IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

IMAGE ANALYSIS USING NEURAL NETWORKS: A REVIEW

[1]Dishant Totade, [2]Ashish Shrivastava,

[1]Research Scholar, [1]Dept. of Computer Science and Engineering

[2]Assistant Professor ,[2]School of Computer Science and Engineering

Lovely Professional University, Punjab , India

***Abstract*:** Image processing is a very important step in the field of classification. Processed images give better results as compared to raw original images .Many of these images have unique features like texture, intensity, patterns, RGB values, gray level, binary image, heat image etc. It is very important to extract these features using some techniques, how to extract and generate these features will be illustrated in this paper. This paper provides an academic database in between 2012 and 2019. It has been observed that the new generation methods in the space area of image analysis have produced remarkable performance. In this paper, various techniques have been discussed with the different concept of neural networks in image analysis.

***Index Terms* : Image processing, neural networks, deep learning, computer vision, image segmentation**.

I. INTRODUCTION

Nowadays AI is very important part of our lives. It is used vastly in computer vision, Robotics, space research, medical diagnosis etc. In every field image is used as a main source of information to take the decisions. Robots are used to take the decisions with the help of video frame which is an image for the path finding, object finding etc. Industrial robot also relies on the image data for the operation in the manufacturing process of the machine. In Space research, image processing has been proved one of the key technologies. NASA, ISRO and other organizations capturing the deep space images with the help of big telescope and apply certain algorithm to analyze the image for threshold segmentation, edge detection and gradient edge direction etc. These algorithms have been given the desired results.

In medical diagnosis there are various type of images used like clinical analysis. Radiography imaging are of two types fluoroscopy and projection radiography , fluoroscopy gives internal structural of the body in real time image , projection radiography is more of an x-ray projecting through human body onto the film. Magnetic resonance imaging (MRI), in which human body is examined under the polarize magnet which excite the hydrogen molecules of water of human tissue and produce a detectable signal in the form of protons. Ultrasound diagnosis provides on surface imaging of the patients body organ. Ultrasound uses high frequency sound wave that gets reflected by the tissues to get 3D imaging of body organ such as during female pregnancy or kidney stone diagnosis. Tomography is an x-ray projection of the body onto the film in which x-ray beam gets stopped by the bone when it passes through the body and do not appear on the x-ray film. This section gives a glimpse of images in various fields and how it works so that you can have an idea of different images during reading this paper. In section 2 we have presented a literature review of various research papers related to image analysis with neural networks and its various architectures and techniques of processing the image and shown in the tables. In section 3, the conclusion of the paper and discussion about future work has been presented

II. LITERATURE REVIEW

The Literature of neural network is very wide. The work done by various researchers in the field of Image Analysis using Neural Network has been discussed in this section. S N Deepa et al. (2012) classified the brain MRI images to detect Malignant and Benign cancer and to segment the tumors using the standard statistical feature extraction technique for image analysis. They have used two neural network architectures for the classification of the brain cancer and then compared the accuracy of those networks. Back Propagation Neural Network in which one hidden layer with the log sigmoid activation function, one input and output layer are used, initially, network takes the input as training data then feed it forward up to the output layer then output is compared with the target value. If the output value is not near to the target value then the network does back propagation of the weight from the output layer to the input layer until the output value is equal to or near to target value. Next is Radial basis function network in which two layers are used, one hidden and one output layer, the hidden layer takes the feature vector as input and apply the Euclidean distance on vector and the test samples, after finding distance the neuron apply the kernel function for mapping the data into multidimensional space, then Gaussian radial function is applied on each neuron, the weight between the hidden and the output layers are used in least mean square method, at last the value of each neuron is passed to the output layer for the summation of the overall network values. Before applying classification techniques author applied some Statistical analysis methods on pixels of the Brain MRI images, i.e. Energy, Correlation coefficient, contract, Entropy, Homogeneity. For the feature extraction methods i.e. Histogram equalization, region isolation, author implemented the methodology in Matlab 2010A on the image dataset from PSG IMS, Coimbatore, the image size is 512 x 512, 12 bit per pixel image from MRI Avanto 1.5t MRI scanner. The input data involves 42 patients (25 abnormal, 17 normal). For training set 30 normal images are used and for abnormality 12 images are used. when both classification methods applied on the data set RBFN, performed well as compared to BPN . RBFN accuracy is 85.71%, sensitivity 72%, specificity 84%. Whereas BPN performance accuracy is 76.19%, sensitivity is 82.3%, and specificity is 88.23% [1].

