IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Medical Image Based Pulmonary Disease Detection Techniques and Features

Urvashi Dubey Amar Nayak

Technocrats Institute of Technology, Excellence Bhopal. M.P. India

***Abstract*—** Health facilities increase to detect and cure various diseases. Medical image diagnosis plays an important role to detect such infection or damages. Out of various kinds of diseases, pulmonary diseases are very common these days. Many researchers are working on this to detect and diagnose the image. This paper has surveyed various approaches adopted by authors to predict the infection class of disease. Paper has summarized work done by researchers to predict the input medical image class either infected or healthy. Further paper has identified the features used for the learning or prediction of pulmonary disease in the body. Comparison of various techniques were also done on the basis of outcomes and limitations. Finally the paper has listed few evaluation parameters to compare techniques of medical image diagnosis.

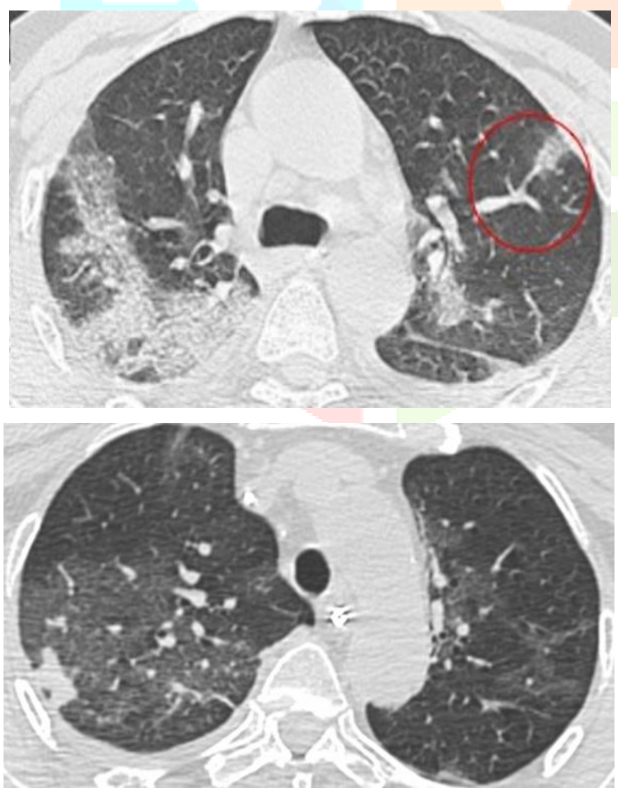
Index Terms— pulmonary disease, Image Processing, Information Extraction, visual processing.

I. INTRODUCTION

A large number of diseases that affect the worldwide population are lung-related. Therefore, research in the field of Pulmonology has great importance in public health studies and focuses mainly on asthma, bronchiectasis, COVID and Chronic Obstructive Pulmonary Disease. For the public health system, the early and correct diagnosis of any pulmonary disease is mandatory for timely treatment and prevents further death. From a clinical standpoint, diagnosis aid tools and systems are of great importance for the specialist and hence for the people's health. CT images of lungs represent a slice of the ribcage, where a large number of structures are located, such as blood vessels, arteries, respiratory vessels, pulmonary pleura and parenchyma, each with its own specific information. Thus, for pulmonary disease analysis and diagnosis, it is necessary to segment lung structures. It is worth noting that segmentation is an essential step in image systems for the accurate lung disease diagnosis, since it delimits lung structures in CT images. People all over the world are experiencing significant negative effects on their physical and mental health as a direct result of the rapid spread of the COVID-19 infection.

When imaging technology becomes more widespread, there is a corresponding increase in the demand for computing solutions that can interpret and analyse the ever more complex images that are produced. As a direct consequence of this, medical image computing has expanded in tandem with the expansion of medical imaging, and there are now a great deal of papers and conferences dedicated to the expansion of this field. The primary objective is to develop computational methods that can make use of the acquired imaging data to derive as much meaningful information as possible in order to improve the outcomes for patients [3]. As a consequence of this, machine learning (ML) algorithms have gained popularity across the entire spectrum of medical imaging, with applications ranging from segmenting regions of interest (segmentation), classifying whole images (classification), extracting characteristics from images (feature extraction), aligning multiple images (registration), and creating images from the raw data provided by the scanner. In short, ML algorithms have been used for everything from segmenting regions of interest to classifying whole images to extracting characteristics from images (reconstruction).

In order to improve the quality of medical care, researchers proposed various learning algorithms for diagnosis. In this study, different visual contents such as edges, histograms, DWT, and DCT, as well as others [3,] were used to extract image features that could then be categorized.



Infected Lungs

Non Infected Lungs

Fig. 1 Lung CT Images.

The learned behaviour of CNN, RNN, DNN, and other learning models was then applied to the extracted features [4]. On a single set of medical images, each model was evaluated to see how well it performed. As a consequence of this, the researchers need to develop a model that is capable of making accurate predictions regarding multiple diseases.

III. Features of Image Classification

Color Feature The image could be a matrix of light intensity values, and each of these intensity values would represent a different kind of colour [7]. The colour feature could be described as: Therefore, the ability to recognise an object's colour is a very important feature, and a low computation price is an essential component of this feature. There are many different image files available, each with its own unique colour format. For example, images can have any number of colour formats, ranging from RGB, which stands for red, green, and blue.

Edge Feature: As an image can be a collection of intensity values, and with the rapid change in the values of a picture, one important feature emerges as the Edge, as shown in figure 4. Edge Feature: As an image can be a collection of intensity values. This feature can be utilised for the detection of a wide variety of image objects, including roads, buildings, and other similar elements [5, 7]. There are several rules that have been developed to effectively illustrate all of the pictures of the image or frames, such as Sobel, perwitt, and canny, among others. Out of all of those algorithms, canny edge detection is one of the most effective algorithms to search out all of the potential boundaries of a picture.

Texture is a property that enumerates qualities such as regularity and smoothness [6]. Texture can be thought of as a degree of distinction in the intensity of a surface's appearance. When compared to the paint house model, the texture model requires an additional step in the process. The feel options based on the colour premise are less sensitive to changes in illumination than the same on edge options.

Corner Feature: In the event that the camera is moving, it is necessary to be able to differentiate between the two frames that are being displayed within the image or frame thanks to the corner feature. This allows the video to remain stable. Therefore, resizing the window in the original text can be accomplished by locating the corner position of the two frames and using that information. This function can also be used to determine the angles still as the distance between the item in the two distinct frames. because they serve a purpose within the image and can therefore be used to trace the objects that are the focus of attention.

DWT Feature: It is a frequency domain feature that is used to transform pixel values in frequency domain having four region first is flat region, other is horizontal edge region, similarly vertical and diagonal edge region [8,9]. [DWT Feature] A combination of low pass and high pass filters was used to obtain each image subsection.

DCT Feature: This is another feature that falls under the frequency domain umbrella. In the top left corner of the image matrix, low frequency values were found to be present. In order to obtain these feature set coefficients, the cosine transformation operation was applied. The Discrete Cosine Transmit, also known as DCT, is an image processing technique that is considered to be both an industry standard and one of the most widely used today. The DCT makes it possible to separate an image into a number of distinct frequency bands, which makes it much simpler to secretly embed data hiding information into the middle frequency bands of an image. Due to the fact that this part of the spectrum is unaffected by either noise or compression, the middle frequency region is utilised here for the purpose of data hiding. The fact that the visual frequency region can be found in the low frequency part of the image is yet another consideration. Therefore, embedding is accomplished by positioning the least significant bit (LSB) of the pixel.

III. Related Work

In [8] authors address this issue and propose a computational framework to determine the risk factors for postoperative PVO (PPVO) from computed tomography angiography (CTA) images and build the PPVO risk prediction model. From clinical experiences, such risk factors are likely from the left atrium (LA) and pulmonary vein (PV) of the patient. Thus, 3D models of LA and PV are first reconstructed from low-dose CTA images. Then, a feature pool is built by computing different morphological features from 3D models of LA and PV, and the coupling spatial features of LA and PV. Finally, four risk factors are identified from the feature pool using the machine learning techniques, followed by a risk prediction model.

In [9] authors proposes a detection method of pulmonary embolism based on the improved faster region-based convolutional neural network (Faster R-CNN) named More Accurate Faster R-CNN (MA Faster R-CNN). A new feature fusion network named Multi-scale Fusion Feature Pyramid Network (MF-FPN) is proposed by extending and adding two bottom-up paths on the Feature Pyramid Network (FPN). It enhances the feature extraction capability of the entire network by transmitting low-level accurate location information, and makes up for the original information lost after multiple down-sampling, strengthens the use of detailed information, which is more helpful to the detection of small object. In the prediction module, the residual block is added before the fully-connected layer to deepen the network and enhance the classification accuracy, named residual prediction module (RPM).

In [10] authors proposed on the identification and diagnosis of Chronic obstructive pulmonary disease (COPD) in the high-risk population, this study proposes a convenient and effective clinical decision support system to help identify patients with COPD in primary health institutions.

The techniques of multilevel feature extraction and concatenation were utilised by N. Noreen et al. in the year 2020 [11] in order to detect early tumour diagnosis. This project utilises the two different models known as Inception V-3 and DensNet201 in order to establish the two different alternative methods for the identification and diagnosis of cancers. The tumour detection characteristics of the inception model were retrieved from the pre-trained inception model V3 and concatenated at the beginning of the process. After that, the SoftMax classifier was used to determine what kind of brain tumour the patient had. In a similar manner, the DensNet201 was utilised to extract features from the Dens Netblock, which were subsequently concatenated and sent through the SoftMax classifier in order to locate the tumour. As a consequence of this, three datasets on class tumours that are available to the public are used to check both modalities.

In the year 2020, Hari Mohan Rai and colleagues will construct a deep neural network for the diagnosis of cancer that has fewer layers and less complexity in its architecture than the U-Net [12]. A data set containing 253 images needed to be sorted into categories of normal and abnormal MRI scans. There was a reason for doing so. Before this, MRI scans were resized, cropped, pre-processed, and supplemented in order to guarantee an accurate outcome and facilitate deep neural network training as quickly as possible.

Malaria is a dangerous and common disease, and Mehedi Masud et al. [13] presented a new algorithm that can detect the presence of the disease. This algorithm was developed specifically as a mobile healthcare option for patients. Convolutional or deep learning architecture, which has been shown to be beneficial in detecting malaria disease in real-time by imputing images and thereby reducing the amount of manual labour required in disease detection, will be the primary focus of this paper. The paper's primary objective is to focus on convolutional or deep learning architecture.

In the paper that Khan et al. [14] wrote, the author proposed a model for COVID-19 diagnosis called "CoroNet." This model is a CNN model that uses radiography images of the chest. The suggested method is based on something called "Xception Architecture," which is a model that has already been pre-trained using ImageNet's dataset. After that, the model is trained using a dataset that was compiled for research purposes from a variety of databases that are available to the general public.

In [15] Only images are used in the reduction of dimension that is proposed by this paper's algorithm, and each image's unique characteristics are preserved to the greatest extent possible. During the process of classifying the objects, the dimensions surrounding their appearance were shrunk, and an edge detection technique was performed in advance to obtain information about the object's appearance. Extracting and crossing three segment images that are adjacent to each other on a window of a certain size results in the production of three distinct images through the process of edge detection. After performing an AND operation on the images that were produced, binary edge information is extracted by eliminating the noise that was caused by the AND operation that was not removed by the difference image. Such produced data is used as an input of the RNN-based learning model, and the reduction data of each segment in a circulation layer is also inputted to detect anomalies in the entire X-ray image. This is done in order to determine whether the X-ray image of COVID-19 is positive or negative.

Year	Technique	Advantage	Limitation
Goyal, S. et. al. [16]	Recurrent Neural Network and LSTM	Efficiently classify X-Ray images with 95% accuracy.	Feature optimization need to be improved.
Chandra Mani Sharma et. al. [17]	synthetic minority over-sampling technique (SMOTE)	98.6 % accuracy of classification of CXR images.	Addition of synthetic data may leads to patient diagnosis.
J. -X. Wu et. al. [18]	Fractional order convolution (FOC)	This work was done on random images and achieved 83.57% accuracy.	Convolutional operations for feature extraction increases execution time.
Y. Choi et. al. [19]	convolutional neural network-bidirectional gated recurrent	92.3 accuracy was achieved.	Convolutional operations for feature extraction increases execution time.
Y. Wang et. al. [20]	Support Vector Machine	Auc of this model for Pleural Line and B-Lines in Lung Ultrasound Images is 0.96.	SVM is not good for multi class.

IV. Techniques of Image Classification

The Image Classification methods are [21, 22]:

1) **Support Vector Machine :** This technique creates a set by using vectors. Use of hyperplanes in a high-dimensional space, which is used to illustrate the classification or statistical regression purposes [20, 24]. The healthy distance between them realized by the use of the hyperplane. SVM uses non-parametric with an approach based on binary classification and capable of handling more input data efficiently. The hyperplane has an effect on accuracy for image classification [24].

2) **Artificial Neural Network (ANN):** Artificial neural networks Artificial neural networks is a form of artificial intelligence that sends out certain signals. Functions that the human mind is capable of an ANN is having a a layered progression or sequence [25]. Every level of the neural network structure comprised of a group of nerve cells called neurons. Neurons can be found in every layer linked by weighted connections to all of the neuronal layers prior to and following the current one. The precision is determined by on the amount of inputs that are available and the layout of the network [26].

3) **Decision Tree:** A decision tree is a graph that is structured like a tree decisions. Each fork reflects the decisions that need to be made created using graphical means [27]. It is a supervised non-parametric analysis approach. It does so by dividing the input into standard classes [28, 29]. The acceptance or rejection of class is made possible through method label at each subsequent stage in the process. The set can be obtained through this method a set of rules that come after classification that need to be comprehended.

4) **Maximum likelihood classifier:** Maximum Likelihood is a type of image classification that requires human oversight method that involves calculating the probability value of individual pixels into account for the purpose of classifying the pixels [30]. Within this method, the probability of every pixel belonging to a certain class can be calculated as calculated. After that, we compare these values to one another. The pixel is this: attributed to the category that possesses the highest probability value. For the purposes of this method, it is presumed that each of the input bands normal distribution.

5) Minimum distance classifier: A supervised image is what the minimum distance classifier is a technique for classifying images in which the pixels are categorized according to how far away their spectra are from the average of the pre-defined categories [27]. This technique begins by calculating the mean vector for the training dataset is used to inform the calculation of each class. Next, utilising the algorithm for finding the shortest distance, the Euclidean the distance of each pixel that has not been classified from the mean vector is calculated. After that, the pixel is assigned to the class that contains the minimum distance. Calculating the distance between a specific pixel and a number of different class mean vectors typically involves using the Euclidean distance: This particular classification method is not overly complicated mathematically, and as a result, it also requires less effort to implement in practice. When compared to all of the other techniques for supervised classification, it has the quickest computation time requirements. The fact that this method only considers the average value makes it less effective than other approaches; as a result, it has fewer advantages technique that is more effective than maximum likelihood.

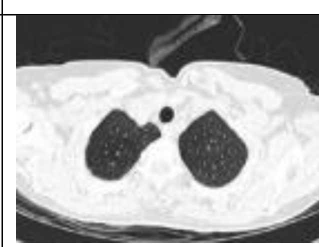
6) Random Forest (RF) : RF is a type of supervised machine learning algorithm [31] that primarily functions to solve problems by classifying data in order to make predictions. This algorithm combines several different Decision Trees, and the more trees it uses, the more precise the results will be. These decision trees are then trained to produce outputs (predictions), and the Random Forest algorithm will choose the best prediction (solution) based on the voting. The decision trees are fed data and trained to produce outputs. A tree of possible choices. The selection of data samples at random from a dataset is the first step in the Random Forest algorithm [32, 33]. After that, a decision tree is constructed for each selected sample, and the results of the predictions made using each tree are compiled. After all of the information has been compiled, a voting procedure is carried out in order to select the most accurate prediction result as the ultimate answer.

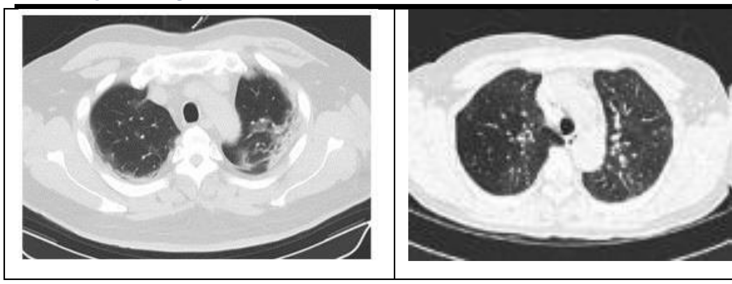
7) Nearest neighbour-based methods : This method can determine whether a given data point is irregular by analysing its proximity to other points, also known as neighbours, or by focusing on the density [34]. The k-Nearest Neighbors algorithm is one of these methods (kNN). The anomaly score, which is the distance between the neighbours, is what this method uses to calculate (data instances). If the score is higher than a certain score level, then it is considered to be anomalous. Checking to see if the density around a data point is low or high can help determine whether or not a point should be considered anomalous based on density. The data point is considered abnormal if the density is low, whereas it is considered normal if the density is high. For instance, an algorithm known as Local Outlier Factor (LOF) is a method that is used in conjunction with k-nearest neighbours to compute the average density ratio of the points and anomaly score in order to determine whether or not there are any abnormal behaviours. This allows the algorithm to determine whether or not there are any outliers in the data. When it comes to determining the density of information in large datasets, LOF performs significantly better than kNN. However, when applied to large datasets, both of these methods have problems with their scalability. The reason for this is that in order to determine nearest-neighbors, both of these methods need to compute the distance between the data points. In addition, the use of these two methods raises the bar for the level of computational complexity required in both the training and the testing phases.

V. Evaluation Parameters

In order to evaluate results there are many parameter such as accuracy, precision, recall, F-score, etc. Obtaining values can be put in the mention parameter formula to get results [34, 35, 36].

Table 2 Sample dataset images.

Infected Images	Healthy Images
	
	
	



Precision =(True_Positive)/(True_Positive+ False_Positive)

Recall =(True_Positive)/(True_Positive+ False_Negative)

F-Measure =(2* Precision*Recall)/(Precision + Recall)

Accuracy =(True_Positive+ True_Negative)/(True_Positive+ True_Negative + False_Positive +False_Negative)

VI. Conclusion

Detection of any heath issue at early stage is good for every living being. Hence medical image diagnosis plays an important role for the early detection. This paper has cover the pulmonary disease detection models proposed by various researcher with the techniques and features adopted. Paper has elaborate various image features used for the image classification in different applications. As classification of input data needs learning, so paper has list techniques of image feature learning and prediction models adopted for pulmonary diseases. In future scholars can apply feature leaning models in various lungs infection diseases like corona, Asthma new variants.

References

1. L. Li, “Using artificial intelligence to detect COVID-19 and community-acquired pneumonia based on pulmonary CT: evaluation of the diagnostic accuracy,” vol. 296, no. 2, 2020, pp 250-262.
2. Rehman, S. Naz, A. Khan, A. Zaib, and I. Razzak, Improving Coronavirus (Covid-19) Diagnosis Using Deep Transfer Learning, MedRxiv, Cold Spring, NY, USA, 2020.
3. Alexander Zotina, Konstantin Simonovb, Mikhail Kurakoc, Yousif Hamadc, Svetlana Kirillova. "Edge detection in MRI brain tumor images based on fuzzy C-means clustering" International Conference on Knowledge-Based and Intelligent Information & Engineering Systems, 2018, pp 1261-1270.
4. M. Li, L. Kuang, S. Xu and Z. Sha, "Brain Tumor Detection Based on Multimodal Information Fusion and Convolutional Neural Network," in IEEE Access, vol. 7,.
5. Y. Pei *et al.*, "Building a Risk Prediction Model for Postoperative Pulmonary Vein Obstruction via Quantitative Analysis of CTA Images," in *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 7, pp. 3127-3138, July 2022, doi: 10.1109/JBHI.2022.3146590.
6. H. Yuan, Y. Shao, Z. Liu and H. Wang, "An Improved Faster R-CNN for Pulmonary Embolism Detection From CTPA Images," in *IEEE Access*, vol. 9, pp. 105382-105392, 2021, doi: 10.1109/ACCESS.2021.3099479.
7. X. Lin *et al.*, "A Case-Finding Clinical Decision Support System to Identify Subjects with Chronic Obstructive Pulmonary Disease Based on Public Health Data," in *Tsinghua Science and Technology*, vol. 28, no. 3, pp. 525-540, June 2023, doi: 10.26599/TST.2022.9010010.
8. Das, A.K., Ghosh, S., Thunder, S. et al. Automatic COVID-19 detection from X-ray images using ensemble learning with convolutional neural network. Pattern Anal Applic 24, 1111–1124 (2021).
9. Hari Mohan Rai, Kalyan Chatterjee. "Detection of brain abnormality by a novel Lu-Net deep neural CNN model from MR images". Machine Learning with Applications 7 October 2020.
10. Mehedi Masud, Hesham Alhumyani, Sultan S. Alshamrani, Omar Cheikhrouhou, Saleh Ibrahim, Ghulam Muhammad, M. Shamim Hossain, Mohammad Shorfuzzaman, "Leveraging Deep Learning Techniques for Malaria Parasite Detection Using Mobile Application", Wireless Communications and Mobile Computing, 2020.
11. Hong Huang, Fanzhi Meng, Shaohua Zhou, Feng Jiang, And Gunasekaran Manogaran. "Brain Image Segmentation Based On Fcm Clustering Algorithm And

Rough Set". IEEE Access Special Section On New Trends In Brain Signal Processing And Analysis, February 6, 2019, pp 12386-12396.

12. N. Noreen, S. Palaniappan, A. Qayyum, I. Ahmad, M. Imran and M. Shoaib, "A Deep Learning Model Based on Concatenation Approach for the Diagnosis of Brain Tumor," in IEEE Access, vol. 8, pp. 55135-55144, 2020.
13. Hari Mohan Rai, Kalyan Chatterjee. "Detection of brain abnormality by a novel Lu-Net deep neural CNN model from MR images". Machine Learning with Applications 7 October 2020, 36111-36114.
14. Mehedi Masud, Hesham Alhumyani, Sultan S. Alshamrani, Omar Cheikhrouhou, Saleh Ibrahim, Ghulam Muhammad, M. Shamim Hossain, Mohammad Shorfuzzaman, "Leveraging Deep Learning Techniques for Malaria Parasite Detection Using Mobile Application", Wireless Communications and Mobile Computing, 2020, pp 68-70.
15. Q. Yan et al., "COVID-19 Chest CT Image Segmentation Network by Multi-Scale Fusion and Enhancement Operations," in IEEE Transactions on Big Data, vol. 7, no. 1, pp. 13-24, 1 March 2021.
16. Goyal, S., Singh, R. Detection and classification of lung diseases for pneumonia and Covid-19 using machine and deep learning techniques. J Ambient Intell Human Comput (2021).
17. Chandra Mani Sharma, Lakshay Goyal, Vijayaraghavan M. Chariar, Navel Sharma, "Lung Disease Classification in CXR Images Using Hybrid Inception-ResNet-v2 Model and Edge Computing", Journal of Healthcare Engineering, vol. 2022, Article ID 9036457, 2022, 1-15.
18. J. -X. Wu, P. -Y. Chen, C. -M. Li, Y. -C. Kuo, N. -S. Pai and C. -H. Lin, "Multilayer Fractional-Order Machine Vision Classifier for Rapid Typical Lung Diseases Screening on Digital Chest X-Ray Images," in IEEE Access, vol. 8, pp. 105886-105902, 2020.
19. Y. Choi, H. Choi, H. Lee, S. Lee and H. Lee, "Lightweight Skip Connections With Efficient Feature Stacking for Respiratory Sound Classification," in IEEE Access, vol. 10, pp. 53027-53042, 2022.
20. Y. Wang, Y. Zhang, Q. He, H. Liao and J. Luo, "Quantitative Analysis of Pleural Line and B-Lines in Lung Ultrasound Images for Severity Assessment of COVID-19 Pneumonia," in IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, vol. 69, no. 1, pp. 73-83, Jan. 2022
21. M. Liang, Z. Li, T. Chen and J. Zeng, "Integrative data analysis of multi-platform cancer data with a multimodal deep learning approach", IEEE/ACM Trans. Comput. Biol. Bioinf., vol. 12, no. 4, pp. 928-937, Jul. 2015.
22. Sunitha Abburu and Suresh Babu Golla, “Satellite Image Classification Methods and Techniques: A Review”, International Journal of Computer Applications (0975 – 8887) Volume 119 – No.8, June 2015
23. Q. Kai, "A Painting Image Retrieval Approach Based on Visual Features and Semantic Classification," 2019 International Conference on Smart Grid and Electrical Automation (ICSGEA), Xiangtan, China, 2019, pp. 195-198.
24. Parveen and A. Singh, "Detection of brain tumor in MRI images, using combination of fuzzy c-means and SVM," 2015 2nd International Conference on Signal Processing and Integrated Networks (SPIN), Noida, India, 2015, pp. 98-102.
25. M. Abdar et al., "Hercules: Deep Hierarchical Attentive Multilevel Fusion Model With Uncertainty Quantification for Medical Image Classification," in IEEE Transactions on Industrial Informatics, vol. 19, no. 1, pp. 274-285, Jan. 2023.
26. Y. Yu, E. Favour and P. Mazumder, "Convolutional Neural Network Design for Breast Cancer Medical Image Classification," 2020 IEEE 20th International Conference on Communication Technology (ICCT), Nanning, China, 2020, pp. 1325-1332.

27. F. S. Al-Ahmadi and A. S. Hames, “Comparison of Four Classification Methods to Extract Land Use and Land Cover from Raw Satellite Images for Some Remote Arid Areas, Kingdom of Saudi Arabia”, JKAU; Earth Sci., Vol. 20 No.1, pp: 167-191 (2009 A.D./1430 A.H.)

28. O. Oktay et al., "Stratified Decision Forests for Accurate Anatomical Landmark Localization in Cardiac Images," in IEEE Transactions on Medical Imaging, vol. 36, no. 1, pp. 332-342, Jan. 2017.

29. W. Yu and Y. Xiaowei, "Application of decision tree for MRI images of premature brain injury classification," 2016 11th International Conference on Computer Science & Education (ICCSE), Nagoya, Japan, 2016, pp. 792-795.

30. Nabila Farnaaz, M.A. Jabbar, Random Forest Modeling for Network Intrusion Detection System, Procedia Computer Science, Volume 89, 2016, Pages 213-217.

31. J. -H. Lee, D. -Y. Kim, B. C. Ko and J. -Y. Nam, "Keyword Annotation of Medical Image with Random Forest Classifier and Confidence Assigning," 2011 Eighth International Conference Computer Graphics, Imaging and Visualization, Singapore, 2011, pp. 156-159.

32. P. Wicaksono and W. Lazuardi, "Random Forest Classification Scenarios for Benthic Habitat Mapping using Planetscope Image," IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 2019, pp. 8245-8248.

33. R. Zhang et al., "A Comparison of Gaofen-2 and Sentinel-2 Imagery for Mapping Mangrove Forests Using Object-Oriented Analysis and Random Forest," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 14, pp. 4185-4193, 2021.

34. Nabila Farnaaz, M.A. Jabbar, Random Forest Modeling for Network Intrusion Detection System, Procedia Computer Science, Volume 89, 2016, Pages 213-217.

35. M. Amer and M. Goldstein, "Nearest-Neighbor and Clustering based Anomaly Detection Algorithms forRapidMiner," Department of Computer Science and Engineering, German University in Cairo. International Conference, pp 315.

36. C. -M. Kim, E. J. Hong and R. C. Park, "Chest X-Ray Outlier Detection Model Using Dimension Reduction and Edge Detection," in IEEE Access, vol. 9, pp. 86096-86106, 2021,