iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 Issue: VII Month of publication: July 2021

DOI: https://doi.org/10.22214/ijraset.2021.36378

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VII July 2021- Available at www.ijraset.com

Survey on Cardiomegaly Detection with Enhanced X-Ray Using Deep Learning

Prof. Vidyashree K P[1], Ashitha. P. Gowda[2], Daneshwari M Ramannavar[3], Suraj M S[4], Niranjan Dev[5]
[1, 2, 3, 4, 5]Dept. of Information Science, VVCE, Mysuru

Abstract: Cardiomegaly is an augmented (enlarged) heart. It is not a disease merely a sign of another condition. Cardiomegaly in the early stages, which is less severe, is called mild Cardiomegaly. These complications may give rise to conditions like Blood clots, Cardiac Arrest and sudden death, Heart failure, Heart murmur. Hence, detecting these kinds of states in the early stages helps in improving medications and reduce complications. In this paper, we will present various approaches available to detect this development using Deep learning and offer an automated system to detect the presence of Cardiomegaly in a patient. For the computerized system, we are using deep learning concepts such as U-Net and VGG16 and image enhancement (image prepossessing is required) methods like Unsharp Masking, Contrast Limited Adaptive Histogram, High-Frequency Emphasis Filtering. Accurate measure of CTR (cardiothoracic ratio) calculations can effectively diagnose the presence of Cardiomegaly in a patient.

Keywords: Deep learning, CNN, U-Net, VGG16, Feature Extraction, Unsharp Masking, Contrast Limited Adaptive Histogram, High-Frequency Emphasis Filtering.

I. INTRODUCTION

As mild Cardiomegaly does not always promptsymptoms, many with an insignificantly enlarged heart areoblivious of this problem. Though, other people may have permanent Cardiomegaly. It is vital to treat it before it causes more intricacies or damage to the heart. Chest X-ray, or CXR, is popularly used in diagnosing unusual conditions in the chest and nearby structures. Radiologists regularly perform cardiothoracic ratio (CTR) measurement on anteroposterior (relating to or directed towards both front and back) chest radiographs to detect Cardiomegaly. This condition is heavily associated with both congenital and congestive heart diseases. Most Picture Archiving and Communication Systems (PACS) include drawing tools to serve calculations CTR; the process is still labor-intensive and time-consuming. Manual labeling of organ boundaries and CTR calculation are predisposed to error and can lead to erroneous diagnoses.

A. Deep Learning

Deep learning is an artificial intelligence (AI) function that replicates the human brain's mechanism in process knowledge and making decision-making patterns. It is a sub-set of machine learning with networks capable of learning neglected from unstructured or unlabeled knowledge. Hence referred to as deep neural learning or deep neural network. Deep learning could be a key technology behind driverless cars, sanctioning them toacknowledge a stop sign or differentiate a pedestrian from a lamppost [19]. A computer model learns to performclassification tasks directly from pictures, text, or sound from deep learning.

Deep learning has grown with the digital era that has increased data in all forms worldwide. Deep learning exhibits significant amounts of unstructured knowledge that will take humans an extended interval to comprehend. Data identified merely as tremendous knowledge [20] is collected from social media, net search engines, and e-commerce platforms. This considerable quantity of information is accessible and shared through fine-tech applications like cloud computing. Deep learning desires innovative computing power. Once deep learning is combined with clusters or cloud computing, it permits development groups to scale back coaching time. Feature engineering is the method of obtaining optionsfrom data to describe the underlying downside. A deep learning method can scan the information to look for options that correlate and blend them to change quicker learning while not being expressly told to do so. Unstructured knowledge is tough to examine for many machine learning algorithms. Deep learning algorithms aretrained using varied knowledge formats and still acquire insights that are relevant to the aim of its training.

With deep learning assistance, well-labeled knowledgeis not needed as deep learning algorithms are intelligent atlearning without guidelines. The quality of its work is never reduced unless the training knowledge involves datathat does not represent the matter they are attempting to resolve.

