, but the other two have (Conv+BN +ReLU, Conv, and Tanh) Multiple types. The AB directed the FEB and SB, which is good for background noise. To paint a precise picture, the RB does reconstruction last. Model training was created using the mean square error [20].

While certain deep learning algorithms yield fantastic results when using artificial noise, most networks do not perform well when dealing with realistic noise-corrupted images. The noise estimation reduction network was proposed in (NERNet). On photos with realistic noise, NERNet reduced noise. The modules for noise estimation and noise reduction were the two modules that made up the architecture. The noise estimation module integrates the pyramid feature fusion [21] and symmetric dilated block [22] – [23] into the noise-level map. The removal module, meanwhile, eliminated noise by utilizing the map of the projected noise levels. The removal module combined global and local information to preserve texture and minute particulars. To create clear images, receiving the noise estimation module's output was the removal module [24]. Without a doubt, CNN efficiently picks up on noise patterns and picture patches. However, the network produced by this learning contains many image patches and training data. The local division and deep conquer network with patch complexity was depicted in reference [25] as a result of the preceding (PCLDCNet). The network was broken up into regional subtasks and instructed locally (based on the patch and conquer block and clear image). The local subtask was coupled with each weighting combination for a noisy patch.

Finally, image patches were organized according to complexity [26], and the modified stacked denoising autoencoders were used to train the k-network [27]. Another issue with a deep learning network is network degradation (The mistake rate increases with layer depth.). Despite ResNet [28] being introduced, this problem still needed to be completely solved. Without the necessity for identity, mapping uses hierarchical residual learning for picture denoising, three sub-networks that make up the network are extraction of features, inference, and fusion. Increasing-dimensional feature maps are represented by patches extracted by a feature extraction sub-network. A huge receptive field is produced by cascading convolutions in the interference sub-network [29]. The cascaded procedure created tolerant errors in noise estimates and discovered noise maps from multi-scale data. The fusion sub-network then fuses the entire noise map to produce an estimation.

In recent times, CNN architectures have found considerable usage in picture denoising. We have suggested an analysis of various CNN image-denoising approaches. To help readers comprehend current developments, several concepts and methodologies were clearly explained. There are several methods for CNN denoising. This study has 144 references in all. According to the study, GAN was the technique for CNN picture denoising that was most frequently utilized. Several extractions and clean picture production techniques utilized the generator and the discriminator. Interestingly, several researchers coupled the GAN and DCNN approaches. Additionally, U-Net and feedforward CNN were employed. Researchers have employed the residual network at various times. The effectiveness and efficiency of the residual network may be a factor in its heavy usage. Researchers adopted the residual network to reduce the number of convolutions in their network. The researcher tried combining noise as an innovative measure (impulse Gaussian noise). Several thorough deep convolutions were needed to decrease mixed noise in photos. Medical photos frequently contain Rician and speckle noise. Pre-trained networks have demonstrated remarkable noise reduction performance in medical images. The Berkeley database was the one that CNN used the most for image denoising [30].

Additionally, CNN methods for image-denoising tasks commonly incorporate residual networks and the attention mechanism. Their widespread acceptance can be attributed to their popularity and image-denoising efficacy. A few issues with CNN image denoising techniques are a need for more issues with unsupervised denoising operations and memory for CNN applications. In conclusion, only a few CNN techniques were applied to medical images. Additional CNN methods that may be used to denoise medical images would be advantageous. Additionally, the writers tried to compile programs and software, but it wasn't available. More RAM allotments being made available for the CNN work will be very

beneficial [30]. The major issues or problem with the existing approaches are that they fails to give appropriate denoised images even in case of applying pre-processing approaches. Also, the error rate is also higher and fails to give promising outcomes.

3. METHODOLOGY

Here, the image denoising is done using auto-encoders composed of two modules: encoder and decoder. The encoding layer maps the provided input data to latent feature representation, and the decoding layer maps the representation towards the original data. Therefore, the number of neurons in the input and output layers is similar. The training label is similar to the input data. With the integration of intermediate layers, the auto-encoding concept is introduced. The encoding concept is performed either in convolutional or fully connected layers. The training for the denoising concept is done with the noisy input data, and the corresponding labels are related to it. When the noisy data is analyzed, latent data representation is needed for the data construction. Therefore, the proposed auto-encoder designs newer data at the output when the corrupted data is provided to the training network. The denoised image is provided to the proposed $L - CNM$ model for evaluating the prediction outcomes.

3.1. Image denoising

Fig 1 depicts the network architecture we've implemented. A module with both an encoder and a decoder is part of it. The encoder uses a convolutional network architecture that is common. The following are the components utilized in the down-sampling process: rectified linear unit (ReLU), two 2×2 max pooling operations with stride 2, two 3×3 convolutional layers (16 and 8 channels of padded convolutions, respectively), and ReLU are all used. Parameters are reduced by using the pooling layers. The input size is chosen for all layers' maximum pooling operations. After the encoder, a bottleneck is added with four channels of padded convolutions and 3×3 convolutional layer. One layer of $3\,x\,3$ padded convolutions with 8 or 16 channels, ReLU and up-sampling operation make up the decoder path. The up-sampling technique is viewed as the pooling technique's opposite.

