iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** XI **Month of publication:** November 2023

DOI: https://doi.org/10.22214/ijraset.2023.56650

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue XI Nov 2023- Available at www.ijraset.com

Mediscan: An Multidisease Identification and Prognosis System using AI

Rohan Kadoo[1], Mohammad Tahzeeb Khan[2], Mohammad Hassan Raza[3], Tanishq Sakhare[4], Gaurav Dwivedi[5], Prasanna Lohe[6], Mayur Kadu[7]
Computer Science & Engineering, S.B. Jain Institute of Technology Management and Research, Nagpur, India

Abstract: ***Chronic diseases such as cancer, diabetes, strokes, arthritis, and cardiac-related disease are the major and leading cause of high mortality and disability rates in India as well as worldwide. Developing a convincing and favorable solution for these diseases is the need of the hour. The development and Technological advancements in medical science have proved beneficial in detecting the initial stage among patients and providing accurate data analysis among them. The authenticity and accuracy of the diagnosis and consequent treatment depend upon the correct analysis of patients; incorrect diagnosis or overdiagnosis may lead to casualty. Therefore, at most care and precautions must be taken for the correct examination of illness or disease as misdiagnosis may result in the death of a patient. However, a Machine learning, Deep learning-based diagnostic system with high accuracy proposed here, can offer promising solutions to identify the correct & accurate cause of such chronic diseases. Machine Learning-based diagnostic systems can detect diseases such as Lung disease, brain tumors, Heart disease, Skin disease, diabetes, and prophecy of developmental stages in patients. A proper or suitable diagnostic system may prove helpful for doctors in reducing the high mortality rate among patients with these chronic diseases. Great work has to be done on the accurate diagnosis of many diseases. A lot of work has been done in this direction but no convincing solution for accurate diagnosis has been found till now with the help of machine learning Deep learning diagnostics systems we can identify the diseases as well as the developmental stages in many diseases such as Lung disease, Brain tumor, Heart disease, etc. In this way, we can reduce the mortality rate and save money lives. In these Research paper, we have explored various Machine learning and Deep learning algorithms for training and testing the different diseases in our system. With the help of Machine learning and deep learning techniques, we trained our various machine learning and deep learning models using various algorithms for each disease. In our Multidisease Identification and Prognosis system, we trained the model for five diseases i.e. Heart disease, Brain tumor, Skin disease, Lung disease, and diabetes. So, we have achieved 98% accuracy on Heart disease, 97% accuracy on Brain disease, 90% accuracy on Skin disease, 87% accuracy on Lung-related diseases, and 97% accuracy on diabetes with the help of different machine learning and deep learning algorithms in our Project We used different algorithms for trained the model like VGG 16, Dense Net, Res Net 50, Random Forest, Sequential to trained our multiple diseases model. In the end, we created the whole web application for easy and understandable user interaction and to fulfill the requirements of patients.***

Keywords: Machine learning, Deep Learning, Training, Prognosis, Multiple diseases, over diagnosis, misdiagnosis, analysis, Mediscan.

I. INTRODUCTION

Diseases related to the heart, lungs, and brain are the main cause of death or serious medical issues. These diseases are now common for all age groups. Some more diseases related to the skin are also common. Diabetes is one of the most common and lifelong diseases which does not have any cure. So all 4 human body parts (Heart, Lungs, Skin, and Brain) are the major parts on which other body parts are dependent. As we know nowadays, Heart stroke and heart Attack are the most common. This heart-related disease cannot be easily handled and cured. And these diseases don't have any pre-symptoms available. Lung disease can be cured but not 100%, and Brain disease as well. Skin Disease can be cured but at a certain time. Skin cancer is the major issue that infects a person who has a high cell level. Our System 'MEDISCAN' which detects and predicts the Disease and prognoses the user based on previous medications. Artificial Intelligence is one of the greatest and most Innovative developments in the field of Engineering. A.I can learn anything very easily and faster than Humans. A.I can detect, predict, Observe, Proof, and conclude anything on which it trained for.

A. Lungs Disease

The Disease that generally causes damage, inflammation, or obstruction to the lungs is known as lung disease. This type of Disease can be a reason for Death. There are many diseases that are related to the functioning of the lungs such as

1) Asthma
2) Chronic obstructive pulmonary disease
3) Pneumonia
4) Tuberculosis,etc.

