

Convolutional Neural Network for Automated Analyzing of Medical Images



A.Cibi, R.Jemila Rose

Abstract: Convolutional Neural Network (CNN) is one of the deep learning algorithms. It is useful for finding patterns in images. Intelligent software automates understanding images and speech. Extracting distinct features, by their own induces intelligent to software for identifying objects, recognizing faces and diagnosing diseases from medical images. With the help of CNN, software on their own acquires the knowledge of patterns from raw data. These developments play a prominent role in medical imaging. Classification, Segmentation and diagnosing are the area where CNN marked its importance. About CNN there has been a large array of improvements achieved in the last few years. We provide a short overview of the role of CNN in medical image analysis. A shallow CNN model is proposed as an automatic diagnosing system. This work specifically concentrates on three key elements: (1) building blocks of convolutional neural networks (2) introduction of various CNN architecture; (3) Challenges in implementing CNN for analyzing medical images.

Keywords: Artificial Intelligence, Convolution Neural Network, Computer Vision, Medical Imaging

I. INTRODUCTION

Digital imaging imparts huge progress in medical imaging modality. Disease diagnosing is benefiting because of the advancement in digital imaging. The hospitals having diagnostic and investigative imaging facilities are producing large amount of imaging data thus causing a huge increase in production of medical image collections [8]. This large collection of medical images needs to be analyzed for diagnosing diseases properly. Development of Computer Aided Diagnosing (CAD) system becomes prominent in the medical field. Radiologists are very good in understanding medical images and write clinical reports. But the advancement of medical images and storage capacity of medical archive systems require clinicians in browsing large datasets. We work to device intelligent software to automate the diagnosing process. We have to make computers to see images like humans see and understand. Looking and understanding images is an easy task for human beings. But it is a tedious task for making computer to look into images and understand.

The difficulty in this is describing formally to computers to recognize images. Artificial intelligence helps computer to learn and understand real world by experience.

Revised Manuscript Received on May 30, 2020.

* Correspondence Author

Ms.A.Cibi*, Assistant Professor, Department of Computer Science and Engineering, Rajalakshmi Engineering College, Chennai.

Dr.R. Jemila Rose, Associate Professor, St.Xavier's Catholic College of Engineering, Nagercoil.

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](http://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

The experiences are defined in terms of concepts with related factors. The relationship between concepts can be showcased by means of hierarchy. Computers can learn complicated concepts from simple concepts build from hierarchy of concepts.

Medical images contain many important features. Extracting these features helps classify the medical images for identifying diseases. Efficient image classification techniques improve diagnosing system. Automated diagnosing system must make use of image features such as color, texture, edges, contrast and intensity to classify images. In recent years, computer vision and machine learning have been working on images to develop efficient techniques in image recognition, image classification, and pattern recognition and object detection. The complexity and non-linearity of image systems makes designing a classification system more complex.

Machine Learning (ML) algorithms have been widely used for classification of images. CNN is a Deep learning (DL) algorithm based on Artificial Neural Network (ANN) predominantly useful for image classification. The literature shows CNN outperforms than the conventional approach of image classification. The main objective of this paper was to study CNN to automate analyzing of medical images. The next sections are organized as follows. Section II gives existing system. Section III describes basic building blocks of CNN. Section IV details milestone architectures of CNN. Section V brief about challenges in designing CNN for analyzing medical images and Section VI concludes the topic.

II. LITERATURE SURVEY

The literature is reviewed to examine available methods that could be used to develop new systems.

Quing et al., [1] presented a customized convolution neural network for classifying lung images. This architecture is a generalized one to classify medical images. It contains one convolution layer, one max-pooling layer, and three fully connected neural layers.

Sameer Khan et al., [2] proposed a deep learning architecture for classifying medical anatomy images. Also, they have compared the experimental result with three milestone architectures LeeNet, AlexNet, and GoogLeNet. It outperforms this three baseline architecture. It is a modified CNN architecture of these three baseline architecture. It shows acceptable performance in classifying anatomy from medical images.

Heba Mohsen et al., [3] used a combination of Discrete Wavelet Transform (DWT) and Deep Neural Network(DNN) for classifying MRI images of the brain into four classes.