Three things must be precisely laid out to any convolutional neural network (CNN) can be trained: the optimization technique, the starting weight values, and the loss function. We used a typical loss function called mean square error to get the projected AE (MSE). The employment of an optimization algorithm, such as a gradient-based one, is required to minimize the loss function. Adadelta, a gradient-based optimization technique, was used to train the suggested AE. The Adadelta algorithm modifies the parameters to change the learning rate by considering bigger updates for uncommon parameters and fewer features linked with often occurring features (i.e., low learning rates). When using this optimizer, no predetermined learning rate is required. Using the gradient descent method for optimization, the network's weight initialization in a CNN is critical. Our method uses Gaussian distribution with beginning weights with standard deviations of $\sqrt{2/N}$, where N represents all of a neuron's input nodes. Here, N equals 48 for our network's $3x3$ kernels and 16 channels.

Journal of Theoretical and Applied Information Technology
31st October 2023. Vol.101. No 20

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

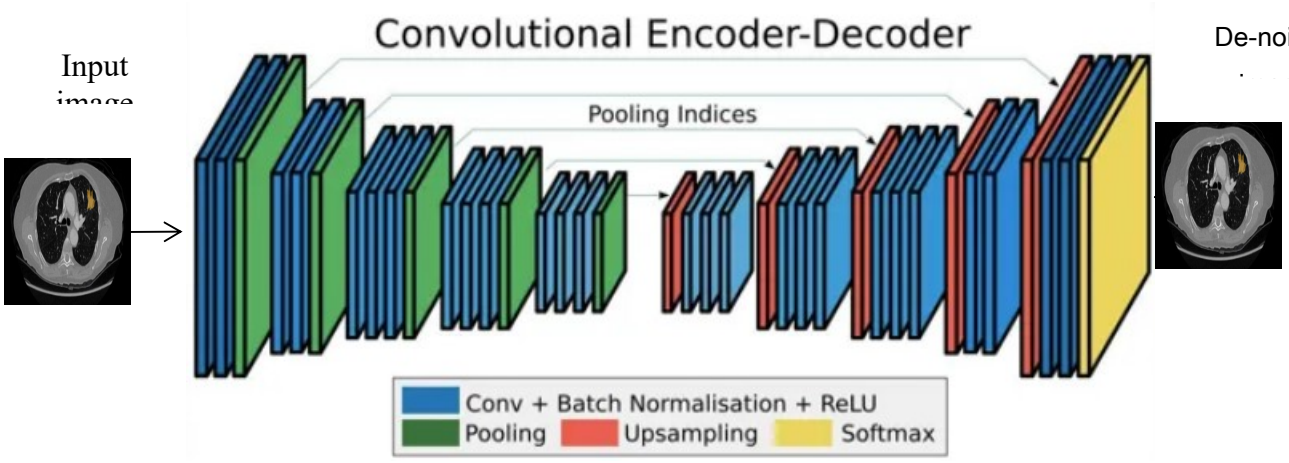

Figure 1: Encoding and decoding network

A large amount of data collection must be prepared if convolutional networks are trained appropriately. Overfitting is prevented by increasing data diversity. Data augmentation is a tactic that gives the network essential invariance and resilience qualities without getting extra data when the training data set size is insufficient. Many augmentation options can be applied to enhance the database's size. CT images can be enhanced using cropping, shifting, horizontal flipping, and rotation (clockwise and counterclockwise). However, the fixed form of the retinal layers and physical limitations force these augmentations to have a restricted range. Real-time and training data augmentation is also possible to avoid excessive memory utilization.

3.2. Classification

We provide a localization technique that estimates landmark locations from a global to local scale. An entirely convolutional neural network makes advantage of patch-based analysis to forecast the locations of several landmarks during global landmark localization. Each landmark's location is fine-tuned using an $L-CNM$ during the subsequent local analysis. The $L-CNM$ used for global and local analysis carries out simultaneous regression and classification for a specific input patch. Regression uses $L-CNM$ to foretell the displacement vectors from any patch's centre to pertinent landmarks. For precise landmark localization, picture patches with more detail are closer to the desired landmark crucial than image patches farther away. Only a few alterations, though, are similarly crucial for precise localization. The median of the iconic places where the anticipated displacement vectors point is used to estimate each landmark's location. The relative value of the landmark localization patches is determined by combining regression and classification. An image patch is classified to determine whether a specific landmark is present. The anticipated displacement vectors are then weighted and averaged using the collected posterior classification probabilities.

A two-stride, $16\ (7\times\ 7\times\ 7)$ kernel convolutional layers is followed by the global $L-CNM$ and has four ResNet blocks. Its foundation is ResNet50. Each ResNetblock has three, four, or six convolutional layer pairs, each having two convolutional layers with kernels of 32, 64, 128, or 256 ($3*3*3$). Between the first and second ResNet blocks in our network are a layer of pooling with a stride and average size of $2*2*2$ voxels. In contrast, every ResNet block in the original ResNet-50 starts with a strung-together convolutional layer. There are two output heads in the network following the four ResNet blocks: one

for classifying the presence of landmarks and another for the regression of displacement vectors. Both output heads have similar designs. A pair of output layers and two 256 −node dense layers are constructed for each head using one-to-one convolutions. A sigmoid function constrains each landmark's scalar output from the classification head to 0 and 1. The regression head produces each landmark's displacement vectors. Although much smaller and concentrated, the idea behind FCNNs for local landmark prediction is similar. Each network comprises average pooling, parallel heads for regression and classification, a ResNet block, and a succeeding ResNet block. The first ResNet block comprises two convolutional layers with a total of 32 $(3 \times 3 \times 3)$ kernels each. The size and stride of a typical pooling voxel are $2 \times 2 \times 2$. Here, 64 $(3 \times 3 \times 3)$ kernels are used in each convolutional layer that makes up the second ResNet block. There are two output heads: a classification output and a regression output are placed after the two ResNet blocks, the same as the global $L - CNM$. All levels use 64 kernels.