B. CNN

CNN represents a massive breakthrough in image recognition. They're most usually used to parse visual imagery and frequently work behind image classification scenes. A CNN learned features with input data and uses 2D convolutional layers. It means that this type of network is exemplary for processing 2D images. Corresponding to other image classification algorithms, CNN uses barely any preprocessing. It means that theycan learn the filters that have to be hand-made in different algorithms. CNN's have an input layer, and an outputlayer, and hidden layers. The hidden layers habitually consist of convolutional layers.

A CNN works by eliciting features from images, henceeliminating the need for old-fashioned feature extraction. The features are not trained; they're learned while the network trains on a set of images. It makes deep learning models extremely specific to computer vision tasks.

CNN's learn feature detection over tens or hundreds ofhidden layers. Each layer raises the complexity of thelearned features.

Convolutional Neural Networks, commonly called CNN, is a deep neural network class most commonly usedin analyzing visual imagery. CNN's are regularized versions of multilayer perceptrons. The "fully- connectedness" of these networks makes them prone to over-fitting data.

Over-Fitting data - Overfitting refers to a model that models the training data too well.

Overfitting happens when a model learns the feature and noise in the training data to the degree that it negatively reshapes the model's performance on new data. It means that the noise or random vacillations in the training data are picked up and learned as the model's concepts. The problem is that these concepts do not apply to new data and negatively impact the model's capacity to generalize. CNN is made up of neurons with learnable biases and weights. Each neuron receives multiple inputs and then takes a weighted sum over them, where it passes it through an activation function and responds with an output.

CNN is a complex feed-forward neural network. CNN's used for image classification and recognition because of its high precision. CNN's are used in many domains, including image and pattern recognition, speech recognition, natural language processing, and video analysis. There are many reasons why CNN's are becoming essential. In traditional models for pattern recognition, feature extractors are hand-designed. Inrecent days CNN is hugely popular because of its architecture. The best thing is there is no need for feature extraction. The system learns to do feature extraction. Thecore concept of CNN is that it uses convolution of images and filters to generate invariant features passed on to the next layer. The following layer features are convoluted with different filters to generate more invariant and abstract features. The process continues till one gets the final feature/output, which is invariant to occlusions.

C. U-NET

U-Net is a convolutional neural network (CNN) first developed by Olaf Ronneberger, Philipp Fischer, and Thomas Brox in 2015, for biomedical image segmentationat the Computer Science Department of the University of Freiburg, Germany [7]. The network is based on the fully convolutional network, which Evan Shelhamer, Jonathan Long, Trevor Darrell developed in 2014[15]. Its architecture was modified and extended to work with fewer training images to comply with more precise segmentation. Segmentation of a 512 × 512 image takes less than a second on a modern GPU. U-net consists of two paths, namely, Contraction Path and Expansion Path. The encoding part is also called a Contraction path. Generally, an encoding part consists of iterated convolutions, i.e., repeatedly passing the image through convolutional layers of 3x3 convolutions (unpadded). Each convolution is followed by a ReLU and a 2x2 max pool layer with stride 2 for downsampling. At each downsampling step, it doubles the number of feature channels. It captures context via a compact feature map. Compact feature map- The feature map is the output ofone filter applied to the previous layer. A given filter is drawn across the entire previous layer, moved one pixel ata time. Each position results in the neuron's activation, and the output will be collected in the feature map. The decoder part, also called the expansion path, consists of upsampling the compact feature map followed by a 2x2 convolution ("up-convolution"). It halves thenumber of features and channels a concatenation with the cropped feature map from the contracting path by 3x3 convolutions, followed by a ReLU. Upsampling of the feature map is done to meet the same size block produced in the previous steps, which then is concatenated to the one on the left, which preserves the localization of the image. ReLU-ReLU is a rectifier is an activation function and is defined as the positive part of its argument: {\displaystyle f(x) =x^ {+} =\max} where x is the input toa neuron. There are many applications of U-Net in biomedical image segmentation, such as brain image segmentation "BRATS" [16] and liver image segmentation "siliver07" [17], CardioXnet(Cardiomegaly detection)[],Variations of the U-Net have also been used in reconstruction of medical image [18].