Amer Al-Badarneh et. al. (2012) in the proposed methodology the author used ANN with interconnection of the neural network similar to human brain. In this study, similar network is used like perceptron which uses a threshold activation function. In KNN, there is k nearest neighbors in the training data of same type, the training sample is compiled into 2D space using the Euclidean distance. Minimum distance is calculated between the data of testing sample, in this study k =3 features are used and considered their minimum distance as a final decision. The MRI images used in this paper were collected from the Harvard medical school, Benchmark dataset is used. It consists of T2-weight with axial 256X256 pixels with 710 images in which there are 615 abnormal images and 95 normal images. The author obtain feature texture of 275 brain images from which 181 are abnormal and 94 are normal each with 256X256, then extracted features are passed to the neural network and KNN, the output of this classifier is measured in accuracy, precision, recall, F-measure. In neural network classifier there are three cross validation folds, in which 10th fold have shown highest accuracy of 98.55% and with percentile split of 66%, the average accuracy of KNN is 98.92% and Individually KNN have shown highest accuracy with the percentile split of 80% is 100% [2].

Riries Rulaningtyas et. al. (2012), for the classification of the tuberculosis mycobacterium, the proposed methods have two stages, feature extraction and then classification using neural network. In first stage, the author used geometry feature extraction techniques in which there are seven attributes perimeter, area, radii, circularity, compactness, and eccentricity, tortuosity on the basis of these attributes features are extracted and passed to the neural network. In the neural network weights are initialized and overall output of the network is calculated, if the output is not equal to or near to target value error is calculated on each epoch and the back propagation is applied and start updating the weight from the backward direction, this process keep on going on each epoch. The research uses 100 sample from which 75 are used for training purpose and 25 are used for testing purpose. The network architecture is optimum with 1 hidden layer with 20 neurons in it with 0.5 learning rate and 0.9 momentum rate. The gradient error is evaluated with 0.0011967 at 5000 epoch and cross validation failed out zero at 5000 epoch, for the analysis of the neural network linear Regressing is applied for the calculation of the output value, how much it is near to the target value, which is 0.99942 (almost 1) which means it is closed to target value. For the future work author suggest using appropriate Image processing technique color segmentation to extract the feature from the bacteria images for classification [3].

Belvin Thomas et. al. (2013). Texture analysis and run length analysis of the color image and using it to classify the oral cancer lenion into 6 groups of malignancy. The proposed method is implemented using Matlab 7.14 and image data set is collected from Himalayan institution of the medical science, Dehradun, India. The images are recorded using Sony cyber shot digital still picture camera (4X optical zoom, 10.1mp). Image dataset consist of 16 images of malignant oral lesion. Four feature extraction techniques are implemented based on their number of optimal features, then using these features back propagation neural network is trained for (GLCM) grey level Co occurrence metric with 36 hidden neuron with accuracy of 89.58%. For (GLRL) grey level run length, technique with 11 optimal features, with 31 hidden neurons of network has given accuracy of 91.66%. With all 61 feature with 16 hidden neurons has given accuracy of 91.66%, with limited selected 11 features with 25 hidden neurons has given accuracy of 97.92% [4] .

Fatma Taher et. al. (2013). In this paper, author used mean shift algorithm and histogram analysis for the segmentation of the sputum cell into nuclei and cytoplasm region. The database has 100 sputum color images which are obtained from Tokyo centre of lung cancer in Japan. The size of each pixel is 768X512, additionally, each image mask is provided of cytoplasm and nuclei area. The result of the proposed method is compared with the result of HNN. For the performance measurement author used true positive, false positive, true negative, false negative, precision, specificity, accuracy. For gray mean shift accuracy 82.07%, precision 58.53%, specificity 82.28% for gray-space mean shift, accuracy 87.21%, precision 59.08%, specificity 89.74% [5].