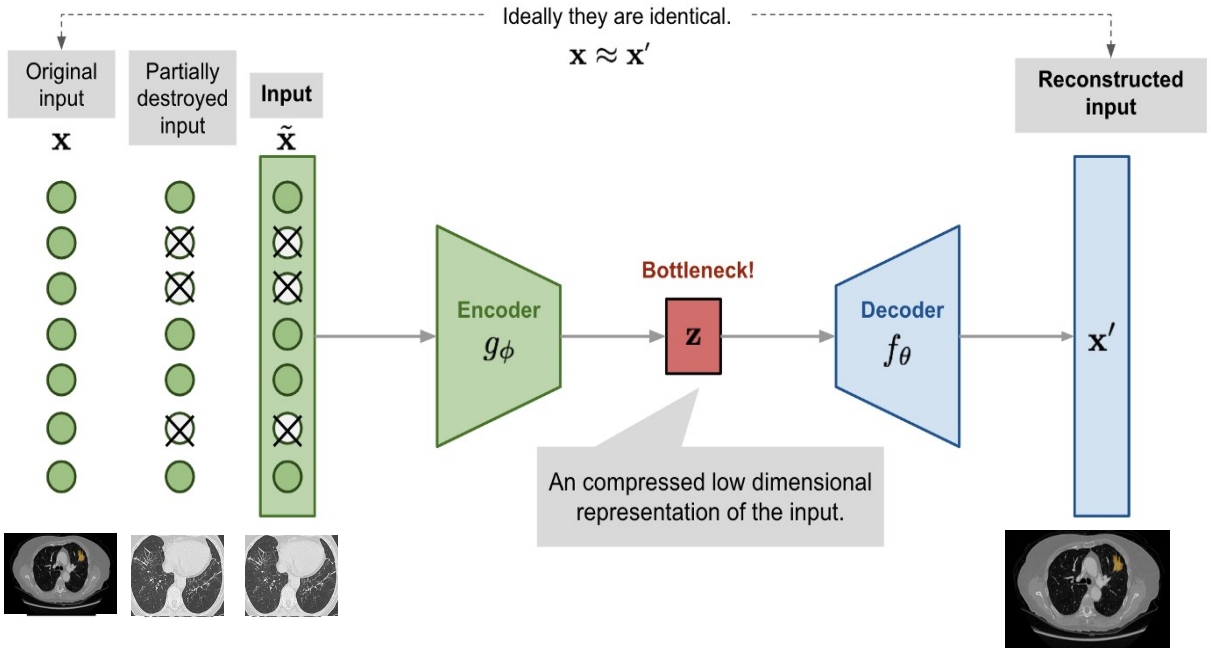

Figure 2: Denoising AE

The $L - CNM$ give each convolutional layer zero padding, and batch normalization is applied after each convolutional layer. Instead of flattened, 3D feature maps' thick layers are implemented as voxel convolutions to enable the application to images of any size (see Fig 2 and Fig 3 for prediction). Unlike the output layers for regression and classification, ReLUs, or rectified linear units, are used by networks to activate. To acquire posterior probabilities between 0 and 1, a sigmoid activation is used for classification and regression and employs a linear activation function. The two components of the adjustments to the loss function made during training were the binary cross-entropy, the mean absolute error between the classification output and reference labels, and the regression output and reference displacements. Displacement vectors with a log transformation are used to calculate the mean absolute error to minimize the influence of input patches placed distant from the landmark on changes to network parameters than input patches placed close to the landmark. Adam was the optimization strategy applied, and 0.001 was the learning rate.

A network can assess input images of various sizes since it is fully convolutional. During global landmark localization, the network generates a variety of displacement vectors and posterior probabilities depending on the input image. The network's outputs are dispersed throughout a grid. Since they are the first convolutional layer with a

stride of two voxels, the total number of pooling layers and stride convolutional layers determine the grid spacing and the average pooling layer of the network, respectively. For the global or local localization step, where n is the sum of the pooling layers and stride convolutional layers, the down-sampling rate is $1/2^n$, and the patch size is 2^n voxels. Patches from an input image are sampled, and a grid with 2^n voxel spacing created especially for a particular network is used.