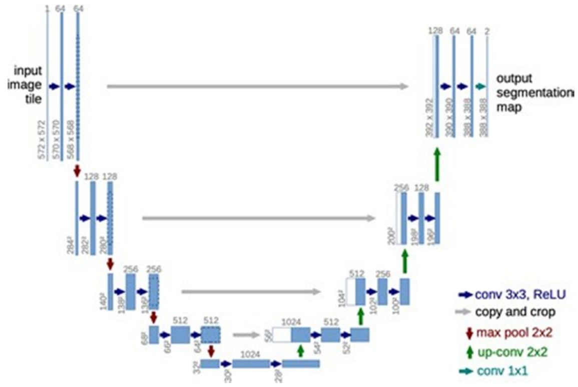


Fig.1: U-NET [7]

D. VGG

To increase the precision of the large-scale imageclassifications, we use VGG16 to recognize image segments more accurately. The architecture of VGG16 consists of Maxpool Layer and five convolutional layers. Figure 2 shows the basic architecture of the VGG16 network. [22] Here, we have two convolutional layers andone Maxpooling layer, then again one Maxpooling layer, and it will go on.

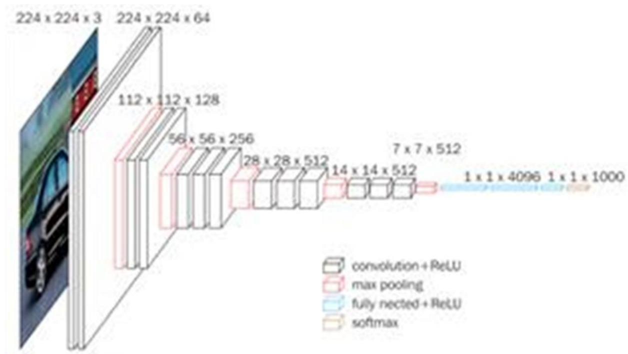


Figure 2: VGG16 - Convolutional Network for Classificationand Detection.

As we see in fig2, we have considered the input image of size 224 × 224 × 3. This image passes through a VGG16 architecture, firstly through convolution layers to keep the size of the input and output feature the same; the convolutional layer uses row and column padding the resolution after this process remains the same. After each convolution block, The MaxPooling process reduces the image's dimensions by 2 × 2. [22]. A variant of the fully- convolutional neural network (CNN), which employs a VGG16 as the encoder, is used for image segmentation. As VGG16 uses 3 × 3 Kernel size more piercingly, the features will be extracted from our input image, comparedto when we have fewer layers. The number of trainable parameters will be $27K^2$ as compared to 7 x 7 kernel size when taken gives $49K^2$ trainable parameters.

II. FEATURE EXTRACTION

Feature extraction is a part of the dimensionality reductionprocess, in which an introductory set of the raw data is split and reduced to more controllable groups. So when we want to process, it will be easier. The most significant characteristic of these large data sets is that they have a large number of variables. These variables need manycomputing resources to process. So Feature extraction helps to get the most beneficial feature from those big datasets by selecting and combining variables into features, thus, efficiently lessening the data amount. These features are easy to process but can still represent the actual data set with precision and authenticity.

