Medical Image Denoising with Recurrent Residual U-Net (R2U-Net) base Auto-Encoder

Shamima Nasrin, Md Zahangir Alom, Ranga Burada, Tarek M. Taha, and Vijayan K. Asari
Department of Electrical and Computer Engineering, University of Dayton, Dayton, OH, USA
Emails: {nasrinm1, alomm1, buradar1, ttaha1, and vasari1}@udayton.edu

***Abstract*—Deep learning (DL) approaches have been applied in different sectors of medical imaging applications, i.e. classification, segmentation and detection tasks and shown superior performance. The DL based generative methods are used for image denoising, enhancement and restoration task. In case of image analysis, image denoising is one of the most crucial pre-processing steps. Recently, there are various DL approaches are applied in image denoising problems and achieved state-of-the-art performance. In this work, we apply recurrent residual U-Net (R2U-Net) based autoencoder model for medical image denoising which is applied for digital pathology, dermoscopy, Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) images denoising tasks. The performance of R2U-Net based auto-encoder model is also evaluated for Transfer domain (TD) between MRI and CT scan images. The experiments have conducted on different publicly available medical image datasets and shows promising denoising results which can be applied in different medical imaging applications.**

Keywords— DL, R2U-Net, Autoencoder, Medical Imaging, and Image denoising.

I. INTRODUCTION

The medical image becoming very popular day by day, because of handiness of patients data. The number of patients are increased and the nature of diseases are also changed over the times. Nowadays, it is a great challenge to provide scientific treatment, for this reason a huge number of medical image processing researchers try to establish a medical images data warehouse, so that doctors can easily make their decision by analysis the previous case history using those warehouse. This research also makes the treatment more accurate, easy, time and cost effective. **[1].** In image processing applications, image denoising is one of the most fundamental process of removing noise from an image. It is a classical technique of image enhancement and restoration tasks. Medical imaging including digital pathology, demos copy, magnetic Resonance Imaging (MRI) and Computed Tomography (CT), ultrasound etc. are susceptible to noise [2]. The most common noise is speckle noise. The medical image with speckle noise reduce the accuracy rate of computer-assisted diagnosis result [3].

There are different image acquisition device or techniques are used for image capturing in pathology, dermoscopy, MRI and CT scan. As a result, the quality of images are also varied due to different environments (lighting, intensity, different artifacts etc.). The performance of Artificial Intelligent (AI) based medical image analysis system or any other classical methods significantly vary with respect to the quality of images. Image denoising is required for proper image analysis and it also ensure the better performance for both human and machine (AI system).

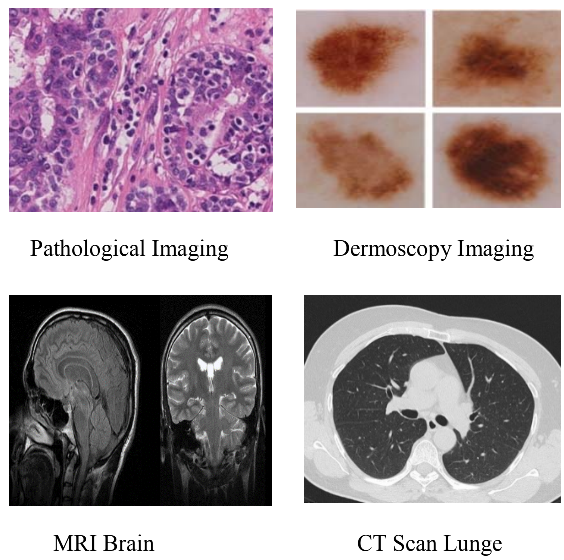


Fig. 1. Example images for different modalities of medical imaging.