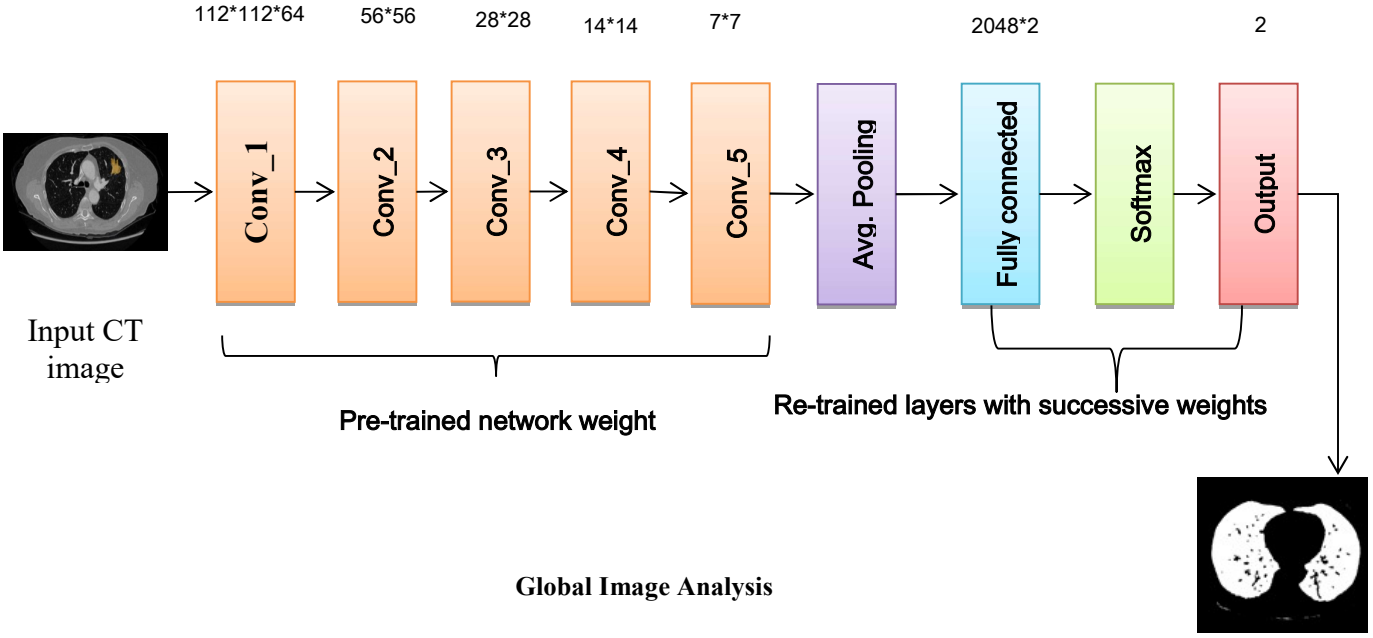

Figure 3: L-CNM architecture

4. NUMERICAL WORKS

This section shows the numerical outcomes of the proposed $L-CNM$, which combines auto-encoding and classification of input CT images. The anticipated $L-CNM$ model employs certain image information as a constraint to enhance the denoising performance. The network intends for lost restoration during denoising using the proposed $L-CNM$ model. The model employs an essential training strategy to train the network model independently for predicting the noise in the preliminary stage and enhancing the accuracy by noise elimination, thus improving the network performance. The proposed $L-CNM$ model works well over certain noise levels and depicts the model's robustness.

4.1. Data and execution details

We used the data set from the Combined Healthy Abdominal Organ Segmentation (CHAOS) challenge, which comprises 51642 abdominal CT slices with a size of $112 * 112$ for the low-dose CT perfusion noise denoising job. Testing involves the utilization of 9000 slices, while 42642 slices are chosen for training. MATLAB 2020a is used to implement our networks. With a 10^{-3} initial learning rate, we optimize the training process using Adam. After 30 epochs, the learning rate drops to 10 4 per cent. There is a batch size of 64. Furthermore, we employ four and eight NVIDIA 2080Ti GPUs to train $L-CNM$. However, only one NVIDIA 2080Ti GPU is needed for testing. To illustrate the potency and robustness of the suggested model, these models are trained using a predetermined blind noise level and noise level.

4.2. Evaluation metrics

Two quantitative metrics in this work, the peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM), are used to evaluate the effectiveness of our method. The following indicates the PSNR.

$$PSNR = 20\log_{10}\frac{255}{RMSE} \quad (1)$$

RMSE is the difference between the denoised images and the noise-free image's root mean square error. A measurement that is more in line with the sensation of sight is the SSIM, and it has the following representation.

$$SSIM\,(x,y) = \frac{(2\mu_x\mu_y)(2\sigma_{xy} +}{(\mu_x^2+\mu_y^2+c_1)(\sigma_x^2+} \quad (2)$$

Where μ_x and μ_y represent an image's mean value, σ_x and σ_y represent its standard deviation, σ_{xy} represents its covariance, and c_1 and c_2 represent two constants. To demonstrate our network's effectiveness, we trained it on both levels of precise and blind noise. To train our model specifically for noise level denoising, we used the noise levels for CT images 15, 25 and 35, respectively. We used random noise levels to train our model between 15 and 35 and 17 and 32 for blind noise denoising on CT images, respectively. We evaluated how well the suggested method performed against several existing denoising techniques like BM3D, D-CNN, DNN, W-NM, and low-rank matrix approximations (LRMA). We once more assessed the performance of the suggested method for low-dose CT perfusion noise denoising applications BM3D, D-CNN, DNN, W-NM, and low-rank matrix approximations (LRMA) methods.

4.3. Qualitative analysis

This work shows various analyses: (a) as the noisy input image with PSNR of 25.67 dB; (b) input low-dose perfusion noise; (c) ground truth of the clean image; and (d) denoised outcomes of various approaches. Fig 4a to 4c and Fig 5a to Fig 5c illustrates an outputs from the methods described above. Visually, the suggested method outperforms other approaches, as is evident. Additionally, the suggested method achieves 1.5 dB greater PSNR, which is higher than D-CNN and illustrates the effectiveness of our technique for denoising CT image noise. From the denoised images, the BM3D and W-NM cannot remove the disturbances adequately. D-CNN frequently loses more tissue detail information, and its texture information is typically not well retained. However, our suggested model can recover features with sharper edges, making them more visually appealing to human eyes.

We employ the techniques mentioned above to eliminate the noises produced during the acquisition of the CT perfusion imaging to test the effectiveness of the procedure suggested for clinical CT images with low noise levels. Since no fundamental ground truth can be used, we utilize visual perception to compare the denoised results of various techniques. We use the BM3D, D-CNN, DNN, W-NM, and low-rank matrix approximations (LRMA) models for blind noise levels for the deep learning methods to directly deal with the noisy images because it is uncertain how noisy the input image is. We established a noise level range for denoising equivalent to the deep neural network-based approaches for the conventional BM3D, D-CNN, DNN, W-NM, and low-rank matrix approximations (LRMA) methods. Fig 4a to Fig 4c displays the outcomes of one sample that was chosen at random. The suggested model produces greater visual results than the other three approaches. Because of the limitations of these models, some noise in the image remains in the denoised results produced by BM3D, D-CNN, DNN, W-NM and LRMA. In terms of overall performance, deep learning-based approaches beat traditional approaches. The proposed model removes more noise and reserves more detailed information, resulting in an acceptable visual effort.