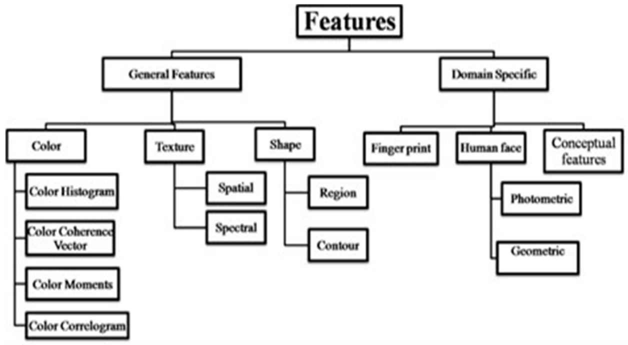


Figure 3: Classification of feature Extraction [21]

The technique of extricating the features is proper when we have a large data set and need to decrease the number of resources without losing crucial or relevant information. Feature extraction helps to lessen the amountof irrelevant data from the data set.
The data reduction helps build the model with more concise machine efforts and increases the machine learning process's learning and generalization steps.

Feature type	Properties	Models
Color Based Features	Impression, expression and construction, RGB, LUV, HSV and HMMD	a. Contourlet Transform b. Steerable Pyramid c. Gabor wavelet Transform d. K-means based CIS
Texture Features	Homogeneity, entropy, contrast, correlation, sum of square variance, spectral and spatial	a. Gaussian Markov Random Field (GMRF) model b. Homogeneous Texture Descriptor (HTD) c. LoG method d. gLoG method e. HLoG method f. Difference of Gaussian (DoG) method g. Support Vector Machine (SVM)
Intensity features	Mean, Median, Standard Variance, Intensity, Skewness	a. Gaussian mixture model (GMM) b. Stochastic model c. Probabilistic model
Human features	Body shape, size, color, age-group, age, gender, height	a. SVM b. Relevance Vector Machine (RVM) c. Prototype Learner (Prot) d. K means e. Histogram of oriented gradients (HoG)
Finger print features	Arches, Loops, Whorls	a. Fuzzy models b. Markov model c. Fingerprint individuality model d. Stochastic model
Conceptual features	Generic product/object knowledge, flexibility, attributes, mutability	a. Generic process model b. Product/object model c. Feature-based model
Text features	Synonymy, polysemy, circularity, irregularity, area, perimeter, roundness	a. Gaussian Markov RMF b. Fractal model c. Probabilistic model d. Simultaneous autoregressive model e. Vector space model

Table 1: Types of features, their properties and models [21]

III.UNSHARP MASKING

Unsharp masking (USM) is an image sharpening technique frequently obtainable in digital image processing software. Its name originates from the fact that the method applies a blurred, or "unsharp," negative image to create a mask of the original image.
The unsharp mask is then blended with the initial positive image, creating an image that is less blurry than the original. The resulting image, although more precise, maybe a less authentic representation of the image's subject. In the circumstances of signal processing, an unsharp mask is generally a linear or nonlinear filter that augments the high-frequency segments of a signal.
Fig 4.1 stewards the original image, and Fig 4.2 stewards the image obtained after employing unsharp masking.

Figure 4.1: Original image

Figure 4.2: Processed image

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VII July 2021- Available at www.ijraset.com

The unsharp filter is a simplistic sharpening operator that enhances edges (and other high-frequency components in an image) via a procedure that subtracts an unsharp or smoothed version of an image from the original image.

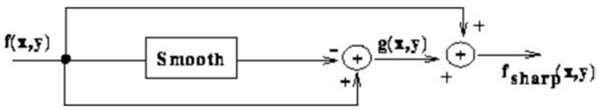


Figure 4.3: Complete working of unsharp maskingFrom the figure we get the equation

$$f_{sharp}(x,y) = f(x,y) + k * g(x,y)$$

More information on unsharp masking can be obtained by referring to Image Enhancement via Adaptive Unsharp Masking by Andera and others [23].

IV. HISTOGRAM EQUALIZATION

***Histogram* -** A digital image histogram is a distribution of its discrete intensity levels in the range [0, L-1]. The distribution is a discrete function h associating to each intensity level: r_k the number of pixels with this intensity: n_k.

Histogram equalization is a method to process images to adjust the contrast of an image by modifying the intensity distribution of the histogram.