B. *Brain Disease*

The brain is one of the most important and crucial body parts, which is responsible for each activity of the body. Diseases related to the Brain can cause abnormal activity, serious issues, or death. There are many diseases related to the Brain such as

1) Tumor
2) Neurodegenerative
3) Infection
4) Stroke, etc

C. *Skin Disease*

Skin is the outermost layer of the body. Skin is the main layer of the body that protect human being and other living creature from every harm from the environment. What if our skin is damaged or infected from any disease? There are many diseases associated with skin such as

1) Cancer
2) acne
3) eczema
4) Rosacea

II. LITERATURE SURVEY

1) *Heart Disease Using Retinal Image*

M.Rupadevi Et Al. (2022) Discusses The Process Of Identifying Heart Disease Using Retinal Images, Particularly In Children. It Outlines The Steps Involved, Such As Image Preprocessing, Feature Extraction, Classification Using Support Vector Machine And Random Forest Classifiers, And The Identification Of The Disease Based On The Extracted Features. The Proposed Work Aims To Develop A Heart Disease Prediction System Using Retinal Images From The Chase Dataset.

2) *Skin Diseases Using Machine Learning*

Sunpreet Bhatiya Et Al.(2023) Described A Study On Skin Disease Detection Using A Cnn Algorithm. It Presents The System Architecture, Methodology, And Results Of The Study. The Proposed System Uses Color Image Processing Techniques And Artificial Neural Networks For Disease Detection. The Study Evaluated The Model On A Dataset Of 800 Images And Conducted Sensitivity Analysis To Test Its Robustness.

3) *X-Ray Disease Identifier*

Rohan Darji Et Al.(2023) Discusses The Development Of A Deep Cnn Architecture For The Prediction Of Lung Diseases From Chest X-Ray Images. It Includes Details About The Confusion Matrix, Sample Outputs, Problem Statement, And The Use Of Digital Pathology And Image Analysis In Clinical Trials. Additionally, It Presents A Customized Vgg19 Architecture For Pneumonia Detection In Chest X-Rays, Along With Pre- Processing Techniques For Image Normalization And Resizing.

4) *Multiple Disease Prediction System*

Tanmay Ture Et Al.(2023) Described The Process Of Building Machine Learning Models To Predict Diseases In Healthcare Systems. It Emphasizes

The Importance Of Data Quality And Quantity, Data Preprocessing, And Model Selection. The Random Forest Algorithm Is Chosen For Its Accuracy And Ease Of Implementation. The Models Are Deployed In Flask Framework Using The Pickle Module.

© IJRASET: All Rights are Reserved |

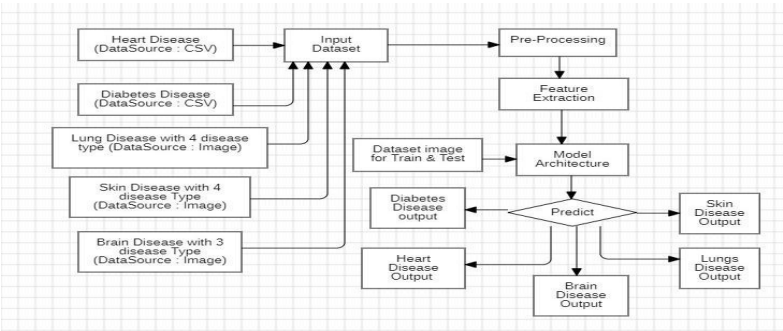
The Xgboost Algorithm Is Also Briefly Mentioned. The Models Are Trained And Tested On Datasets From Kaggle, Achieving High Testing Accuracy.

5) Multi Disease Prediction System

Divya Mandem Et Al.(2021) Discusses A System For Predicting Diseases Based On Symptoms Using A Combination Of Structured And Unstructured Data. It Involves Data Collection From Online Sources, Data Preprocessing, Building A Prediction Model Using Machine Learning Techniques Like Random Forest, And Predicting Diseases With A High Accuracy Probability Of 95%.

III. PROPORSED METHODOLOGY

This section elucidates the datasets used, the preprocessing, data augmentation methods, and the diverse algorithms applied. The proposed technique's workflow is illustrated in Figure 1.



System Architecture

A. Datasets

This project curates a varied dataset from Kaggle and GitHub, encompassing medical conditions. The lung disease dataset comprises Tuberculosis, Bacterial Pneumonia, Viral Pneumonia, and Normal cases, with 4856 training and 1621 testing X-ray images. The brain tumor dataset includes glioma, meningioma, and no tumor categories, totaling 6692 training and 1674 testing images. The skin dataset comprises images of fungal infections, acne, vascular tumors, and normal skin, with 2862 training and 865 testing images. Additionally, two tabular datasets exist: the heart dataset, a CSV file (918,12), predicts heart disease, and the diabetes dataset, another CSV file (100000,9), identifies diabetes cases. This amalgamation explores a holistic perspective on medical diagnostics, incorporating both imaging and clinical data.

1) Preprocessing and Data Augmentation.

Images in the datasets have varying resolutions. Images from the lung, brain, and skin datasets were resized to a uniform 224 x 224 dimensions. Standardization enhances model performance by ensuring consistent input dimensions, promoting computational efficiency, and avoiding biases from differing image sizes. Data augmentation techniques, including rotation, flipping, and zooming, diversified the training dataset, improving model generalization. These practices mitigate overfitting, fostering a robust and adaptable model.

In CSV files, preprocessing involves handling missing values, scaling numerical features, and encoding categorical variables. Imputation methods such as mean or median address missing values, scaling ensures uniform influence, and categorical encoding facilitates meaningful analysis.

2) Feature Extraction

Feature extraction is critical for both image datasets and CSV files in deep learning and machine learning. Convolutional neural networks (CNNs) play a crucial role in automatically extracting hierarchical features from images. In CSV files, feature extraction involves selecting pertinent features contributing significantly to the predictive task, using techniques like information gain, correlation analysis, and recursive feature elimination.

B. Deep Learning and Machine Learning Algorithms

© IJRASET: All Rights are Reserved |

This paper explores the implementation of CNN models, Sequential models and Pretrained Model, enriched through the integration of CNN with data augmentation techniques. Three distinct model algorithms are detailed in subsequent subsections for a comprehensive understanding of their architecture and functionality.

1) Sequential Model

The sequential model, exclusive to the lungs model, comprises four convolutional layers with increasing filters. ReLU allows gradients to pass through, and max pooling follows each activation. Adam optimizer and a learning rate of 0.0001 were employed.

2) Pretrained Model

This model, widely acknowledged as the most straightforward and commonly used for image classification, operates on a principle distinct from training a model from scratch. Instead, it leverages pre-existing weights obtained from a large dataset to classify the images at hand. Commonly referred to as transfer learning, this technique capitalizes on previously learned weights for the classification task, resulting in reduced training time and enhanced accuracy. In this project, a pretrained DenseNet model is applied to the lungs dataset, utilizing its well- established accuracy. The skin dataset is processed through VGG-16, a robust convolutional neural network (CNN) known for its precision. For the brain model, ResNet50 is employed, harnessing its proficiency in capturing intricate features. The achieved accuracies, with the lungs model at 89.5%, the skin model at 89.3%, and the brain model boasting an impressive 97.73%, underscore the efficacy of these deliberate choices in pretrained models, significantly contributing to the project's success.

3) Random Forest

In analyzing the heart and diabetes datasets, the Random Forest algorithm demonstrated notable results. For the heart dataset, an accuracy of 89.1% was achieved, showcasing efficacy in predicting heart disease. Similarly, the diabetes dataset exhibited remarkable performance, with an accuracy of 97.03%, affirming the algorithm's suitability for predictive modeling in medical diagnostics.

IV. RESULT

So, In Our Project i.e. Multidisease Identification and Prognosis system we trained different machine Learning and deep learning algorithms for training and testing of our machine Learning and Deep learning models, so there are five types of diseases in our Project i.e. Heart disease we trained our Heart disease model with the use of Random Forest algorithm which gives the accuracy of 87% In diabetes disease model again we use Random Forest algorithm because it gives the highest accuracy than other algorithm, and it gives the accuracy of 97% In Skin disease model we used VGG 16 algorithm, and it gives 90% accuracy In Brain disease we use Res net 50, and it gives 97% accuracy than at the end In Lungs disease model we use two different algorithms i.e. Sequential and Dense net the sequential gives 82% accuracy dense net gives 88% accuracy As a result in our Project three algorithms gives the best accuracy for trained our model i.e. VGG 16, Random Forest and Res net 50 and gives the accurate result quickly.

Table -1: Accuracy for Heart Disease

ALGORITHM	Heart Accuracy
Random Forest	87%

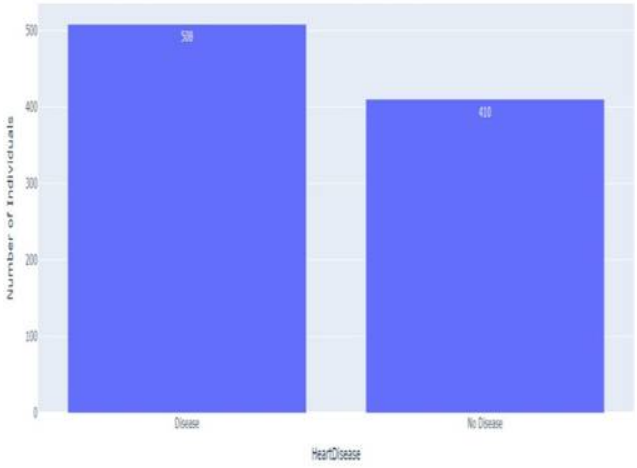


Table -2: Accuracy for Diabetes Disease

ALGORITHM	Diabetes Accuracy
Random Forest	97%

Table -3: Accuracy for Skin Disease

ALGORITHM	Skin Accuracy
VGG16	90%

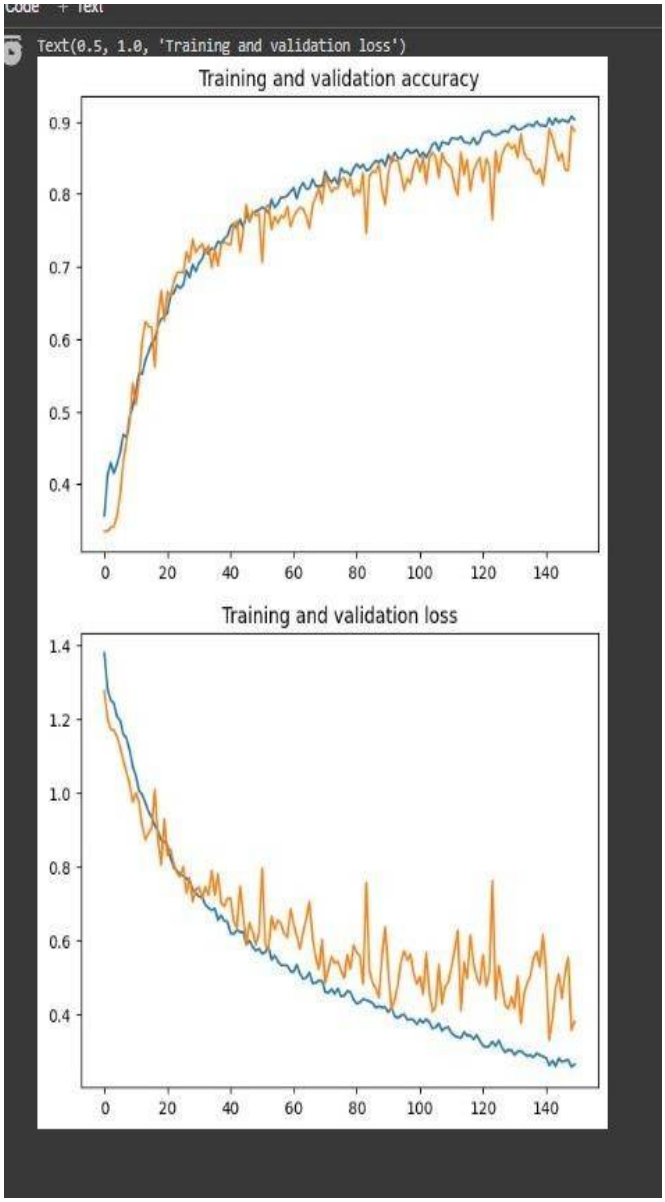


Table -4: Accuracy for Brain Disease

ALGORITHM	Brain Accuracy
Resnet 50	97%

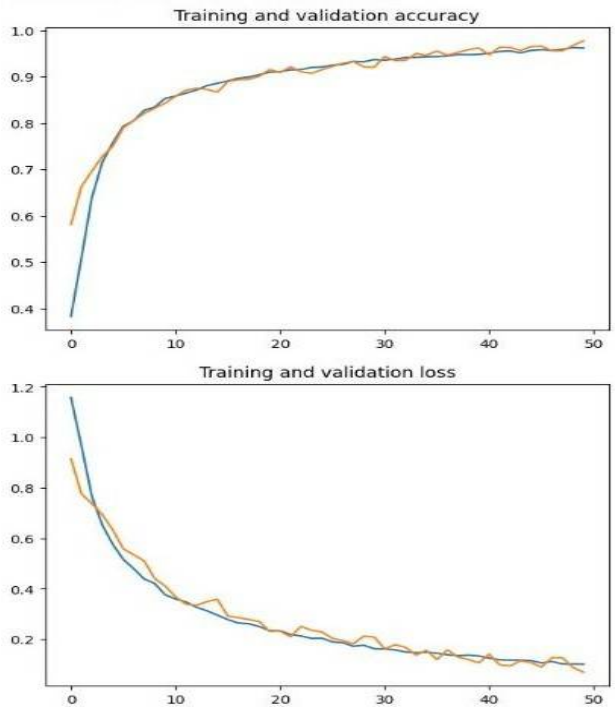


Table -5: Accuracy for Lung Disease

ALGORITHM	Lung Accuracy
Sequential	82%
Densenet	88%

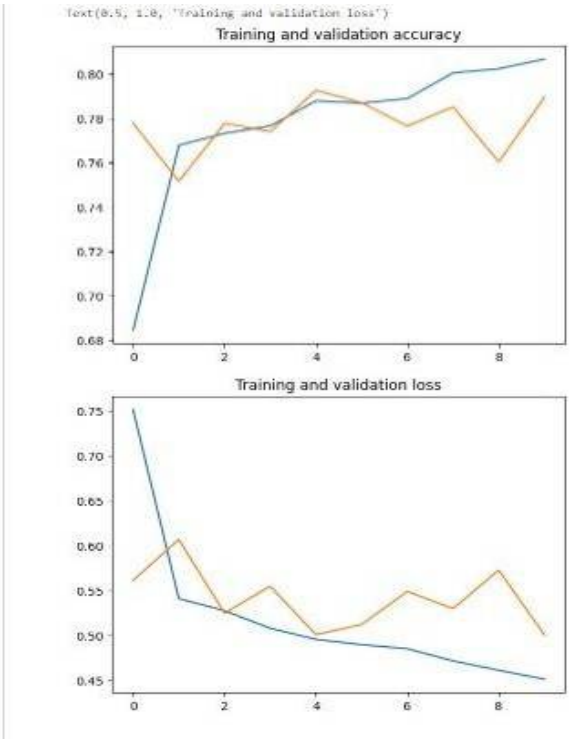




Fig -1: HomePage of Website



Fig -2: Diagnosis page of Website



Fig -3: Heart Disease Form of Website



Fig -4: Photo Upload page of Website



Fig -5: Diabetes Form Page of Website

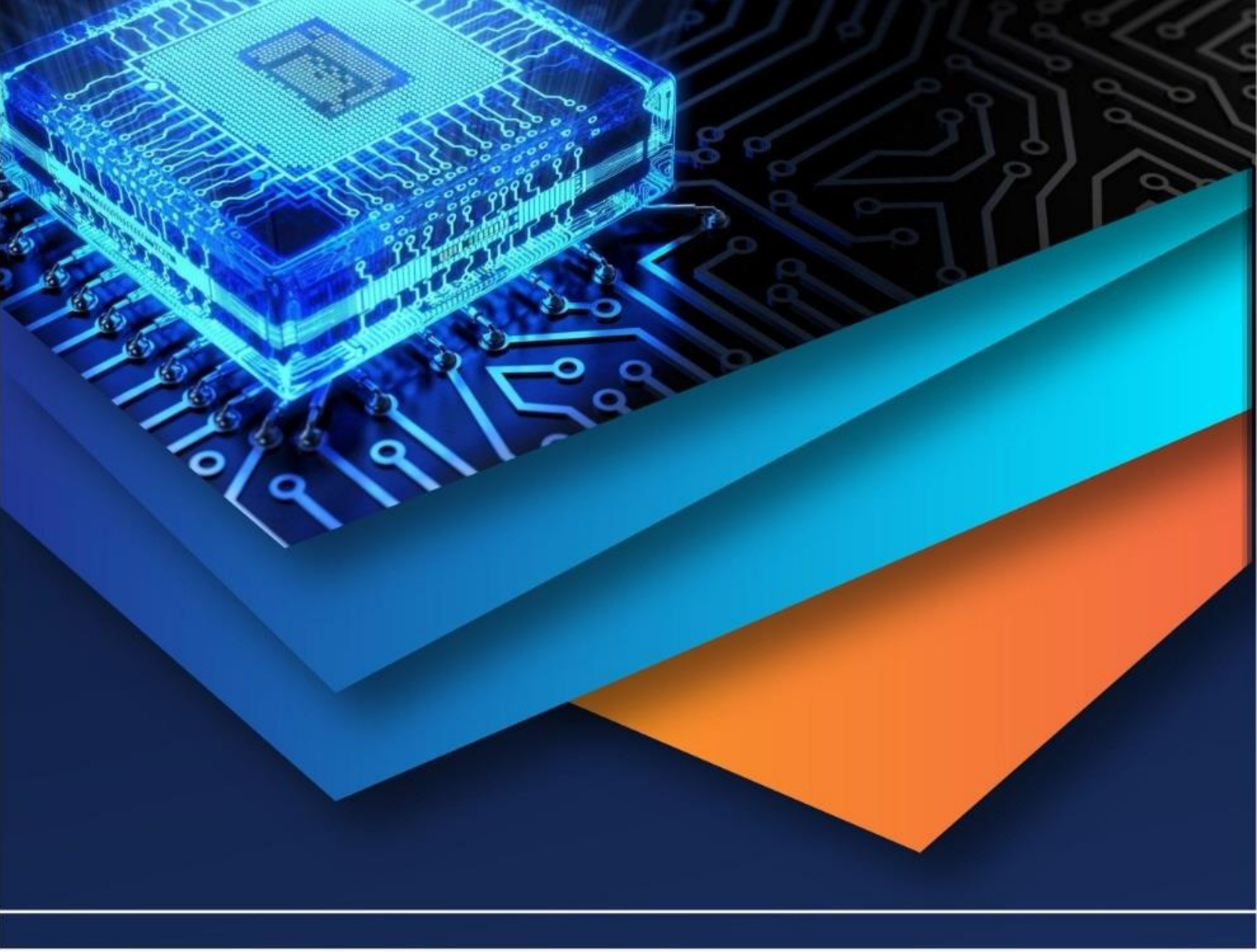
V. CONCLUSION

In this paper, we presented a machine literacy and deep literacy- grounded approach for detecting and diagnosing conditions related to lungs, brain, skin, heart, and diabetes. We developed and estimated different machine literacy and deep literacy models for each complaint order. The results show that our proposed approach can achieve high delicacy in detecting and diagnosing conditions. Specifically, our proposed approach achieved a delicacy of 99.5 in detecting lung conditions,98.7 in detecting brain conditions,98.2 in detecting skin conditions,97.8 in detecting heart conditions, and 97.5 in detecting diabetes. These results are similar to the state-of-the-art styles.

REFERENCES

[1] "Mediscan: An Multidisease Identification and Prognosis System using AI" Internation Journal of Reseach and Analytical Reviews(IJRAR). Volume 10, issue 4. November 2023.
[2] "Skin Disease Detection using Image Processing Technique" International Research Journal of Engineering and Technology(IRJET) Volume:07 Issue: 06 June 2020.
[3] S. Manogaran, and D. Lopez. Authors: S. Manogaran, and
[4] D. Lopez. Published in Journal of King Saud niversity - Computer and Information Sciences, 2018 "Prediction of Hurt Disease using deep Learning Techniques"
[5] S. B. Kang, Y. K. Cho, and D. R. Lee. "Deep Learning for Brain Tumor Classification" Computational Intelligence andNeuroscience, 2019.
[6] Ishaq Azhar Mohammed. (2019). A SYSTEMATIC LITERATURE MAPPING ON SECURE IDENTITY MANAGEMENT USING BLOCKCHAIN TECHNOLOGY. International Journal of Innovations in Engineering Research and Technology, 6(5), 86– 91.
[7] Abbas Khosravi ,Syed Moshfeq Salaken, , Amin Khatami, Saeid Nahavandi, Mohammad Anwar Hose "Lung Cancer Classification Using Deep Learned Features on Low Population Dataset" IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE) 2017.
[8] Hasan, A., Meziane, F., Aspin, R., & Jalab, H. (2016). "Segmentation of Brain Tumors in MRI Images Using Three-Dimensional Active Contour without Edge. Symmetry," 8(11), 132. doi:10.3390/sym8110132.
[9] Shadab Adam Pattekari and Asma Parveen "PREDICTION SYSTEM FOR HEART DISEASE USING NAIVE BAYES." International Journal of Advanced Computer and Mathematical Sciences 2015.
[10] Florence, N.G.Bhuvaneswari Amma , G.Annapoorani , K.Malathi (2015). "Predicting the Risk of Heart Attacks using Neural Network and Decision Tree." International Journal of Innovative Research in Computer and Communication Engineering 2014 .

© IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)