4.4. Quantitative outcomes

The PSNR and SSIM measurements are then used to assess the effectiveness of different models. The experiment results on blind and particular perfusion noise levels are shown in Table 1 to support the usefulness of the suggested network. In this part, we analyze these models at noise levels 17, 22, and 32. Our proposed model outperforms the two established techniques of BM3D, D-CNN, DNN, W-NM and LRMA in terms of PSNR and SSIM for reducing noise. Additionally, the PSNR for the D-CNN model powered by deep neural networks rises from 36.99 to 37.18 by 0.19, from 35.82 to 36.01 by 0.19, and from 34.08 to 34.31 by 0.23, respectively, highlighting the adaptability of the recommended approach once more. The PSNR increases at each of the three noise levels for the blind noise denoising job from 37.19 to 37.47 by 0.28, 35.17 to 35.38 by 0.21, and 33.53 to 33.86 by 0.33, respectively. It provides additional evidence

for the effectiveness of denoising, multi-supervision and progressive training approaches are suggested. As we can see from the results above, our suggested model performs better on blind and specific perfusion noise levels than BM3D, D-CNN, DNN, W-NM and LRMA. The success of the suggested approach is shown by the fact that, despite noise level 32, the SSIM of the proposed model performs only marginally worse than other approaches. The total denoising performance of the anticipated model is still superior to other methods. Fig 5a to 5c shows how the PSNR and SSIM vary. When several models are used, the training epoch increases. After 50 epochs, D-CNN outperforms other approaches demonstrating that the proposed network is simple to train for convergence.

Table 1: PSNR and SSIM computation

Types	Methods	Noise level = 15		Noise level = 25		Noise level = 35	
		PSNR (dB)	SSIM	PSNR (dB)	SSIM	PSNR (dB)	SSIM
Specific model	BM3D [5]	21.8	0.37	16	0.13	15.5	0.14
	D-CNN [30]	21.7	0.36	16.1	0.13	15.6	0.14
	DNN [4]	36.5	0.89	33.6	0.84	31.8	0.80
	W-NM [7]	36.8	0.89	33.7	0.85	31.9	0.80
	LRMA [10]	36.5	0.90	33.8	0.85	31.9	0.81
	L-CNM	**37**	**0.91**	**34.2**	**0.86**	**32**	**0.82**
Blind model	BM3D [5]	35.5	0.85	32.7	0.79	31	0.63
	D-CNN [30]	35.7	0.87	33	0.80	31.5	0.78
	DNN [4]	35.6	0.88	33.08	0.83	31.5	0.76
	W-NM [7]	35.8	0.89	34	0.84	32	0.77
	LRMA [10]	35.5	0.90	35	0.855	32.5	0.78
	L-CNM	**37**	**0.91**	**35**	**0.86**	**33**	**0.83**

We compare the proposed model results with other renderings in this example to demonstrate the value of multi-supervision data for the difficulty similar to the CT noise denoising experiment part, noise denoising with low-dose perfusion. The two evaluations of PSNR and SSIM for predictions and several renderings are shown in Fig 4 and Fig 5.

The proposed model performs better than the other three renderings, proving that combining different noise scales during training might produce more fruitful outcomes. Again, the PSNR and SSIM of render-3 outputs aren't necessarily higher than those of render-1 and render-2, emphasizing the importance of including hierarchical noise information for better prediction and the importance of multi-supervision information for improving denoising performance. We re-evaluate the model's

performance after confirming the effects of the proposed network for refinement, the method of progressive training, and the various enormous channels. The results are shown in Table 1. The proposed model performs worsen than the refinement network.

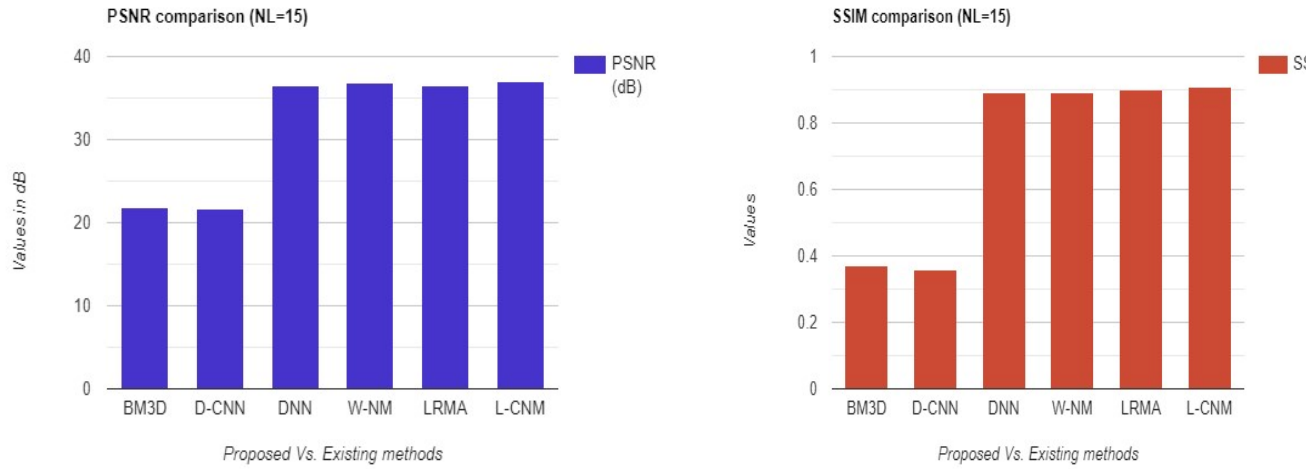

Fig 4a. PSNR and SSIM comparison for NL =15 (specific model)