This technique aims to give a linear trend to the cumulative probability function associated with the image.The cdf is a cumulative sum of all the probabilities lyingin its domain, given by:

$$cdf(x) = \sum_{k=-\infty}^{x} P(k)$$

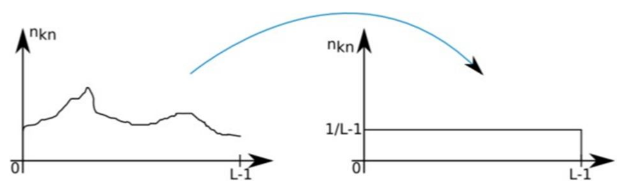


Figure 4.4: Technique to perform histogram equalization

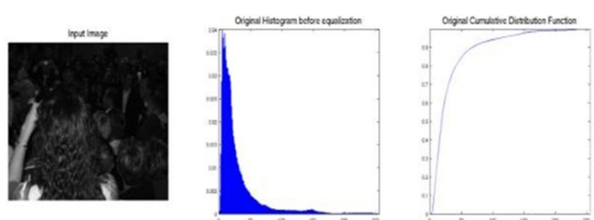


Figure 4.5: Before histogram equalization

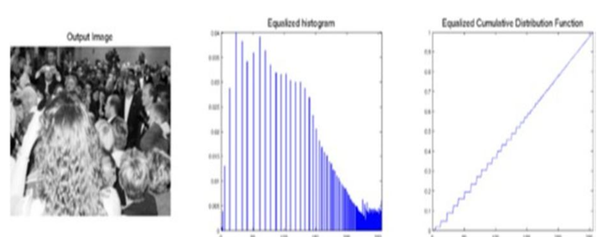


Figure 4.6: After histogram equalization

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VII July 2021- Available at www.ijraset.com

V. CONTRAST LIMITED ADAPTIVE HISTOGRAM

This method is a histogram-based method used toenhance contrast in images. This technique calculates a histogram for the region around each pixel in the image, enhancing the local contrast and enhancing the edges in every region.

It differs from conventional histogram equalization in the respect that the adaptive method calculates several histograms, each corresponding to a discrete section of the image, and uses them to redistributethe image's lightness values. It is, therefore, suitable for improving the local contrast and improving the definitionsof edges in each region of an image.

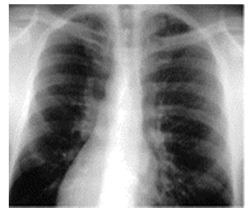


Figure 5.1: Original image **Figure 5.2:** Processed image

VI. HIGH FREQUENCY EMPHASIS FILTERING

High-frequency Emphasis filtering is a method that uses Gaussian High Pass Filter to articulate and accentuate the edges. The edges tend to be represented in the high-frequency spectrum since they have more extreme changes of intensity. This technique provides a low contrast image, and the use of Histogram Equalizationis required to enhance both sharpness and contrast.

Steps to be followed while applying High- Frequency emphasis filtering are:

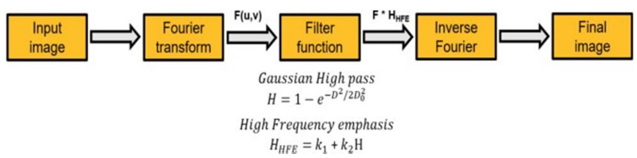


Figure 6.1: Fourier-domain filtering to apply High-frequencyemphasis filter

Step 1: Load Image

Step 2: Compute the Fast Fourier transform and FFT Shift of the original image

Step 3: Compute the HFE filter using a Gaussian High-Pass filter.

Step 4: Apply the HFE filter (by multiplying HFE with the FFT of original image)

Step 5: Perform inverse Fourier transform and generate an image to view the results.

Figure 6.2: Original image Figure 6.3: Processed image

VII. LITERATURE REVIEW

One of the main things to do before cardiomegaly detection is image pre-processing. Now by image pre-processing, we mean image enhancement. Many researchers have proposed new techniques for image enhancement. For more information on image pre- processing, we would like to refer to a Review on Image Enhancement Techniques []. Reference: A review on Image Enhancement techniques, Haris Ačkar, Ali AbdAlmisreb, Mohamed A. Saleh.