This method resembles the CNN architecture. Also, this method produces great accurate results with low hardware specifications.

S.Pereira et al., [4] proposed the automatic segmentation of brain tumor in MRI images based on CNN. They used a kernel of size 3×3 . It has 11 layers. In that 6 layers are convolution layer, 2 layers are max-pooling layer and 3 fully connected layers. Also, this method is evaluated using the BRATS 2013 and 2015 databases. It won first prize in 2013 and second prize in 2015.

Alex et al., [5] presented a deep convolutional neural network to classify 1.2 million high-resolution images into 1000 different classes. The architecture contains 8 learned layers, five convolutional layers, and three fully connected layers. The error rate is very low compared to other conventional methodology. In this method in order to make the training faster non-saturating neurons and a very efficient GPU were used.

Eli Gibson et al., [6] proposed NiftyNet an open source deep learning platform. Using this platform much applications of image processing can be done. The range of applications includes segmentation, regression, image generation, and representation learning applications.

Jianpeng et al., [7] presented a Synergic Deep Learning model using multiple deep convolutional neural networks (DCNNs). Dual DCNN is used to learn simultaneously and correct the errors. So if one DCNN made incorrect classification the correct classification of other DCNN helps to update the result.

Adnan et al., [8] proposed a deep learning framework for Content Based Medical Image Retrieval (CBMIR) systems. It can classify multimodal images of human organs. They used to classify 24 different types of human organs. The system produces 99% accuracy.

Tao et al., [9] proposed a deep learning framework for cervical dysplasia diagnosis. They used CNN for extracting features from multimodal information acquired. It automatically finds cervical dysplasia using image and non-image modalities through back propagation.

Thijis et al., [10] compared state-of-the-art mammography CAD system and pre trained CNN architecture. From the study CNN outperformed the conventional CAD system.

Suresh et al., [11] proposed a feature extraction method based on CNN. This automatic feature extraction system is a modification of Le-Cun architecture used for classifying handwritten images. The experimental results show that the proposed model produces 99% accuracy.

III. BUILDING BLOCKS OF CNN

A. Convolution Layer

The mathematical operation convolution is applied in at least one of the layers of the CNN. Convolution is a linear operation on two functions of a real valued argument. The typical convolution operation is denoted by $X=I*K$, where I is the input image and K is the kernel. The output X is the feature map of the input image. The kernel is the parameters of the learning algorithms. Slide the kernel on the input image to extract different features. Convolution is the first operation in all CNN algorithms. It follows pooling layer. Convolution and pooling act as giant filters. If we give an image of size $224 \times 224 \times 3$ pixels as input to a fully

connected layer having 100,000 perceptron then it requires $224 \times 224 \times 3 \times 100,000 = 15,052,800,000$ parameters. It is highly complicated to process. Convolution and pooling help to reduce some features in the image which are may not require to train.

B. Pooling Layer

The feature maps obtained from the convolution layer is feed to the pooling layer. It is a simple local operation. The pooling function fills the output from the convolution layer with a summary of values calculated using neighboring pixels. Two prominent pooling functions are max pooling and average pooling that summarizes the average presence of a feature and the most activated presence of a feature respectively. Small movements in images give us different values in convolution layer. This may turn into different feature maps. Apply down sampling to overcome this problem. This down sampling is done by the pooling layer

C. Activation Function

Complex patterns of images can be learned by activation functions. Appropriate activation function for the problem in our hand improves the learning rate of the CNN. The activation functions can be broadly classified into two categories linear activation functions and Non-Linear activation functions. Activation functions are used in the hidden layer and output layer. Which neuron should be activated is decided by the activation function. This is done by calculating weighted sum and adding bias with it. List of the activation functions are Linear, Sigmoid, Tanh, RELU and Softmax. RELU is mostly used in CNN. If anyone is not particular about which activation function to be used then uses RELU.

D. Dropout Regularization

Preventing neural network from over fitting is known as dropout. Deep learning neural network requires large number of dataset to train the network. If the training dataset is small it over fits quickly. So randomly removing neurons helps to tune the network. Dropout Regularization optimizes the network.

E. Batch Normalization

CNN helps computers to learn interesting features by itself. This learning by itself is done by batch normalization. These layers are typically placed after activation layers, producing normalized activation maps by subtracting the mean and dividing by the standard deviation for each training batch [15]. Normalize the input output mapping based on new input.

IV. THE ARCHITECTURE OF VARIOUS CNN

A. LeeNet

Yan LeCun, Leon Bottou, Yosuha Bengio and Patrick Haffner proposed LeeNet[12] to recognize handwritten digits and machine printed characters in 1990's. The LeeNet architecture contains five layers. The first layer is a convolution layer with 6 filters of size 5×5 . The input size is 32×32 grayscale image.

After passing through the first layer the size of the image changes from $32 \times 32 \times 1$ to $28 \times 28 \times 6$.

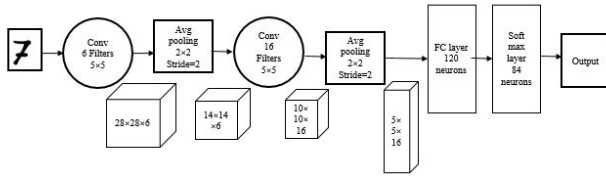


Fig. 1. Architecture of LeeNet

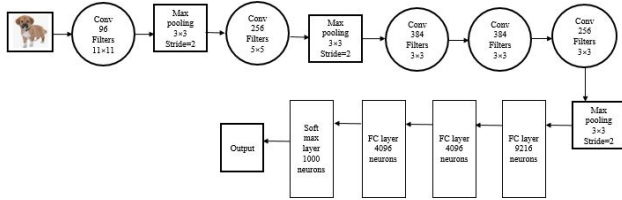


Fig. 2. Architecture of AlexNet

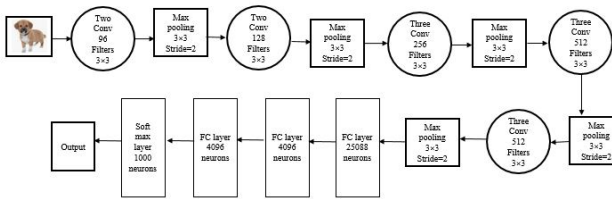


Fig. 3. Architecture of VGG

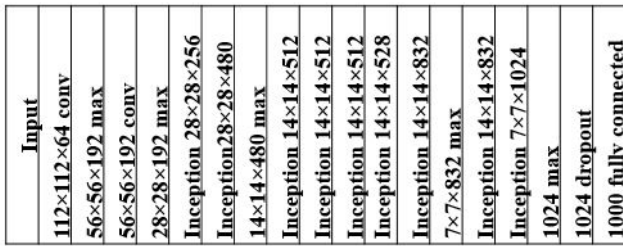


Fig. 4. Architecture of GoogleNet

Second layer is an average pooling layer with filter size 2×2 and a stride of two. Now the image size gets reduced to $14 \times 14 \times 6$. The third layer is again a convolution layer with 16 filters of size 5×5 . The fourth layer is again an average pooling layer with a filter size of 2×2 . The fifth layer is a fully connected convolutional layer with 120 filters of size 1×1 . Finally, a softmax classifier classifies the images. The architecture is shown on Fig. 1.

B. AlexNet

Alex Krizhevsky, Geoffrey Hinton and Ilya Sutskever designed [5] a neural network architecture called AlexNet. It won the Image Classification Challenge (ILSVRC 2012). It is used to classify RGB images into 1000 classes. It can accept a $227 \times 227 \times 3$ size RGB image. First layer is a convolutional layer with 96 filters of size 11×11 and stride of 4. Then the AlexNet applies maximum pooling with a filter of size 3×3 and s stride of 2. Then again the second layer is a convolutional layer with 256 filters having size 5×5 . Then again maximum pooling layers with filter size 3×3 and a stride of 2. The next three layers are convolutional layers with filter size 3×3 . The three convolutional layers are followed by a maximum pooling layer with filter size 3×3 and s stride 2. The sixth layer is flattened through a fully connected layer with 9216 feature maps each of size 1×1 . Then there are two fully connected layers with 4096 units. The last layer is a softmax layer

which classifies into 1000 classes. The architecture is illustrated in Fig. 2.

C. VGG16

Visual Geometry Group (VGG) proposed VGG16 and VGG19 by the team Karen Simonyan and Andrew Zisserman [13]. It consists of twelve convolutional layers and 4 fully connected layers and for classification a softmax layer. It takes an image of size $224 \times 224 \times 3$ which is then passed to two convolutional layers with 64 filters of size 3×3 . Then max pooling with a stride of 2 is used which reduces the size of the input image to $112 \times 112 \times 64$. Then the third and fourth layers are again convolutional layers with 128 filters of size 3×3 . Then as like previous a maximum pooling layer with a stride of 2 is applied. Hence the image gets reduced to $56 \times 56 \times 128$. The next two layers are convolutional layers with 256 filters of size 3×3 . It is then followed by a maximum pooling layer with filter size 3×3 and a stride of 2. Then two sets of three convolutional layers are done from seventh to twelveth layers. 512 filters of size 3×3 are used. Next, a flattening fully connected layer with 25088 filters is applied. Then two fully connected layers with 4096 neurons are applied. Then finally a softmax layer which classifies into 1000 classes is used. The architecture is shown on Fig. 3.

D. GoogLeNet

GoogLeNet is 22 layer deep convolutional neural network architecture for classifying high resolution images [13]. Inception module is an important layer in GoogLeNet. In AlexNet and VGGNet the filter size is fixed in each convolutional layer. But in GooLeNet different sizes of convolutional layers with max pooling are stacked together in the inception layer. This helps to find different kinds of features from an image. Also, a 1×1 convolution operation is inserted for dimension reduction. Global Average pooling is used at the end of the network to map the feature vectors accurately.

E. ResNet

ResNet [18] is a 152 layer deep residual learning algorithm for image classification. When cascading the network with more layers the training error rate increases after a particular point of time. Residual block is added to ResNet to overcome this problem. A typical residual block is shown in Fig. 5. This replaces identity mapping by skip connections. The input from the previous layer is appended with the output to reduce the degradation problem.

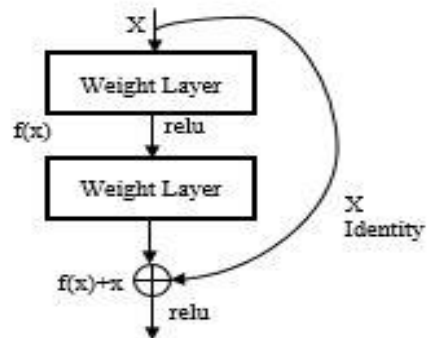


Fig. 5. Residual Block of ResNet

V. EXPERIMENT AND RESULT

A shallow convolutional neural network have designed to work with TCIA (The Cancer Imaging Archieve) public database.

The proposed CNN have totally three layers. Two convolution layers followed by max pooling layers and one dense layer with 120 neurons. The architecture diagram is shown in Fig. 6. Training is done using sample of 851 MRI images with 20% for testing. Adam optimizer is used to minimize the error rate. This gives 94.6% validation accuracy. The training accuracy and testing accuracy are shown using Fig. 7. The training loss and testing loss are shown in Fig. 8. The confusion matrix of the predicted stages of cancer is shown in Fig. 9.

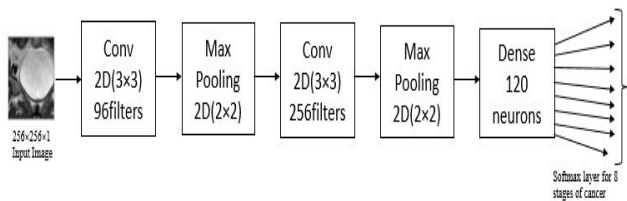


Fig. 6. Architecture of Proposed CNN

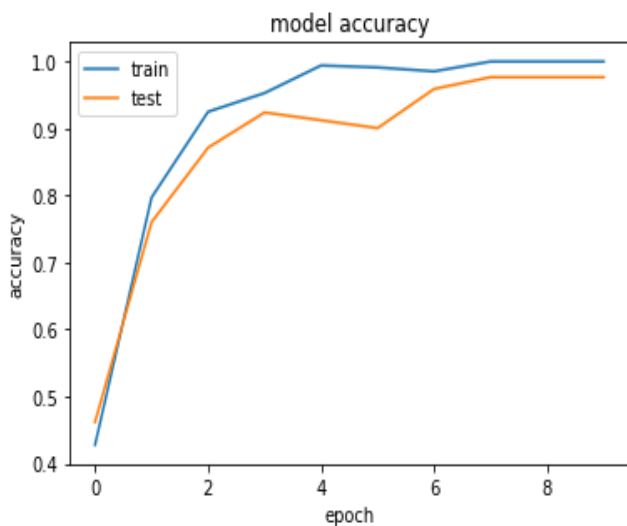


Fig. 7. Training and Testing Accuracy

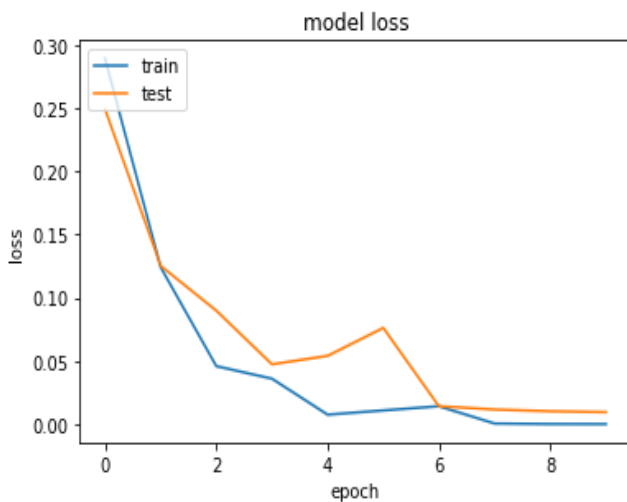


Fig. 8. Training and Testing Loss

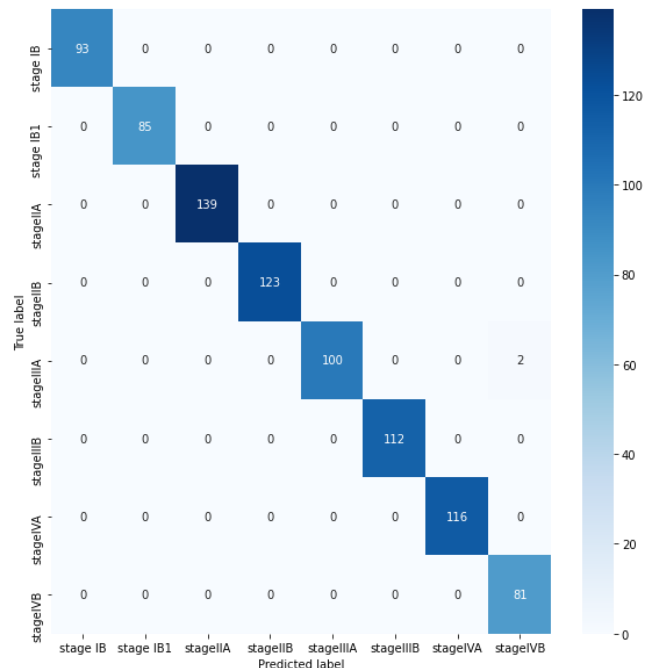


Fig. 9. Confusion Matrix

VI. CHALLENGES

Designing efficient automated tool for diagnosing medical images has the following challenges.

- intra-class variation and inter-class similarity
- diversity of imaging modalities
- Semantic gap that exists between the low level visual information captured by imaging devices and high level semantic information perceived by human.
- Tissue appearance is challenging to characterize
- Describing formally how radiologist recognize images
- Features that characterize high level information

VII. CONCLUSION

In this work a comprehensive study of CNN for automated diagnosing of medical images was done. Comparison of various CNN architecture and study of the layers of CNN was extensively done to carry out future work.

REFERENCES

1. C. C. Reyes-aldasoro *et al.*, "Medical Image Classification with Convolution Neural Network," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2, no. July, pp. 15–22, 2018.
2. S. Khan and S. P. Yong, "A deep learning architecture for classifying medical images of anatomy object," *Proc. - 9th Asia-Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. APSIPA ASC 2017*, vol. 2018-Febru, no. December, pp. 1661–1668, 2018.
3. H. Mohsen, E.-S. A. El-Dahshan, E.-S. M. El-Horbaty, and A.-B. M. Salem, "Classification using deep learning neural networks for brain tumors," *Futur. Comput. Informatics J.*, vol. 3, no. 1, pp. 68–71, 2018.
4. S. Pereira, A. Pinto, V. Alves, and C. A. Silva, "Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images," *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1240–1251, 2016.
5. G. E. H. Alex Krizhevsky, Ilya Sutskever, "ImageNet Classification with Deep Convolutional Neural Networks," *Handb. Approx. Algorithms Metaheuristics*, pp. 1–1432, 2007.

6. E. Gibson *et al.*, "NiftyNet: a deep-learning platform for medical imaging," *Comput. Methods Programs Biomed.*, vol. 158, pp. 113–122, 2018.
7. J. Zhang, Y. Xie, Q. Wu, and Y. Xia, "Medical image classification using synergic deep learning," *Med. Image Anal.*, vol. 54, pp. 10–19, 2019.
8. A. Qayyum, S. M. Anwar, M. Awais, and M. Majid, "Medical image retrieval using deep convolutional neural network," *Neurocomputing*, vol. 266, pp. 8–20, 2017.
9. I. N. B *et al.*, "Multimodal Deep Learning for Cervical Dysplasia Diagnosis," *Miccai 2016*, vol. m, pp. 388–397, 2016.
10. T. Kooi *et al.*, "Large scale deep learning for computer aided detection of mammographic lesions," *Med. Image Anal.*, vol. 35, pp. 303–312, 2017.
11. S. Dara and G. K. Tumma, Priyanka , NR Eluri, "Feature Extraction In Medical Images by Using Deep Learning Approach," vol. 120, no. 6, pp. 305–312, 2018.
12. Y. B. and P. H. Yann LeCun, Leon Bottou, "Gradient-Based Learning Applied to Document Recognition," *IFAC Proc. Vol.*, vol. 15, no. 1, pp. 19–24, 1998.
13. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc.*, pp. 1–14, 2015.
14. S. Liu *et al.*, "Deep Learning in Medical Ultrasound Analysis: A Review," *Engineering*, vol. 5, no. 2, pp. 261–275, 2019.
15. A. S. Lundervold and A. Lundervold, "An overview of deep learning in medical imaging focusing on MRI," *Z. Med. Phys.*, vol. 29, no. 2, pp. 102–127, 2019.
16. A. Maier, C. Syben, T. Lasser, and C. Riess, "A gentle introduction to deep learning in medical image processing," *Z. Med. Phys.*, vol. 29, no. 2, pp. 86–101, 2019.
17. S. Wang and R. M. Summers, "Machine learning and radiology," *Med. Image Anal.*, vol. 16, no. 5, pp. 933–951, 2012.
18. Kaiming He, Xiangyu Zhang, Shaoqing Ren and Jian Sun, "Deepresidual Learning for Image Recognition", Microsoft Research, 2015

AUTHORS PROFILE



Ms.A.Cibi is an Assistant Professor in the department of Computer Science and Engineering, Rajalakshmi Engineering College, Chennai. She received her Bachelor's degree in Computer Science and Engineering in 2007 from St.Xavier's Catholic College of Engineering, Nagercoil. Then she received her

Master's degree in Software Engineering in 2009 from College of Engineering, Guindy Campus, Anna University, Chennai. Her research interests are Medical Image Analysis, Computer Vision and Artificial Intelligence and Deep Learning.



Dr.R. Jemila Rose is as Associate Professor in St.Xavier's Catholic College of Engineering, Nagercoil. She received her Ph.D. degree in faculty of Information and Communication Engineering, Anna University, Chennai in 2016. She received her

Master's degree in the Department of Computer and Information Technology in 2007 at Centre for Information Technology and Engineering, M.S.University, Tirunelveli. She received her Bachelor's degree in Information Technology Department in 2005 at Sun College of Engineering, Nagercoil. She has organized many International Conferences. Also she has authored and co-authored more than 20 publications. She is supervising research scholars of Anna University. Her research interests Medical Image Processing, Denoising, Segmentation, Computer Vision and Artificial Intelligence.