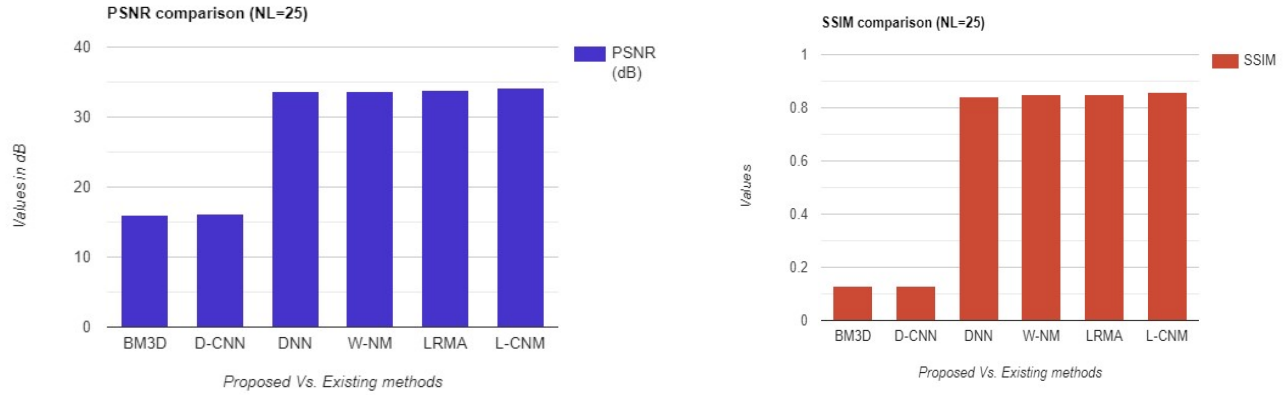

Fig 4b. PSNR and SSIM comparison for NL =25 (specific model)

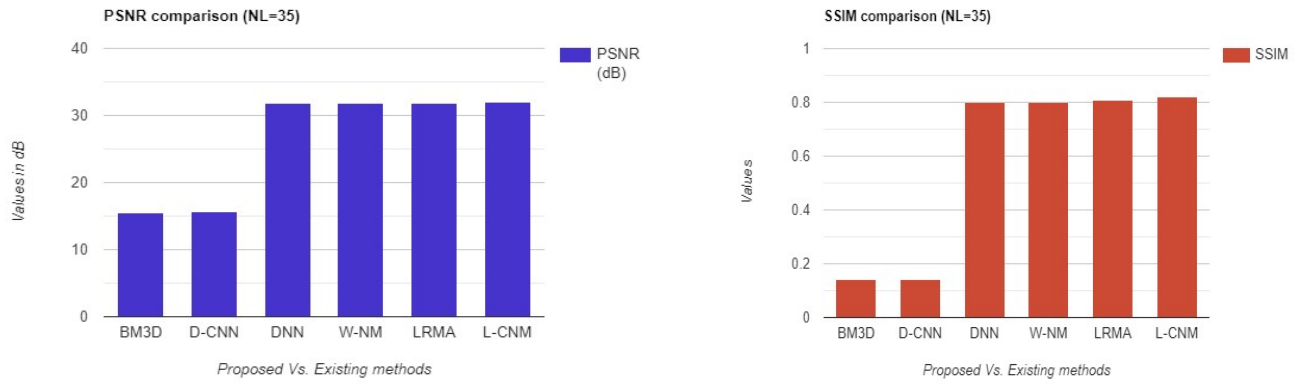

Fig 4c. PSNR and SSIM comparison for NL =25 (specific model)

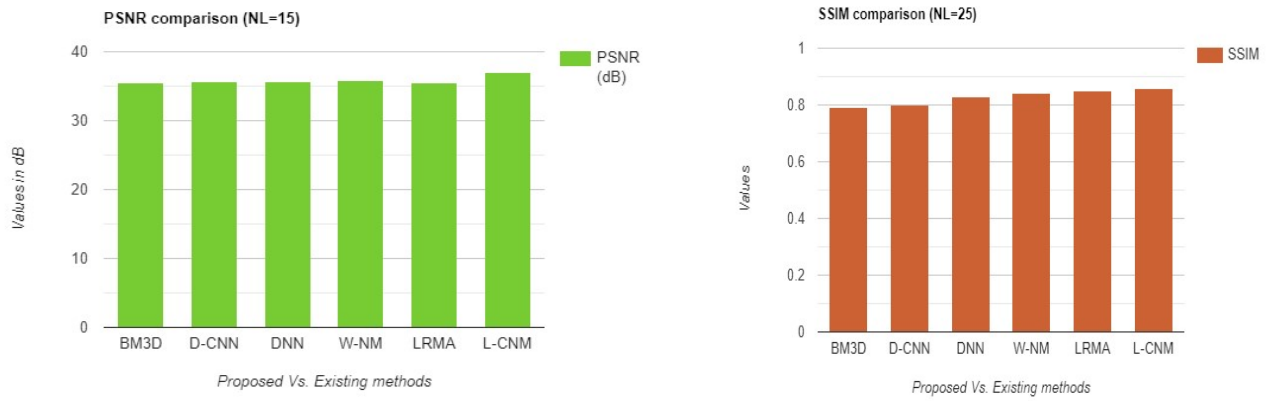

Fig 5a. PSNR and SSIM comparison for NL =15 (Blind model)