Chen Ding et. al. (2014) have used the histogram and threshold values of image to segment the vasculature of retinal eye image, initially, the color RGB image of retinal is converted into the grey level image. In the drive dataset there are 20 retinal image and 20 artificial segmented images, for the training of neural network the parameters used by author is statistically histogram properties of the gray level images and supervised threshold values of artificially segmented images as input. The threshold values is decide based on the different distributions on the grey level through the pixel, once the threshold is figure out then for segmentation g(i,j) = { 1 f(i,j)>=T

, else 0 } is computed. These threshold methods are applicable to some images only because image brightness and color is not distributed evenly, that's why multiple thresholds are computed, for multiple threshold values blocking of image is used, in which 9 x 9 matrix is formed of pixels for the correlation supervised threshold. This 9x9 matrix is slide throughout the pixels of image for the segment of blood vessels by choosing the minimum error threshold by comparing at the same position with the manually segmented image pixels. Input vector for the neural network is of 1 x 256 matrixes, as histogram parameter and 1 X 1 matrix of threshold values, input and output size is 256 with 256 nodes, 200 iterations with one hidden layer the accuracy of classifier and NN is 97.43% [6].

Jaspinder Kaur et. al. (2014) in this paper, two different feature extraction techniques is used to train the feed forward back propagation neural network for the preprocessing of the images. Noise is removed with the help of median filter and for lung segmentation optimal threshold is used. Two feature extraction techniques are textural and statistical. Textural parameter like area for calculating particular pixel intensity value, energy for the intensity of pixel brightness, correlation among the pixels, homogeneity for the closeness among adjacent pixels, entropy for randomness, statistical feature parameter are mean, standard deviation, skewness, kurtosis. The database contains 547 images of 10 patients, which is publicly available provided by 'lung images database consortium'. The size of image is 512X512 resolutions, for network training the input layer takes 6 elements from textural parameters and 6 element form statistical parameters are selected and compared the result of both technique. The result from this methodology shows the accuracy of 98% for textural parameter and 94.5% for statistical parameter. Future work of this proposed methodology is to apply this technique on MRI, X-ray image and to extend the database and to finding of tumor stage [7].

Mingchen Gao et. al. (2016) in this paper, author used manually labeled pixel data for the classification of lung into five classes i.e. ground glass, fibrosis, micro nodule, emphysema, healthy. Interstitial lung disease - data-set have manually annotated images of lung by the radiologist. The problem with data-set is that it is only single class annotated. So author used fully connected (CRF) conditional random field to classify the unlabeled region of interest. The classifier predicts the labeled annotation with the accuracy of 92.8% [8].

Y.Hong et. al. (2017) used stacked BCN Brief convolution network for the segmentation label of hearts coronary artery. Luminal and wall label are predicted for the segmentation. They have used pre trained BCN auto-encoder as it gives better segmentation result. They also proposed a technique of probability amplified feature for the dimension of label mask in the image, it is increasing in the accuracy. The architecture of BCN having a set of compressing and decompressing stages in the network. There are 3 convolution layers, 1 max pooling, 2 pixel striding layer and same in decompressing stage but max pooling is replace with up sampling. For the denoising of sonograph image auto-encoder and for the label prediction BCN is used both were implemented in serial stage to get the desired output. The dataset consist 50 patient of Coronary computed tomography angiography (CCTA) intravascular ultrasound (IVUS) dataset. It is splitted into 40 sets into training and 10 sets into validation. Each set consists of 3 view of artery image i.e. axial, coronal, sagittal view. The result of this prediction is measured in Dice Similarity coefficient and it gives accuracy for luminal of artery 0.63 and for wall of artery 0.78 [9].

Dr. Z. Faizal khan et. al. (2017) solved the error occurred in the lung segmentation process, by proposing the textural base echo state neural network. The proposed methodology have two stages, first extracting the features with the help of Gray Level Co-Occurrence Matrix (GLCM) attribute and then in second stage passing those features to neural network. In GLCM, energy of pixels is calculated which is summation of square of pixels intensity value. Homogeneity returns the closeness of pixels value distribution. Contrast are the difference of each pixel intensity value, correlation returns how much a pixel is related to its adjacent pixel. These features give input to echo state neural network. In echo state neural network, there is input layer, reservoir, and output layer. Nodes between the hidden layer and input layer are connected to output layer which is architecture replica of recurrent network. They have also proposed the algorithm of ESNN which figures out input and connections randomly. The proposed algorithm for the segmentation of the image gives the accuracy of 97.8%, with sensitivity and specificity of 90.3 %, 98.2% respectively. The lung data-set used is taken from LIDC (Lung Image database Consortium) database [10].

Dimitris Perdios et. al. (2018) proposed a technique to improve the quality of ultrasound images. Author used simulated platform to create a training set. The data-set name is PICMUS (frame work for ultra sound image). In this, low-quality images and high-quality images are re-constructed from single plane wavelength to multiple plane wavelengths on 2D plane. Using of simulated data allow control over distribution of variable echogenicity. Later this data is used to train the U-net neural network. U-net architecture is in the form of U shape. It performs recurrent of every left side layer to right side layer with auto-encoder structure. The encoding part performed on left side and decoding part performed on right side. The structure is build of 5 decomposition layers. Author used three types of U-net and compared the result i.e. U-net16, U-net32, and U-net64. Compared to the single plane wavelength ultrasound image all three U-net enhance the CNR (Contrast-to-noise) ratio of image by 117%, 125%, 138% respectively [11].

Goran Jakimovski et. al. (2018) deep neural network is used to recognize the lung cancer by magnetic resonance imaging scan that is determined by multiple pattern of tissue. They have used unsupervised sparse auto-encoder machine to accelerating DNN to 90.1%. DNN requires data-set that is predetermined to recognize pattern. It works with algorithm to classify these images into predetermined classes. They used multi-stages scheme to overcome the over fitting problem. Author used skin tissue cancer to examine DNN and put into piles of yes or no (yes there is cancer or no). Data is of a CT images have taken form local medical hospital of 70 patient where 58 having cancer and 12 not having cancer. These images are used for 10% testing and 90% for training in a subset [x,y], Shuffled and converted into binary class matrix so it can be fed to neural network. DNN requires large data-set, it gets it from parameter of argument flipped, rotate, skewed so that more version of same image used in training process. By developing these parameter, the Deep neural network examine [12].

Priya et. al. (2019) Two-step convolution neural network as deep learning framework has been proposed for extraction of road from aerial or satellite image to discuss the term such as accuracy, precision, recall and f-score. CNN model composed of two steps in which U-net is used in each step. Image passed to the first U-net which converts color aerial to gray image and then passed to second U-net extracts the roads in the image. Training of model needs data-set which has original size if 1500 X 1500 pixel and at least 5% pixels of image should be road pixel . further cropped to 512 X 512 pixel to get less complexity. Two-step CNN model works with U-net each step have multiple convolution layer and have left and right side . left side goes downward direction and right side upward direction and these two sides are attached with entire connected layer. This model work with biomedical segmentation. Leaky Relu (Rectified Linear Unit) gives good precision measure and ELU (Exponential Linear Unit) Gives good recall measure . Massachusetts data were used in this model and implemented with keras along the tensor-flow as a back-end . This CNN model solved two problem, precision gave exact pixel throughout all predicted pixel and recall gave exact pixel throughout all real pixel. [13]

Table 1. Review between 2011 to 2019

Author	Technique use	Data-set	Results/Accuracy
S N Deep et al [1] (2012)	Statistical feature with BPN and RBFN	PSGIMS Coimbatore	RBFN : 85.71% BPN : 76.19%
Amer Al-Badarneh et al [2] (2012)	KNN and Perceptron	Harvard medical school, Benchmark dataset	Neural network: 98.55% KNN : 98.92%
Riries et al [3] (2012)	Geometric feature with BPN	Review	Review
Belvic Thomas et al [4] (2013)	GLCM and GLRL With BPN	Himalayan institution of medical science ,Dehradun	GLCM : 89.58% GLRL : 91.66%
Fatma et al [5] (2013)	Mean Shift Algorithm and Histogram analysis	Tokyo center of lung cancer , Japan	MSA **:** 82.07%
Chen ding et al [6] (2014)	Review	20 retinal images, 20 artificial segmented	Neural Networks:97.43%
Jaspinder Kaur et al [7] (2014)	Textural and statistical feature with BPN	Lung database Consortium	Texture parameter:98% Statistical parameter: 94.5%
Ming chen et al [8] (2016)	CRP (Review)	Review	Classifier : 92.8%
hong et al [9] (2017)	Stacked Brief CNN	50 patient of IVUS dataset	Luminal prediction : 0.63 Wall prediction : 0.78
Dr .Z Faizal khan et al [10] (2017)	Textural base Echo state Neural Network	Lung image database consortium	Accuracy : 97.8%
Dimitris perdios et al [11] (2018)	Review	PICMUS	Review
Goran Jaikimovski et al [12] (2018)	Sparse Auto- Encoder	MRI of 70 patient, 58 having cancer	Review
Priya et al [13] (2019)	Two- set CNN of U-net	Massachusetts data	Review

III. METHODOLOGY FOR SEGMENTATION

For the Image analysis we have discussed various techniques and utilizing various data-set in different area field. The method that we are discussing in this paper have two step implementation , first apply the edge detecting on the image after that apply Mask R-CNN (Masked Region Based Convolution Neural Network) this will give the exact desire segmented region . R-CNN is a instance segmentation technique which is used to segment the object in the image.

3.1 Architecture of image segmentation using edge extraction with Mask R-CNN

Edge extraction with Mask R-CNN will give the better result for desired object segmentation in image . For edge detection apply mean threshold (t) pixel intensity function(f) on image . x,y are the co-ordinates of pixels. The equation (1) is specified below.

$$f(x,y) = \{\ 1 \ \text{ if } t >= I(x,y) \text{ else } 0\ \} \qquad (1)$$

This function will extract the edges. Further by adjusting the threshold , function will give the desire intensity edges. After extracting the edges, pass the image to mask R-CNN . Mask R-CNN involves two stages, in first stage appropriate backbone neural network model like res-net , vgg etc is used and trained for the object detection that needs to be segmented in the final output but the last layer of this backbone is remove because we only want the pixel matrix of feature map and then passed to the next region proposal layer this help to find the relevant region of object .Then in second step , network predicts the bounding box over the object with the help of regression as define in class set of object and then, at the end their is fully connected layer with the Roi-Align function which use bilenear interpolation to extract mask region of object from the bounded box feature map. Figure 1. of the network is shown below.

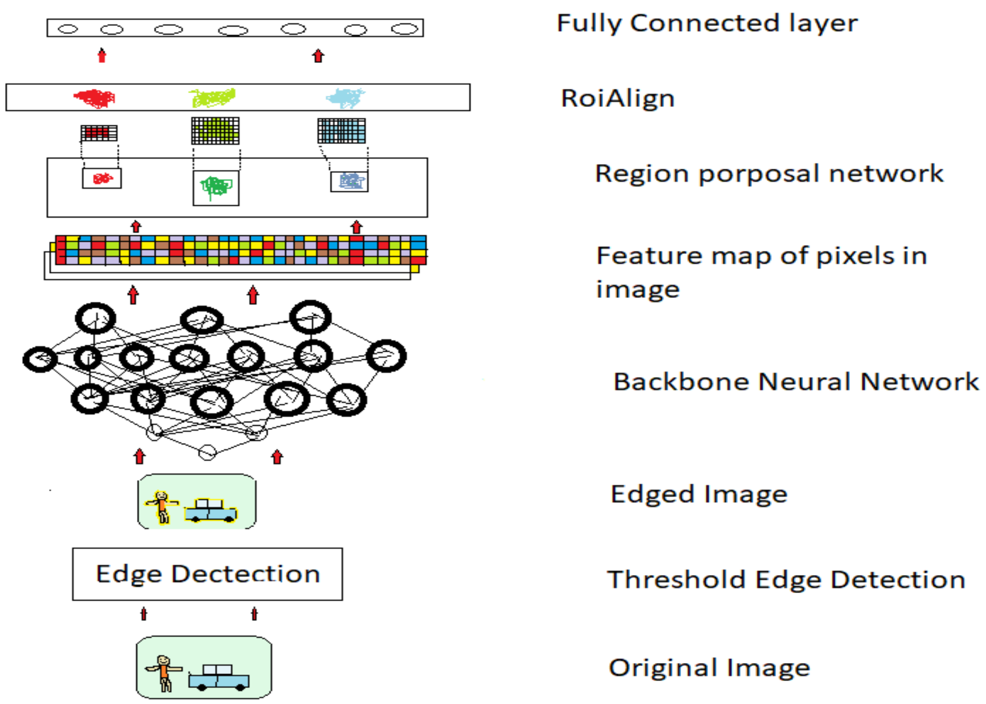

Figure 1 . Segmentation using edge detection with mask R-CNN

Backbone Neural Network : It is basically a standard neural network model which is used to train the network for desired object that needs to be located e.g dog ,cat, car, person etc but last fully connected layer of this model is remove so that we get only the feature map of the object.

Pixel of feature map: It is a pixels collection of the object in image, that needs to be segmented and have densely feature map.

Region Proposal network: The main purpose of this layer is to generate the bounded box around the densely pixel feature map so that the RoiAlign would easily extracted the region that needs to be masked segmented.

RoiAlign: It is very important step in Mask RCNN . Standard Roi pooling extract feature from each Roi but with this unwanted pixels also get consider. The problem with Roi Pooling layer is if Roi doesn't align perfectly with mapped feature the pixel-to-pixel alignment would be disorder. RoiAlign is improved version of Roi pooling , it transforms accurately the features of Roi into fixed size vectors without including any unnecessary pixels with the same interval of Roi pixel. Thus, the network produces perfectly masked images.

IV. CONCLUSION AND FUTURE WORK

From the Literature Review , it is found that there are various image processing techniques and they give different results with neural networks . The most used network for classification is feed forward back propagation neural network because it has a tendency to adjust the weight according to desired output. Back propagation neural networks work best with segmented images and give the best result . From the review it has been observed , most researchers prefer gray level images for the classification approach because textural features can easily be extracted from those images based on the gray intensity value which is 0 to 255(0 means most black and 255 means most white). As a segmented image gives the best result we have given a technique which gives a good segmented image for both color and gray level image.

In the future, research would be to use more appropriate image processing for the gray level image and try to add multiple hidden layers in the back propagation neural network.

References

[1] S. N. Deepa and B. A. Devi, "Artificial neural networks design for classification of brain tumour," 2012 Int. Conf.Comput. Commun. Informatics, ICCCI 2012, 2012.

[2] A. Al-Badarneh, H. Najadat, and A. M. Alraziqi, "A classifier to detect tumor disease in MRI brain images," Proc. 2012 IEEE/ACM Int. Conf. Adv. Soc. Networks Anal. Mining, ASONAM 2012, pp. 784–787, 2012.

[3] R. Rulaningtyas, A. B. Suksmono, and T. L. R. Mengko, "Automatic classification of tuberculosis bacteria using neural network," Proc. 2011 Int. Conf. Electr. Eng. Informatics, ICEEI 2011, no. July, pp. 17–20, 2011.

[4] B. Thomas, V. Kumar, and S. Saini, "Texture analysis based segmentation and classification of oral cancer lesions in color images using ANN," 2013 IEEE Int. Conf. Signal Process. Comput. Control. ISPCC 2013, 2013.

[5] F. Taher, N. Werghi and H. Al-Ahmad, "Automatic Sputum Color Image Segmentation for Lung Cancer Diagnosis," KSII Transactions on Internet and Information Systems, vol. 7, no. 1, pp. 68-80, 2013.

[6] C. Ding, Y. Xia, and Y. Li, "Supervised segmentation of vasculature in retinal images using neural networks," IEEE Int. Conf. Orange Technol. ICOT 2014, pp. 49–52, 2014.

[7] J. Kaur, N. Garg, and D. Kaur, "An automatic CAD system for early detection of lung tumor using back propagation network," 2014 Int. Conf. Med. Imaging, m-Health Emerg. Commun. Syst. MedCom 2014, pp. 257–261, 2014.

[8] M. Gao et al., "Segmentation label propagation using deep convolutional neural networks and dense conditional random field," Proc. - Int. Symp. Biomed. Imaging, vol. 2016-June, pp. 1265–1268, 2016.

[9] Y. Hong et al., "Coronary luminal and wall mask prediction using convolutional neural network," Proc. - Int. Symp. Biomed. Imaging, pp. 1049–1052, 2017.

[10] Z. F. Khan, A. S. Al Sayyari, and S. U. Quadri, "Automated segmentation of lung images using textural echo state neural networks," 2017 Int. Conf. Informatics, Heal. Technol. ICIHT 2017, pp. 1–5, 2017.

[11] D. Perdios, M. Vonlanthen, A. Besson, F. Martinez, M. Arditi, and J. P. Thiran, "Deep Convolutional Neural Network for Ultrasound Image Enhancement," IEEE Int. Ultrason. Symp. IUS, vol. 2018-October, no. 20, pp. 1–4, 2018.

[12] G. Jakimovski and D. Davcev, "Lung cancer medical image recognition using deep neural networks," 2018 13th Int. Conf. Digit. Inf. Manag. ICDIM 2018, pp. 1–5, 2018.

[13] P. Singh and R. Dash, "A Two-Step Deep Convolution Neural Network for Road Extraction from Aerial Images," 2019 6th Int. Conf. Signal Process. Integr. Networks, SPIN 2019, pp. 660–664, 2019.