Newer studies have shown various advanced methods for image enhancement like the use of Colour Limited Adaptive Histogram for image enhancement [11],

Unsharp Masking [12], and adaptive Unsharp masking [13] also High-Frequency Emphasis Filtering [14]. Image segmentation is done in various ways, and numerous papers are published on image segmentation.

An image repository was searched for the nearest neighborof the patient's sample X-ray image using KNN in early work. A SIFT flow algorithm was used to align and transform lung boundary from the nearest neighbor image to the sample [1]. We refer readers to a broad survey by Zaitoun [4] for further information on early image segmentation approaches. More information on segmentation and its working can be found in P. Moeskpos's [11] IEEE paper. Various researchers have provided research papers to find the CTR ratio by segmenting the patient X-Ray; Ebenezer and Rao [2] applied an Euler number-based approach to finding the best threshold separating the two lungs from the background. Candemir et al. The word "data" is plural, notsingular.

"Automatic heart localization and radiographic index computation in chest x-rays" [3] register and compares theinput chest image with the most similar image in the modeldataset. The similarity is then measured by calculating the Bhattacharyya distance of the X-ray and intensity histograms. A correspondence map is then calculated using a scale-invariant feature transform (SIFT) flow algorithm to compute a transformation matrix applied to the model mask to transform it into the input image space. CTR valueis then calculated from the boundaries of lung and heart masks. Recent studies have shown that Deep learning concepts can be used for precise segmentation and CTR calculation.

For example, U-Net to extract lung and heart boundaries [5]. This approach obtained 93.75% accuracy on cardiomegaly detection task, on the dataset of 103 images from NIH Chest X-ray Dataset. Li et al. [7] used U-Net to segment heart and lung masks. They implemented a Conditional Random Field to the masks to smooth region boundaries and calculate CTR by measuring the cardio and thoracic diameters from the lung and heart masks. They performed the test on 5, 000 posteroanterior (PA) chest X-ray images from the Radiology Imaging Centre in their hospital. Furthermore, they obtained 95.3% accuracy on cardiomegaly detection.

U-Net [8] uses a deep learning model to tackle pixel- wise segmentation tasks accurately with great speed on various segmentation tasks. By cardioXNet [9], we know that CTR (Cardiothoracic Ratio) can be measured using Segmentation with U-net, Dense-Net.

The general CTR values to be considered parameter can be found in a paper published in GMJ Medical Journal[10], which gives the approximate CTR for the general, male, and female populations (1989 radiographs).

VIII. CONCLUSION

Many researchers have invested their time in image segmentation for bio-medical uses and have come up with various techniques from an early stage. However, the CTR calculation given by Isarun li et al. [22] Inspired us to study this topic further and create an automated system that can detect the omnipresence ofCardiomegaly in a patient.

REFERENCES

[1] Ahmed Dallal, Chirag Agarwal, Mohammad Arbabshirani, Aalpen Patel, and Gregory Moore. “Automatic estimation of heart boundaries and cardiothoracic ratio from chest x-ray images” page 101340K, 03 2017

[2] J Ebenezer and ACS Rao. “Computer aided analysis of chest x- ray images for early detection of cardiomegaly using euler numbers”, 2017.

[3] Sema Candemir, Stefan Jaeger, Wilson Lin, Zhiyun Xue, Sameer Antani, and George Thoma. “Automatic heart localization and radiographic index computation in chest x-rays”. In Georgia D. Tourassi and Samuel G. Armato III, editors, Medical Imaging 2016: Computer-Aided Diagnosis, volume 9785, pages 302 – 309. International Society for Optics and Photonics, SPIE, 2016.