Fig 5b. PSNR and SSIM comparison for NL =25 (Blind model)

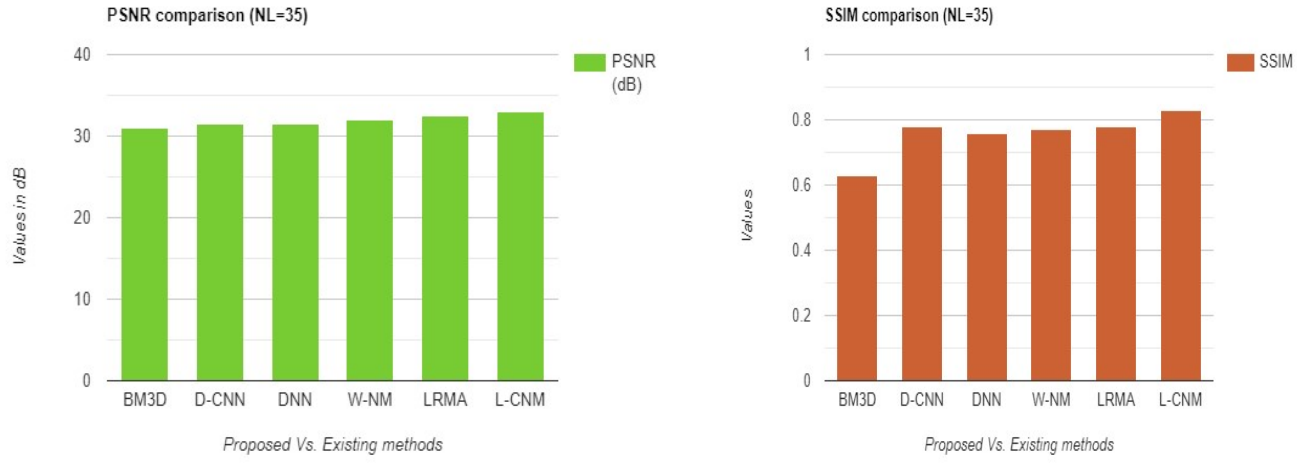

Figure 5c:. PSNR and SSIM comparison for NL =35 (Blind model)

At three different noise levels, the PSNR lowers from 37.18 to 37.14, 36.01 to 35.98, and 34.31 to 34.28, respectively. Without progressive training, the SSIM of the proposed model with channel 94 dropped at the same noise levels from 0.9308 to 0.9305, 0.9147 to 0.9143, and 0.8860 to 0.8789, respectively. It proves that improving network performance through progressive training is effective. Furthermore, the anticipated model with the big channel outperforms existing models with the small channel in terms of performance. The model achieved 36.94, 35.85, and 34.10 at each of the three noise levels, by 0.24, 0.16, and 0.21, respectively, as shown by the comparison in Table 1. It represents a significant improvement in the areas of image denoising. Doing so confirms the significance of using many channels while denoising.

5. CONCLUSION

We created a cascaded multi-supervision convolutional neural network using a progressive training method to remove CT low-dose perfusion noise. Under the supervision of the most recent convolutional layers, to forecast latent noise, the proposed model is used and employed during denoising operations where some missing details are recovered using the proposed L-CNM network. The proposed model is trained separately to forecast the first noise before working together with the refined network to train it using progressive training, which improves network performance. The suggested network may learn hierarchical features obtained from inner layers with varied supervision levels, so resolving the issue with current algorithms that exclusively use data from the anticipation of the noise-free image by the final convolution layer. CT image denoising requirements for some problems with known and unknown noise levels, we evaluated the performance of the suggested technique. According to experimental results, the suggested model performed well when measuring the global structural similarity index (SSIM) and the peak signal-to-noise ratio (PSNR) with unknown and particular noise levels. In the future, this work is further extended by the adoption of optimization and hybrid classification approaches to enhance the global and local prediction outcomes.

REFERENCES

[1] Brenner and E. J. Hall, ‘‘Computed tomography—An increasing source of radiation exposure,’’ New England J. Med., vol. 357, no. 22, pp. 2277–2284, 2007.

[2] Liu, H. Shangguan, Q. Zhang, H. Zhu, H. Shu, and Z. Gui, ‘‘Median prior constrained TV algorithm for sparse view low-dose CT reconstruction,’’ Comput. Biol. Med., vol. 60, pp. 117–131, May 2015.

[3] Cui, Z.-G. Gui, Q. Zhang, H. Shangguan, and A.-H. Wang, "Learning-based artefact removal via image decomposition for low-dose CT image processing," IEEE Trans. Nucl. Sci., vol. 63, no. 3, pp. 1860–1873, Jun. 2016.

[4] Chen et al., ‘‘Low-dose CT with a residual encoder-decoder convolutional neural network,’’ IEEE Trans. Med. Imag., vol. 36, no. 12, pp. 2524–2535, Dec. 2017.

[5] Liu et al., ‘‘3D feature constrained reconstruction for low-dose CT imaging,’’ IEEE Trans. Circuits Syst. Video Technol., vol. 28, no. 5, pp. 1232–1247, May 2018.

[6] Ge et al., ‘‘Structure-based low-rank model with graph nuclear norm regularization for noise removal,’’ IEEE Trans. Image Process., vol. 26, no. 7, pp. 3098–3112, Jul. 2017.