In image denosing, there are various methods exist, ranging from model based on partial differential equation (PDEs) [4-6], domain transformations such as wavelets [7], dual tree complex wavelet and shrinkage with Wiener filter technique [8], Discrete cosine transformation (DCT) [9], Bayesian least squares (BLS) - Gaussian scale mixture (GSM) [10], Block-maching and 3D filtering (BM3D) [11] etc., and a family of models exploiting sparse coding techniques [12-14]. All methods share a common goal, expressed as $C = A + B$ where C is the noisy images produced as a sum of origianl images A and noise B. Most methods try to approxiamate A using C as close as possible. In most cases, B is assumed to generated from a well defined process.

In this paper, we proposed a network referred to as Recurrent Residual U-Net (R2U-Net) based auto-encoder model.The standard R2U-Net performs concatenation operations between encoding and decoding units whereas R2U-Net based autoencoder model do not use concatenation operations that helps to decrease computational cost. In every single steps, our input follwed recurrent residual convolutional opeation.

The main contributions of this work are summarised as follows:

- Examine best R2U-Net architecture for Auto-Encoder model in medical image enhancement tasks.

- Denoise digital pathology, dermoscopy, magnetic Resonance Imaging (MRI) and Computed Tomography (CT) images using R2U-Net based Autoencoder.
- Evaluate the performance for Transfer Learning (TL) between MRI and CT image enhancement problem.

The remaining of this paper is oganised as folllows, section 2 related work, section 3 defines autoencoder and its variants , section 4 provides the proposed method, section 5 shows the qualative experimental results of this paper, section 6 conclude with the conclusion and future work.

II. RELATED WORKS

Nowadays, the deep learning based models have shown promising performance in different modalities of medical imaging and computer vision tasks in [15,16].In the last few decades, there are different alogorithms have been proposed with varying their denosing performances. Some of the classical methods have already mention in Section 1.
There are several deep leanring methods have been proposed. The Deep convolution neural network (DCNN) is learning-based methods that, the latent clear image can be achieved by separating the noise image from the contaminated image. The gradient clipping approach has applied during training for preventing the gradient explosions and enable to enable the network to converge faster. The experiments have done on different types of medical images where they achieved state-of-the-art perofrmance and also stated that their methods has the ability of suppressing different noise with different noise factors by means of one single denoising model [17] .The Enhanced Convolutional Neural Denoising Network (ECNDNet) is used for image denoising where the residual learning and batch normalization techniques to address the problem of training difficulties and accelerate the convergence of the network. In addition, the dialated convolutions are applied for larging the context information and reducing the compuatoinal cost [18].

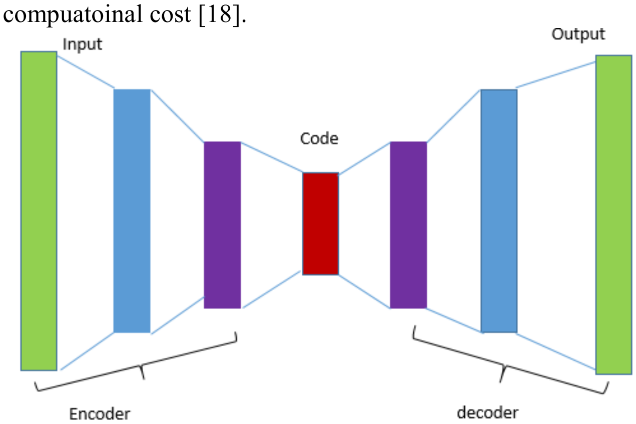


Fig. 2. Basic autoencoders

.An image denosing method has proposed with sparse coding and denoising auto-encoder (DA) with pretrained weigtths. In addition, the blind inspainting approach has introduced which help to remove the complex patterns such as superimposed text from an image [19] .An convolutoinal auto-encoder based efficient medical image denoising method has proposed that combines to boost sample size for increasing medical image deonisng performance [20].

Autoencoder and its variants : An autoencoder is an unsupervised learning technique for neural networks that learns efficient data representations(encoding) by training the network to ignore signal “noise”. The basic autoencoder network (Fig. 2.) has three layers: the input, a hidden layer for encoding, and the output decoding layer. Now suppose we have only a set of unlabeled training samples $\{x^{(1)}, x^{(2)}, x^{(3)}, \dots \dots\}$ Where $x^{(i)} \in R^n$. Using backpropagation, the unsupervised algorithm continuously trains itself by setting the target output values to equal the inputs,i. e., it uses $y^{(i)} = x^{(i)}$. This forces the smaller hidden encoding layer to use dimensional reduction to eliminate noise and reconstruct the inputs [21] .

Denoising Auto Encoder (AE): Denoising autoencoders are an extension of the basic autoencoder, and represent a stochastic version of it. Denoising autoencoders attempts to address identify function risk by randomly corrupting input (i.e. introducing noise) that the autoencoder must then reconstruct or denoise explained in [20] . However, in this implementation, we have used Recurrent residual U-Net (R2U-Net) base auto-encoder model for image denoising. The R2U-Net model has proposed very recently for segemntation tasks which has applied and shown promising resutls for image segmeantion problems including retinal blood vessel segmeantion, skin cancer segmeantion, and lung segmenation in [22].

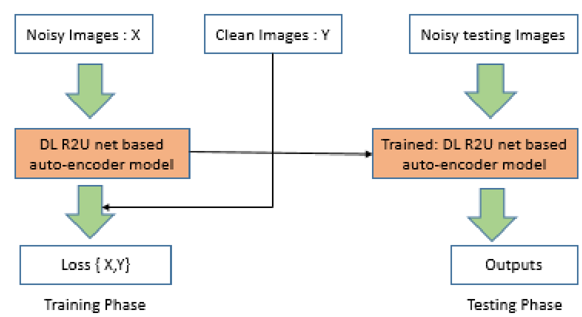


Fig. 3. Training and testing approach of proposed DL based denoising method.

III. METHODS

An end-to-end image denosing method with R2U-Net autoencoder model is shown in Fig. 3 . The entire process consists of two different phases including training and testing phase. In the training phase, we trained our model with noisy images (X) follwed by clean images (Y) with adding gaussian noise. The equation of generating noisy images from clean images are as follows:

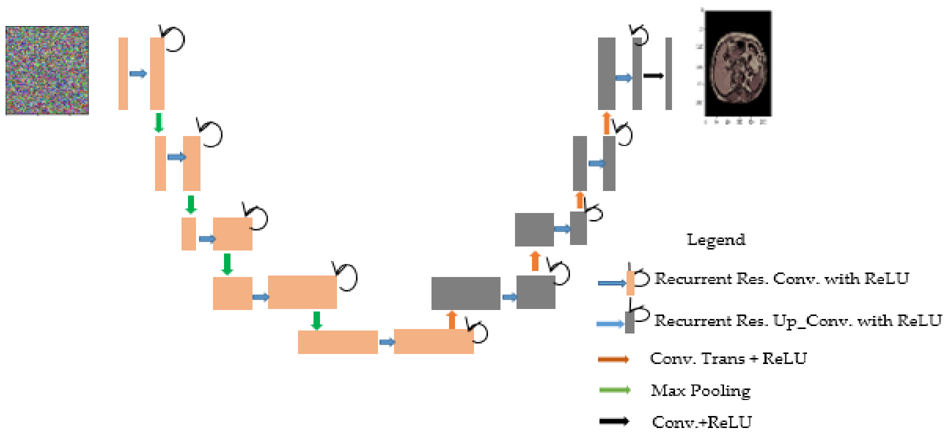


Fig. 4. Proposed R2U-Net based auto-encoder model.

$$noisy_{x_{training}} = training\_image + noise_{factor} * np.random.randn * (x_{training\_image.shape})$$

The noise factors of 0.1, 0.3, and 0.5 are used in this impelementation. In the training phase, at every single step we calculate the loss function{X,Y} and try to propagate the loss in through the network during backproparation with respect to the epochs. After training completing the phase, we tested our model with noisy images with the equations stated above. The trained model then produces very similar ouputs close to the input imges. Here , we have used 90 percent of images for taining phase and rest of the 10 percent of images are used for testing the model. These samples are selected randomly for training and testing phases.

N*etwork architecture*

The convolutional architecture between encoding and decoding units remain same as R2U-Net.

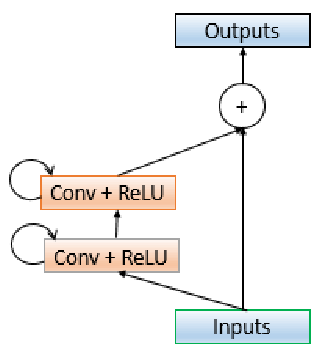


Fig. 5. Basic Recurrent Residual Convolutional Units (RRCU).

Inspired by the R2U-net based autoencoder model, we have proposed here modified R2U-net based autoencoder model for image enhancement task. The proposed R2U-net based autoencoder is shown in Fig.4. Accoding to the Fig.4, our proposed model consist with two parts, namely encoding and decoding. Here orange part function is encoding and yellow part operation is decoding. In standard R2U-Net has concatennation operation which is interlinked between encoding and decoding unit. We have removed interconnection and concatennation operation that reduce our operational cost. The basic convolution unit of this model is shown in fig. 5 .The architecture is U-Net with forward recurrent convolutional layers with residual connectivity is called R2U-Net.

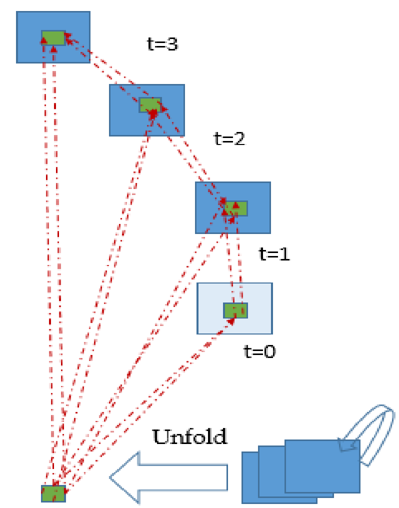


Fig.6. Unfold recurrent convolutional operations.

Fig. 6 represent the unfold RCL layers with different times steps. Here, t=3 (0~3), the recurrent convolutional operation that include one single convolution layer follwed by three sub-sequential recurrent convolutional layers.

Network configuration: The network configuration of our prosed model is as follows : 3→32→64→128→256→512→ 256→128→64→32→3, kernel size is 3x3 with recurrent residual convolution layer, activation funtion is ReLU and in the last decoding step the activation function was sigmoide. In whole model, we use adam optimizer. In every single step we calculation loss fuction and propagate the loss and try to generate the original images. Here we use binary cross entropy loss and the learning rate was 0.0001. In every epochs our batch size was 8 and for all analysis the total number of epochs was 150.

IV. EXPERIMENTAL SETUP

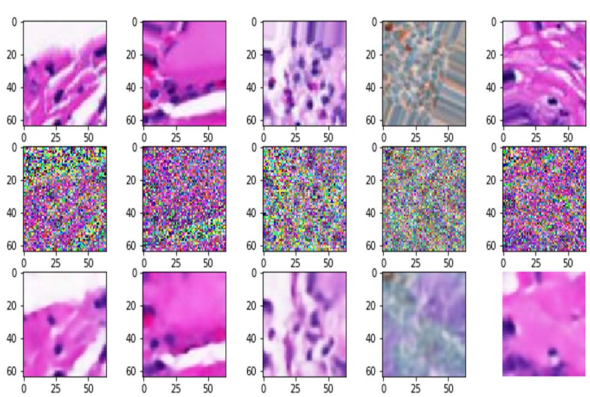
To demonstrate the performance of the R2U-net based auto encoder model, we have tested on four different medical imaging modalities. These include pathology, dermoscopy, MRI and CT images. For these experimental implementations, the Keras and Tensor Flow frameworks are used on a single GPU computer with 56G of RAM and an NVIDIA GEFORCE GTX-980 Ti.

A. Database Summary

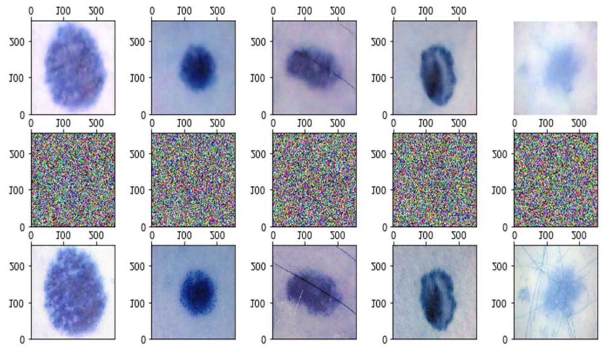
The experiments have been conducted on four different modalities of medical image denoising problem including: pathological image, dermoscopy images, Magnetic Resonance Imaging (MRI) images, and Computed Tomography (CT) image. First, the pathological images are taken on lymphoma classification datasets in [23]. Secondly, the dermoscopy samples are taken from ISIC 2016 dataset in [24]. Third, the MRI image are taken from Kaggle competition in [25]. Fourth, CT images are also taken from Kaggle competition in [26]. In all cases, 256×256 pixel images are used.

V. RESULTS AND DISCUSSION

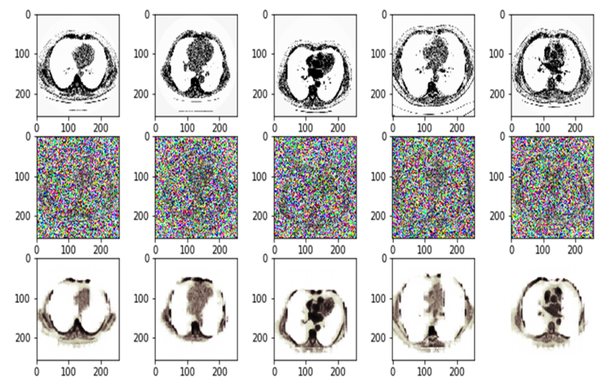
We took clean images as an input, then add some Gaussian noise with input images and use the deep learning model to generate the output as close as possible to the input image. Here, we have used different noise factors in this implementation. For pathological images, we have used 0.1, the noise factor 0.3 is used for dermoscopy and CT scan images, and for MRI images, 0.5 noise factor is used. We trained two models including standard R2U-Net based model and R2U-Net based auto-encoder model with publicly available datasets. Firstly, we trained and tested standard R2U-Net based auto encoder model with the pathology images, the output accuracy was 79.1%.



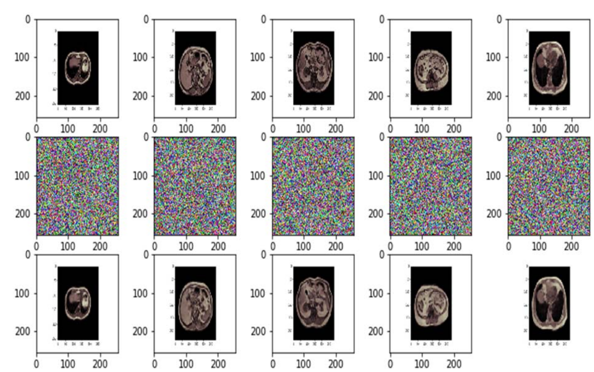
(a) Pathological image (noise factor 0.1)



(b) Dermoscopy image (noise factor 0.3)



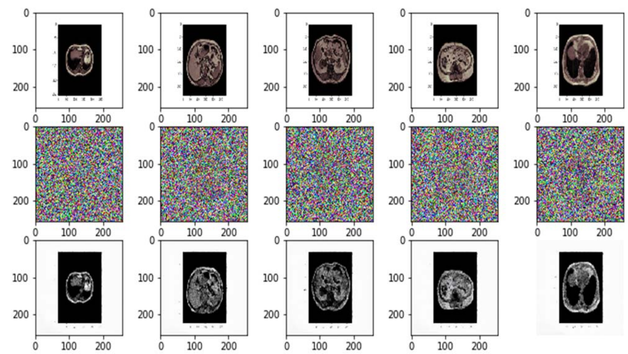
(c) CT scan images (noise factor 0.3)



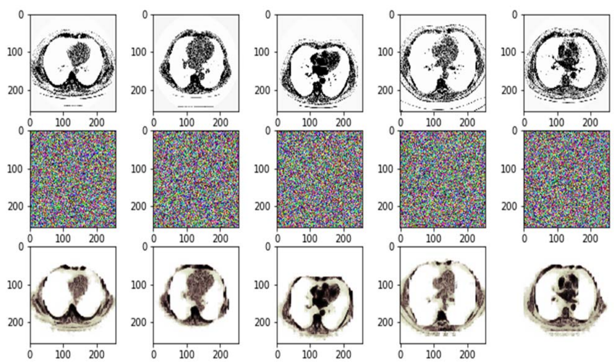
(d) MRI images (noise factor 0.5)

Fig. 7. Qualitative result for R2U based auto encoder model experiment: the first row shows the inputs, second rows show image with noises, and third columns show the experimental results.

Secondly, we trained and tested the same dataset by our proposed model without concatenation between encoding and decoding units and we got little better result than previous, it was 79.9%. The rest of the datasets, we just trained and tested with our proposed model. The qualitative results for R2U-Net auto-encoder model are shown in the Fig.7. The Fig 7 (a), 7(b), 7(c) and 7(d) are shown the outputs for the pathology, dermoscopy image, CT images, and MRI images respectively.



(a) MRI images noise factor 0.3.



(b) CT scan images noise factor 0.3

Fig.8. Qualitative result of transfer learning between MRI and CT images.

Another important aspect of our proposed model is transfers learning between two different domains of medical image denoising including MRI and CT images. We trained our model with MRI images and tested with CT scan images, vise-versa trained with CT scan images and tested with MRI images. Our model shown worthy performance. The qualitative result of transfer leaning between CT images and MRI images are shown in the Fig. 8 (a) and (b). In case of transfer learning, both training and testing phases, the noise factor was 0.3.

VI. Conclusion and Future Work

In conclusion, it can be stated that, we achieved promising image enhancement results for medical image denoising, i.e. pathology, dermoscopy, MRI and CT scan images. The performance are evaluated for different noise factors and have also investigated the performance of transfer learning between MRI and CT image denoising problems. In future, we would like to implement and examine the performance of DL based enhancement model for pathological image de-blurring.

References

[1] Kaur, Prabhpreet, Gurvinder Singh, and Parminder Kaur. "A Review of Denoising Medical Images Using Machine Learning Approaches." *Current medical imaging reviews* 14, no. 5 (2018): 675-685.

[2] Sanches, Joo M., Jacinto C. Nascimento, and Jorge S. Marques. ”Medical image noise reduction using the SylvesterLyapunov equation.” IEEE transactions on image processing 17.9 (2008): 1522-1539.

[3] Maini, Raman, and Himanshu Aggarwal. "Peformance evaluation of various speckle noise reduction filters on medical images." *International Journal of Recent Trends in Engineering* 2, no. 4 (2009): 22.

[4] Perona, Pietro, and Jitendra Malik. ”Scale-space and edge detection using anisotropic diffusion.” IEEE Transactions on pattern analysis and machine intelligence 12.7 (1990): 629-639.

[5] Rudin, Leonid I., and Stanley Osher. ”Total variation based image restoration with free local constraints.” Image Processing, 1994. Proceedings. ICIP-94., IEEE International Conference. Vol. 1. IEEE, 1994.

[6] Subakan, Ozlem, et al. ”Feature preserving image smoothing using a continuousmixtureoftensors.”2007IEEE11thInternationalConference on Computer Vision. IEEE, 2007.

[7] Coifman, Ronald R., and David L. Donoho. Translation-invariant denoising. Springer New York, 1995.

[8] Naimi, Hilal, Amel Baha Houda Adamou-Mitiche, and Lahcène Mitiche. "Medical image denoising using dual tree complex thresholding wavelet transform and Wiener filter." *Journal of King Saud University-Computer and Information Sciences* 27, no. 1 (2015): 40-45.

[9] Yaroslavsky, Leonid P., Karen O. Egiazarian, and Jaakko T. Astola. ”Transform domain image restoration methods: review, comparison, and interpretation.” Photonics West 2001-Electronic Imaging. International Society for Optics and Photonics, 2001.

[10] Portilla, Javier, et al. ”Image denoising using scale mixtures of Gaussians in the wavelet domain.” IEEE Transactions on Image processing 12.11 (2003): 1338-1351.

[11] Dabov, Kostadin, et al. ”Image denoising by sparse 3-D transformdomain collaborative filtering.” IEEE Transactions on image processing 16.8 (2007): 2080-2095.

[12] Elad, Michael, and Michal Aharon. ”Image denoising via sparse and redundant representations over learned dictionaries.” IEEE Transactions on Image processing 15.12 (2006): 3736-3745.

[13] Olshausen, Bruno A., and David J. Field. ”Sparse coding with an overcomplete basis set: A strategy employed by V1?.” Vision research 37.23 (1997): 3311-3325.

[14] Mairal, Julien, et al. ”Online dictionary learning for sparse coding.” Proceedings of the 26th annual international conference on machine learning. ACM, 2009.

[15] Alom, Md Zahangir, Tarek M. Taha, Chris Yakopcic, Stefan Westberg, Paheding Sidike, Mst Shamima Nasrin, Mahmudul Hasan, Brian C. Van Essen, Abdul AS Awwal, and Vijayan K. Asari. "A state-of-the-art survey on deep learning theory and architectures." *Electronics* 8, no. 3 (2019): 292.

[16] Litjens, Geert, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen Awm Van Der Laak, Bram Van Ginneken, and Clara I. Sánchez. "A survey on deep learning in medical image analysis." *Medical image analysis* 42 (2017): 60-88.

[17] Zhang, Fu, Nian Cai, Jixiu Wu, Guandong Cen, Han Wang, and Xindu Chen. "Image denoising method based on a deep convolution neural network." *IET Image Processing* 12, no. 4 (2017): 485-493.

[18] Chunwei Tian, Yong Xu “ Enhanced CNN for image denoising” CAAI Trans. Intell. Technol., pp. 1–7, 2019

[19] Xie, Junyuan, Linli Xu, and Enhong Chen. "Image denoising and inpainting with deep neural networks." In *Advances in neural information processing systems*, pp. 341-349. 2012.

[20] Gondara, Lovedeep. "Medical image denoising using convolutional denoising autoencoders." In *2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW)*, pp. 241-246. IEEE, 2016.

[21] Hinton, Geoffrey E., and Ruslan R. Salakhutdinov. "Reducing the dimensionality of data with neural networks." *science* 313, no. 5786 (2006): 504-507.

[22] Alom, Md Zahangir, Chris Yakopcic, Mahmudul Hasan, Tarek M. Taha, and Vijayan K. Asari. "Recurrent residual U-Net for medical image segmentation." *Journal of Medical Imaging* 6, no. 1 (2019): 014006.

[23] Janowczyk, Andrew, and Anant Madabhushi. "Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases." *Journal of pathology informatics* 7 (2016).

[24] Gutman, David, Noel CF Codella, Emre Celebi, Brian Helba, Michael Marchetti, Nabin Mishra, and Allan Halpern. "Skin lesion analysis toward melanoma detection: A challenge at the international symposium on biomedical imaging (ISBI) 2016, hosted by the international skin imaging collaboration (ISIC)." *arXiv preprint arXiv:1605.01397* (2016).

[25] https://www.kaggle.com/navoneel/brain-mri-images-for-brain-tumor-detection

[26] https://www.kaggle.com/dineshsellathamby/lung-segmentation