[4] Nida M. Zaitoun and Musbah J. Aqel. “Survey on image segmentation techniques.” Procedia Computer Science, 65:797 – 806, 2015.

[5] Que, Z.Tang, R.Wang, Z.Zeng, J.Wang.M.Chua, T.S. Gee, X. Yang, and B. Veeravalli. Cardioxnet: "Automated detection for cardiomegaly based on deep learning”. pages 612–615, July 2018.

[6] “Automated Cardiothoracic Ratio Calculation And Cardiomegaly Detection Using Deep Learning Approach”: Isarun Chamveha1, Treethep Promwiset1,2, Trongtum Tongdee3, Pairash Saiviroonporn, and Warasinee Chaisangmongkon.

[7] O. Ronneberger, P. Fischer, and T. Brox “U-net: Convolutional networks for biomedical image segmentation”, in International conference on medical image computing and computer-assistedintervention, 2015, pp. 234–241.

[8] “CardioXNet: Automated Detection for Cardiomegaly Based on Deep Learning”.Qiwen Que1,4, Ze Tang1,4, Ruoshi Wang1,4, Zeng Zeng2,4,5, Jie Wang3,

Ma`hew Chua1, Teo Sin Gee2, Xulei Yang2, Bharadwaj Veeravalli3.

[9] "Establishing the Cardiothoracic Ratio Using Chest Radiographs in an Indigenous Ghanaian Population: A Simple Tool for Cardiomegaly Screening" : Y B Mensah,1 K Mensah,2 S Asiamah,1 H Gbadamosi,1 E A Idun,3 W Brakohiapa,1 and A Oddoye4.

[10] P. Moeskops, M. A. Viergever, A. M. Mendrik, L. S. de Vries, M. J. N. L. Benders, and I. Išgum, "Automatic segmentation of mr brain images with a convolutional neural network" IEEE Transactions on Medical Imaging, vol. 35, no. 5, pp. 1252–1261, May 2016.

[11] Mixture "Contrast Limited Adaptive Histogram Equalization for Underwater Image Enhancement", Muhammad Suzuri Hitam, Wan Nural Jawahir Hj Wan Yussof, Ezmahamrul Afreen Awalludin, Zainuddin Bachok.

[12] A Generalized "Unsharp Masking" Algorithm Guang Deng.

[13] Image Enhancement via "Adaptive Unsharp Masking" AndreaPolesel, Giovanni Ramponi, and V. John Mathews.

[14] "High Frequency Emphasis Filter" Instead Of Homomorphic Filter, Brian Nevils, Tory Mimbs , Abhimanyu Sailesh, Dr. Naima Naheed.

[15] "Fully Convolutional Networks for Semantic Segmentation", Evan Shelhamer, Jonathan Long, and Trevor Darrell, Member, IEEE,2016.

[16] "MICCAI BraTS 2017: Scope ,Section for Biomedical Image Analysis (SBIA), Perelman School of Medicine at the University of Pennsylvania". www.med.upenn.edu, 2018-12-24.

[17] "siliver07": www.sliver07.org, 2018-12-24.