[7] Gu, Q. Xie, D. Meng, W. Zuo, X. Feng, and L. Zhang, "Weighted nuclear norm minimization and its applications to low-level vision," Int. J. Comput. Vis., vol. 121, no. 2, pp. 183–208, Jan. 2017.

[8] Dong, C. C. Loy, K. He, and X. Tang, “Image super-resolution using deep convolutional networks,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 38, no. 2, pp. 295–307, Feb. 2016.

[9] S. Sindhu, M. Saravanan, “An optimised extreme learning machine (OELM) for simultaneous localisation and mapping in autonomous vehicles,” International Journal of System of Systems Engineering, vol. 13, No. 2, pp. 140–159, June. 2023.

[10] Wang, B.-C. Yin, J. Du, C. Liu, X. Tao, and G. Hu, “Fast and robust detection of anatomical landmarks using cascaded 3D convolutional networks guided by linear square regression,” in Chinese Conference on Biometric Recognition. Springer, 2018, pp. 599–608.

[11] Wang, B. C. Ooi, X. Yang, D. Zhang, and Y. Zhuang, ‘‘Effective multimodal retrieval based on stacked auto-encoders,’’ Proc. VLDB Endowment, vol. 7, no. 8, pp. 649–660, Apr. 2014.

[12] Che, S. Purushotham, K. Cho, D. Sontag, and Y. Liu, ‘‘Recurrent neural networks for multivariate time series with missing values,’’ Sci. Rep., vol. 8, no. 1, p. 6085, Dec. 2018.

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

[13] Kim and K. Chung, ‘‘Neural-network based adaptive context prediction model for ambient intelligence,’’ J. Ambient Intell. Humanized Comput., vol. 11, no. 4, pp. 1451–1458, Apr. 2020.

[14] Cadena, A. Dick, and I. D. Reid, ‘‘Multi-modal auto-encoders as joint estimators for robotics scene understanding,’’ in Proc. 12th Robot., Sci. Syst., 2016, pp. 1–9.

[15] Wu, K. Kim, G. E. Fakhri, Q. Li, A cascaded convolutional neural network for x-ray low-dose ct image denoising, arXiv preprint arXiv:1705.04267

[16] Amini, R. Kafieh, and H. Rabbani, “Speckle noise reduction and enhancement for OCT images,” in Retinal Optical Coherence Tomography Image Analysis. Singapore: Springer, 2019, pp. 39–7

[17] Stankiewicz, T. Marciniak, A. D ̨abrowski, M. Stopa, P. Rakowicz, and E. Marciniak, "Improving segmentation of 3D retina layers based on graph theory approach for low-quality OCT images," Metrology Meas. Syst., vol. 23, no. 2, pp. 269–280, Jun. 2016.

[18] Kafieh, H. Rabbani, and G. Unal, “Bandlets on oriented graphs: Application to medical image enhancement,” IEEE Access, vol. 7, pp. 32589–32601, 2019.

[19] Jorjandi, H. Rabbani, Z. Amini, and R. Kafieh, “OCT image denoising based on asymmetric normal Laplace mixture model,” in Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC), Jul. 2019, pp. 2679–2682.

[20] Amini and H. Rabbani, “Optical coherence tomography image denoising using Gaussianization transform,” J. Biomed. Opt., vol. 22, no. 8, 2017, Art. no. 86011.

[21] Shi et al., “DeSpecNet: A CNN-based method for speckle reduction in retinal optical coherence tomography images,” Phys. Med. Biol., vol. 64, no. 17, Sep. 2019, Art. no. 175010.

[22] Sherubha, "Graph-Based Event Measurement for Analyzing Distributed Anomalies in Sensor Networks", Sådhanå(Springer), 45:212, https://doi.org/10.1007/s12046-020-01451-w.

[23] Sherubha, “An Efficient Network Threat Detection and Classification Method using ANP-MVPS Algorithm in Wireless Sensor Networks”, International Journal of Innovative Technology and Exploring Engineering (IJITEE), ISSN: 2278-3075, Volume-8 Issue-11, September 2019.

[24] C. Kavitha, Saravanan M, “Filter-Based Ensemble Feature Selection and Deep Learning Model for Intrusion Detection in Cloud Computing”, *Electronics* 2023, *12*(3), 556, https://doi.org/10.3390/ electronics12030556, January 2023.

[25] Kande, R. Dakhane, A. Dukkipati, and P. K. Yalavarthy, “SiameseGAN: A generative model for denoising of spectral domain optical coherence tomography images,” IEEE Trans. Med. Imag., vol. 40, no. 1, pp. 180–192, Jan. 2021.

[26] Charte, F. Charte, S. García, M. J. del Jesus, and F. Herrera, “A practical tutorial on autoencoders for nonlinear feature fusion: Taxonomy, models, software and guidelines,” Inf. Fusion, vol. 44, pp. 78–96, Nov. 2018.

[27] Sivaswamy, “Shared encoder based denoising of optical coherence tomography images,” in Proc. 11th Indian Conf. Comput. Vis., Graph. Image Process., Dec. 2018, pp. 31–35.

[28] Alansary, O. Oktay, Y. Li, L. Le Folgoc, B. Hou, G. Vaillant, K. Kamnitsas, A. Vlontzos, B. Glocker, B. Kainz et al., “Evaluating reinforcement learning agents for anatomical landmark detection,” Medical image analysis, vol. 53, pp. 156–164, 2019.

[29] Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.

[30] Wolterink, R. W. van Hamersvelt, M. A. Viergever, T. Leiner, and I. Išgum, “Coronary artery centerline extraction in cardiac CT angiography using a CNN-based orientation classifier,” Medical image analysis, vol. 51, pp. 46–60, 2019.