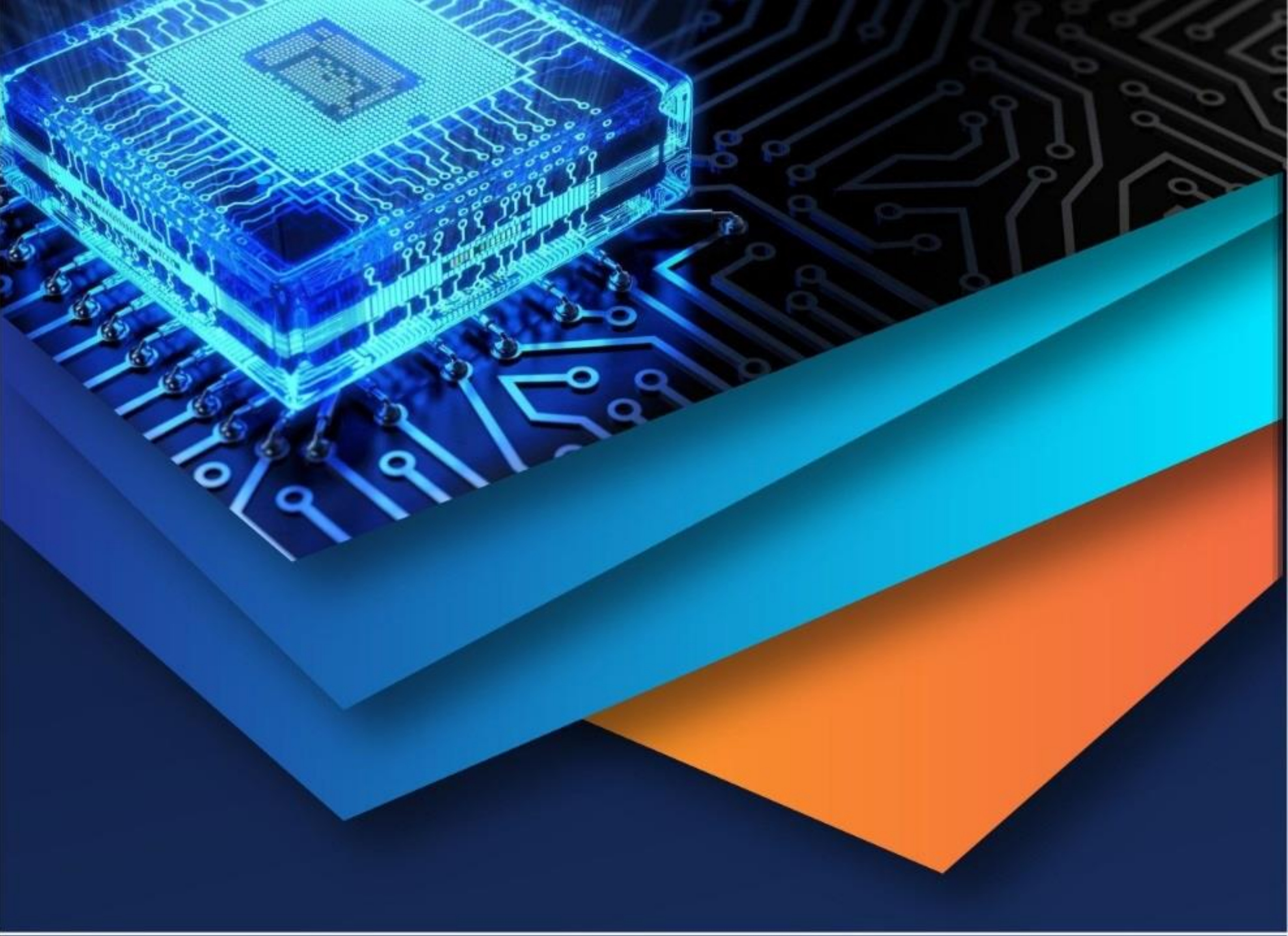
[18] "Separation of water and fat signal in whole-body gradient echo scans using convolutional neural networks",Jonathan Andersson, Håkan Ahlström, and Joel Kullberg.

[19] "A Deep Learning Approach for Network Intrusion Detection System", Quamar Niyaz, Weiqing Sun, Ahmad Y Javaid, and Mansoor Alam, 5th march 2015.

[20] *"**A deep learning approach to structured signal recovery",** Ali Mousavi; Ankit B. Patel; Richard G. Baraniuk, 2 oct. 2015.*

[21] ***"Feature Extraction: A Survey of the Types, Techniques,Applications"**, Ayodeji Olalekan Salau, Shruti Jain,2019.*

[22] ***"Automatic detection of lumen and media in the IVUS images using U-Net with VGG16 Encoder"**, Chirag Balakrishna, Sarshar Dadashzadeh, and Sara Soltaninejad.*



10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY