© NOV 2022 | IRE Journals | Volume 6 Issue 5 | ISSN: 2456-8880

A Deep Learning Based Multiple Chronic Disease Detection Model

FATI OIZA OCHEPA[1], JOHN PATRICK[2], MALIK ADEIZA RUFAI[3], ADAMU ISAH[4]
[1, 2, 3] *Department of Computer Science, Federal University Lokoja, Nigeria*
[4] *Department of Computer Science, Hussaini Adamu Federal Polytechnic, Kazaure, Nigeria*

Abstract*—The rapid advancement of technology in the health sector has paved the way for automating health related processes. Diagnosing of diseases is one of the most important and sensitive tasks performed by health practitioners that if not done efficiently, can lead to dire consequences for the patients. This study developed and implemented a model for diagnosing four life threatening diseases; pneumonia, malaria, breast cancer and skin cancer using Deep learning. The datasets used for this study was acquired from Kaggle and features were selected using the hybrid technique. A Convolutional Neural Network (CNN) model was deployed by using 80% of the data for training while the remaining 20% served for validation of the model. Based on the model, a web-based application diagnostic tool was deployed to enable patients gain easy access to efficient diagnosis. The model was assessed by using performance metrics such as precision, recall and F-measure. The overall accuracy of the model when tested on the four diseases dataset was 86.33%, 96.0%, 95.38% and 88.45% for pneumonia, malaria, breast cancer and skin cancer detection respectively.

***Indexed Terms*— Chronic Disease, Convolutional Neural Network, Deep Learning, Detection, Diagnosis.**

I. INTRODUCTION

Data in the health sector has become more comprehensive and in abundance over the last few decades. There has been a significant increase in the amount of information stored in Electronic Health Record (EHR) systems [1]. These records such as prior diseases history, lab reports, demographics, prescriptions, and procedures are gathered at the time of each health checkup in either structured or unstructured data form. The possibility to explore previously hidden information from large amount of image data used for medical diagnostics provides a great opportunity to improve healthcare quality. Recurring patterns are hard to identify because there are numerous patients and patients history. However, these patterns can be used to develop advanced algorithms and help increase disease detection speed while reducing the burden on healthcare providers.

The advancement in healthcare technology has made acquisition process of medical images easier. Though conventional medical imaging approaches have proved highly accurate for decades, advances in machine learning techniques have enabled a revolution in deep learning [2].

Deep learning is a subclass of machine learning technology that explores unstructured patterns in data and is inspired by how the human brain works. Deep learning models have a high possibility to succeed because the models are trained on hierarchy representation and they can also extract and organize unique attributes hence, they don't need any prior domain knowledge. Computational algorithms based on Convolutional Neural Networks have shown outstanding success in various fields, such as computer-aided diagnosis, speech recognition, and image classification.

[3] suggested that diagnostic errors has developed as a significant patient safety concern, yet they are difficult to detect and identify. The 14th International Diagnostic Error in Medicine Research Summit, an expert summit, debated on the difficulties encountered in defining and measuring errors during diagnosis. Among the challenges discussed during the summit are difficulties in determining error when the disease or diagnosis is evolving over time and in different care settings as well as accounting for a balance between

under-diagnosis and over-diagnosis [4]. The best way to address this problem is to implement machine learning or deep learning models to assist with disease detection.

This research aims at implementing machine learning model that will be used in the detection of Pneumonia, Malaria, Breast cancer and Skin cancer diseases. Deploy the model into a simple graphical interface that can be used by anyone and evaluate the performance of the developed model using appropriate metrics like confusion matrix, precision and recall. The development of a machine learning model for the early diagnosis of these four chronic diseases has become vital in order to lower the cost of conducting a test, the manpower required and consequently reduce the high death rate caused by these four diseases.

II. LITERATURE REVIEW

Machine learning models for the detection of chronic diseases have been proposed as reliable methods to reduce the errors resulting from diagnosis. This study propose multiple chronic disease detection model. Four diseases namely pneumonia, malaria, skin cancer and breast cancer will be used for the purpose of this study.

A. Pneumonia

According to World Health Organization [5], Pneumonia as an infectious disease is single handedly responsible for causing the most infection related deaths in children worldwide. Pneumonia deaths hit a total of 740,180 deaths accounting for 14% of all deaths in children under the age of 5. The use of machine learning in diagnosis can serve as an intervention in reducing the mortality rate of pneumonia related deaths.

In one of the comprehensive studies in the field, [6] implemented a novel, 18-layer deep sequential convolutional neural network-based model that is proven to outperform the state-of-the-art system for this task. A chest X-ray images dataset consisting of 5,856 X-ray images was utilized for the training and testing of the model. The model performs the 'normal' vs. 'pneumonia' classification task with the classification accuracy of 0.9439 which is around 1.6% better than the state-of-the-art system. The model yielded high sensitivity however, the model's limitation was that it yielded lower than the expected specificity (i.e. 0.86). Further improvements was suggested by investigating in details the application of techniques such as transfer learning, fine tuning and other relevant data augmentation.

As against the use of pre-trained networks, [7] built CNN model and models in Ensemble Learning from scratch to determine the presence of pneumonia using the chest X-ray dataset. This is after the SMOTE (Synthetic Minority Over-sampling Technique) method was used to deal with the imbalanced dataset of bacterial pneumonia and viral pneumonia classes. For each classification problem, two different deep learning methods which are CNN and ensemble learning was used and 95% average accuracy was obtained for each model, for binary classification and 78% and 75% average accuracy was obtained for each model respectively for multi class classification problem.

[8] explored some training techniques that helped the network learn better when making use of an unbalanced dataset (fewer cases of COVID-19 along with more cases from other classes). Their study proposed a neural network via concatenation of the Xception and ResNet50V2 networks by means of utilizing multiple feature extraction from the aforementioned networks. The proposed network was tested on 11302 images to report the actual accuracy achievable in real circumstances. The average accuracy of the proposed network for detecting COVID-19 cases was 99.50%, and the overall average accuracy for all classes was recorded at 91.4%.

The primary focus of these authors was to put forth the performances of different simple CNN architectures and select the best architecture based on optimum corresponding minimum loss and maximum accuracy which can serve as a viable tool for physicians and the medicine community to correctly identify and diagnose viral, bacterial, fungal-caused and community acquired pneumonia given only the chest X-ray of the patient. [9] achieved this by the utilization of convolutional neural networks (CNNs) of varying configurations on a machine learning based binary classification task with a given dataset of chest X-rays that depicts affected and unaffected cases of pneumonia.

[10] proposed the use of artificial intelligence for clinical image diagnosis of pneumonia. A dataset of 5216 train chest x-ray images and 624 test chest x-ray images for classification was sourced from Kaggle, with 2 classes as normal and pneumonia. Five mainstream network algorithms were used to classify diseases in the dataset and compared the individual results, from the results obtained, an improvement was implemented on MobileNet’s network structure and thereby achieved a higher accuracy rate than other methods.

The research conducted by [11] demonstrated the use of machine learning algorithms to process chest X-ray images in order to support the decision making process in determining the correct diagnosis. Specifically, the authors focused on the use of deep learning algorithm based on convolutional neural network in order to build a processing model. This model had the task of helping with a classification problem that is, detecting whether a chest X-ray shows changes consistent with pneumonia or not, and classifying the X-ray images in two groups depending on the detection results.

B. Malaria

Malaria is deadly but preventable disease that has taken the lives of an average of half a million persons every year. An estimated 240 million new malaria cases were reported in 2020, 627, 000 of those cases lead to malaria related deaths in 85 countries [12].

The presence of malaria in a blood sample is typically detected using microscopic blood cell examination. In 2010, about 167 million blood films were examined for malaria using microscopy, which was less expensive and uses less time than diagnosing with polymerase chain reaction. Despite its widespread use, microscopic diagnosis has a number of difficulties, it is prone to human error and the fact that malaria has a correlation with poverty and it is prevalent in low-income countries, where most laboratories and diagnostic facilities lack state-of-the-art testing facilities. Malaria has a significant influence on the health care system in underdeveloped countries because it leads to a global pathology shortage. Deep learning algorithms are used to assess medical images in today's computer-aided systems. Around the world, there is a movement to simplify diagnostic methods with the help of various machine learning approaches to assist human doctors in making the correct diagnosis [13].

The evaluation metric accuracy and loss along with 5-fold cross validation was used by [14] to compare and select the best performing architecture using the National Institute of Health named NIH Malaria Dataset. In order to maximize the performance, existing standard pre-processing techniques from literature were deployed to pick the best performing model. In addition, holdout test was conducted to verify how well the proposed model generalizes on unseen data. The best model achieves an accuracy of 97.77%±0.007.

A model proposed by [15] used the malaria dataset from Kaggle to detect cells from images of multiple cells in thin blood smear on standard microscope slides and classify them as either infected or uninfected with early and effective testing using image processing. And also perform classification on the infected cell image using machine learning. The model yielded an accuracy of 96.3%.

[16] compared the diagnostic performance of two rapid diagnostic tests (RDTs), parasite lactate dehydrogenase (PLDH) based Optimal IT and parasite aldolase base Dr. Grey’s with the gold standard microscopy in pregnant women. Blood samples of 113 pregnant women were collected from two hospitals in Lagos and screened for malaria parasite by Giemsa stained smear and two malaria rapid diagnostic tests, Optimal IT and parasite aldolase based standard microscopy in pregnant women. The results showed that, out of 113 blood samples screened, 23 (20.4%), 51 (45.1%) and 14 (12.4%) were positive by microscopy, Aldolase and PLDH respectively. Seventeen (33.3%) and 34 (66.7%) of the aldolase positive samples were positive and negative by microscopy respectively. Eleven (21.6%) of the aldolase positive samples were positive for both microscopy and PLDH. Of the positive samples, 34 (66.7%) were negative for both microscopy and PLDH. Three (13.0%) were positive by microscopy but negative by aldolase and PLDH (RDTs).

[17] In their research raised questions regarding the quality of RDT test results in health facilities, this prompted a comparative study of case detection based on RDTs and microscopy in line with clinical symptoms. The results showed that RDTs were more reliable in diagnosing malaria than microscopy while using PCR. [17] recommended that further studies will be required to test the accuracy and compliance of definitive diagnostic test for detection and management of malaria infections in clinical settings.
An overview of the current diagnostic methods used in malaria was conducted by [18]. The outcome of the overview revealed that the conventional microscopic examination of blood films remains the ideal malaria diagnostic tool as it is cheap and offers sensitivity unrivalled by any other field deployable method. The study also posited that while RDTs are convenient but could lead to over diagnosis and over treatment. Nucleic acid techniques offer the best sensitivities, they are however expensive, out of reach to most and are limited to clinical research.

C. Skin Cancer

Melanoma of the skin is the 17th most common type of cancer worldwide, the 13th most common cancer in men and the 15th most common cancer in women. It is quiet challenging to estimate its comprehensive occurrence for certain reasons. Multiple sub-type of skin cancers are in existence and this pose problems of data collection. Not all categories are properly documented because most of such cases are successfully treated via surgery or ablation. There were more than 150,000 new cases of melanoma of skin in 2020 [19].

An ensemble approach was adopted by [20] to provide a diagnostic for skin cancer using images and meta-data. The ensemble of classifiers had 13 convolutional neural networks (CNN) used on the nine classes of dataset. Two approaches were developed to handle the outlier classes. The study proposed an efficient method that has high accuracy in properly classifying the skin cancer images and detect the outlier classes.
The research report of [21] detailed the diagnosis of skin lesions that has been performed on a complete test image. It was found to work well with an accuracy rate of 98.7%. With a very simple and practical automatic diagnosis attained through well-defined segmentation (K-mean & PSO), feature extraction (SURF) and classification techniques (ANN). The utilization of K-mean with PSO helped to enhance the quality of the image and thereby an increase in detection accuracy.
[22] proposed a model designed into three phases comprising of data collection and augmentation, designing model and finally prediction. The researchers used multiple AI algorithms like Convolutional Neural Network and Support Vector Machine in combination with image processing tools to form a better structure, leading to higher accuracy of 85%.

A two neural network classifier was adopted by [23] namely Back-propagation neural network (BNN) and Auto-associative neural network (AANN). Recognition accuracy of the 3- layers back-propagation neural network classifier was pegged at 91% and auto-associative neural network at 82.6% in the image database that included dermoscopy photos and digital photos. The analysis of work was implemented using the MATLAB software.

[24] proposed a skin cancer detection system using support vector machine for early detection of skin cancer disease. It is more advantageous to patients. The diagnosing methodology proposed used image processing methods and Support Vector Machine (SVM) algorithm for classification. The dermoscopy image of skin cancer was taken and subjected to various pre-processing technique for noise removal and image enhancement. Then the images underwent segmentation process using Thresholding method, afterwards, some features of the images were extracted using GLCM methodology. These extracted features were given as the input to the classifier by classifying the given image into cancerous or non-cancerous images.

[25] presented a computer aided method for the detection of Melanoma Skin Cancer using Image Processing tools. The input to the system was the skin lesion images, the images were analyzed to determine the presence of skin cancer by applying novel image processing techniques. The Lesion Image analysis tools checks for the various Melanoma parameters like Asymmetry, Border, Color, Diameter,(ABCD) by using texture, size and shape analysis for the image segmentation and feature extraction stages. The

extracted feature parameters were used to classify the image as Normal skin and Melanoma cancer lesion.

D. Breast Cancer

In 2020 there were 684,996 deaths from breast cancer globally. In 2022, there will be an estimated 287,850 new cases of invasive breast cancer diagnosed in women; 2,710 cases diagnosed in men [26].

Identification of breast cancer can be conducted using either mammography, MRI scans, CT scans, ultrasound or nuclear imaging. However, none of the aforementioned methods can guarantee a 100% accurate prognosis. The most common approach of tissue-based diagnostics is staining. This involves staining components such as hematoxylin and eosin (H&E) in order to color tissue elements. This approach has several flaws that can be solved using deep learning approaches. A number of research work has been undertaken in the prediction of tumor class using variety of deep learning algorithms. While some authors used pre-built networks like AlexNet, MobileNet and Inception, several others scholars have proposed other approaches for identifying breast cancer.

To enhance improved detection of breast cancer, [27] proposed a system: ABUS technique that used mammography images combined with pitfalls compared with mammography images alone. The study showed that the ABUS technique recorded sensitivity 30% sensitivity higher than the mammogram but 20% less than HHUS while ABUS was 30% less specific than mammogram and 20% less specific than HHUS. The accuracy of HHUS was 20% more than that of both mammogram and ABUS. [28] proposed a hybrid approach for breast cancer diagnosis by reducing the high dimensionality of features using linear discriminant analysis (LDA), and then applying the new reduced feature dataset to Support Vector Machine. The focus of the study was to integrate these machine learning techniques with feature selection/feature extraction methods and compare their performances in order to identify the most suitable approach. The proposed approach obtained an accuracy of 98.82%, sensitivity of 98.41%, specificity of 99.07% and area under the receiver operating characteristic curve of 0.9994 after testing on the Wisconsin Diagnostic Breast Cancer (WDBC) Dataset.

An enhanced double thresholding mammography image segmentation approach proposed by [29] implemented its enhancement by adding borders as contours to the final segmented images of the original image thereby making it easier for physicians to have more accurate diagnosis. This technique, according to the study, can also be generalized on mammograms obtained via x-rays or other biomedical image sources. [30] calculated the texture features from mammogram using Gray Level Co-occurrence Matrix (GLCM) along 0°. From the calculated features, most effective features having large contribution to achieve the desired output were chosen and applied to Artificial Neural Network (ANN) for training and classification. The technique was applied on the mini-MIAS database and the overall sensitivity, specificity and accuracy achieved by using the proposed system was 99.3%, 100% and 99.4% respectively.

Most of the Literature reviewed focused on the detection and classification of four diseases namely; Malaria, Pneumonia, Skin Cancer and Breast Cancer using images based on Convolutional Neural Networks.

III. PROPOSED METHODOLOGY

A. The Dataset

The proposed multiple chronic disease detection model made use of four distinct datasets of the chronic diseases chosen for this study, the datasets were acquired from Kaggle. The pneumonia model made use of 1000 pair sets of normal chest x-ray and pneumonia chest x-ray. The X-ray image is labelled L for left chest x-ray and R for right chest x-ray. The malaria images obtained from Kaggle consisted of 1200 images for malaria infected cells and 1200 images for uninfected cells. 1300 images of unaffected breast cancer and affected breast cancer cells respectively was also acquired, unaffected cell were signified as 0 and affected signified as 1.The skin cancer dataset comprised of both images and a CSV file of melanoma skin cancer type. The CSV file contained already extracted pixel features of 2351 columns and their respective images of 10016.

B. Feature Selection

Though it is of great essence to choose the right model for the data, it is more paramount to choose the right data to feed into the model, as the choice of data selected impacts the model's overall performance. Feature selection is a technique undertaken to reduce the input variables in a model by selecting only the most relevant data. This process consequently allows for noise to be expunged from the data allowing the model to take in only data that is expected to have significant positive impact on the classification model's performance.

For this study, the hybrid technique was implemented. The hybrid method takes advantage of the advanced accuracy of the Wrapper feature selection method and the high computational productivity of the Filter feature selection method [31]. The hybrid of the wrapper and filter methods generated a crossed approach that significantly reduced the number of features, improved the classification accuracy and reduced the time and cost of computation. The four diseases to be detected in this study are namely Pneumonia, Malaria, Skin cancer and Breast Cancer, the model was built using CNN. Dataset Preparation
In developing the model, each of the dataset was split into two sets; which are the training set and testing set. 80% of the dataset was used for training, while 20% of the dataset used for testing. The training data was then fit into the machine for training and the test set used against the training set for accuracy.

C. Dataset Preparation

In developing the model, each of the dataset was split into two sets; which are the training set and testing set. 80% of the dataset was used for training, while 20% of the dataset used for testing. The training data was then fit into the machine for training and the test set used against the training set for accuracy.

D. System Design

The system design aimed at satisfying user requirements and to elucidate the workability of the system and its possible outcome. The system design for this study encompasses the logical and physical design detailing the steps from image acquisition, training of the model to the stage of disease classification using the developed model. Using design guides such as data flow diagram, use case diagram and activity diagram for the model, a web based application system was built.

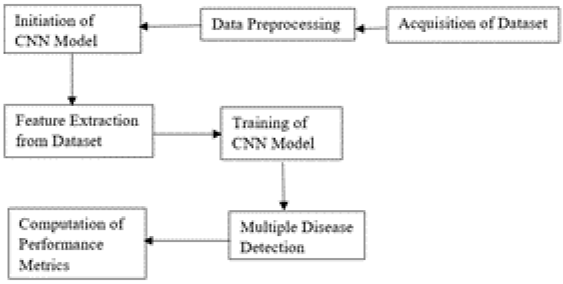

Fig. 1. Process flow diagram of the detection model

The system aims to be flexible enough by allowing users create their accounts and interact with the detection system by supplying relevant image data as required by the system. The doctor who doubles as the administrator los in to the system to receives images and make diagnosis based on the symptoms detected on user images.

IV. RESULTS AND DISCUSSION

The proposed multiple disease prediction model has been implemented and the results obtained from the evaluation of the machine learning model are presented in this section. The study proposed a deep learning based detection model for four chronic diseases which are pneumonia, malaria, skin cancer and breast cancer.

A. Performance Evaluation

An assessment of the efficiency of this system was performed using multiple metrics to evaluate how efficient the model is. Because performance is crucial in the design and use of a model, it strives to ensure that the model gets the optimum performance. The under listed metrics provided below gives us information on the quality of the results that was achieved in this study.

Precision. Precision refers to the ratio of all patients correctly detected to have a chronic disease to all the patients detected to have chronic disease (either true or false).

$$Precision = \frac{True\ Positive}{True\ Positive + True\ Negative} \quad (1)$$

Recall. Is the ratio of the actual number of correctly detected patients with chronic disease to the total number of patients with chronic diseases.

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \qquad (2)$$

F – Measure. Provides a means of combining the properties of both precision and recall into a single measure. It is the harmonic mean between precision and recall.

$$F - Measure = 2 * \frac{Recall * Precision}{Recall + Precision} \qquad (3)$$

Accuracy. Is the ratio of all the output that was correctly detected to all the occurring cases in the data set.

$$Accuracy = \frac{True\ Positive + rue\ Negative}{True\ Pos. + False\ Pos. + True\ Neg. + False\ Neg.}$$

(4)

B. Results

Below is the results on the accuracy, precision, recall and f-measure of the model on the four chronic disease (pneumonia, malaria, breast cancer and skin cancer) datasets.

Table I. Accuracy of the model on the four chronic disease (pneumonia, malaria, breast cancer and skin cancer) datasets

Disease	*Accuracy (%)*
Pneumonia	86.33
Malaria	96.0
Breast Cancer	95.38
Skin Cancer	88.45

Table I above shows the accuracy of detection that was obtained from the chronic disease detection model. The accuracy obtained for pneumonia, malaria, breast cancer and skin cancer disease detection were 86.33%, 96.0%, 95.38% and 88.45% respectively. This shows that our model performed best in the detection of malaria closely followed by detection of breast cancer.

Table 2. Result of Precision, Recall and F – Measure.

Disease	*Precision*	*Recall*	*F-measure*
Pneumonia	0.875	0.816	0.844
Malaria	0.998	0.955	0.975
Breast Cancer	0.938	0.986	0.960
Skin Cancer	0.925	0.860	0.891

Table II depicts the results of the other performance evaluators (precision, recall and F-measure) used in assessing the overall performance of the model.

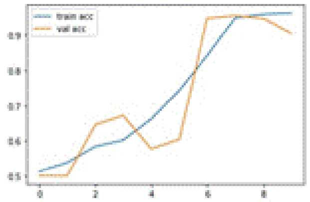

Fig.2. Malaria training and validation accuracy

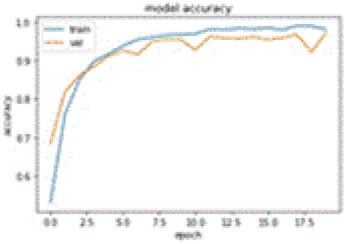

Fig. 3. Skin cancer training and validation accuracy

Fig. 2 and Fig. 3 shows the accuracy obtained in detection of malaria and skin cancer respectively after training and testing of the detection model.

Fig. 4. The Multiple chronic disease detection web application main page

Fig.4 above displays the main page of the web based for the multiple disease detection, the main page has dedicated links for the four diseases being used in this study. A new user is also expected to register and login from this page as shown on Fig. 5 below.

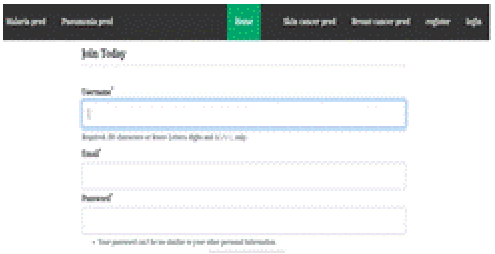

Fig. 5. User login details request form.

Fig. 6. User image input interface

On login in, a registered user can upload image data to the relevant detection link for diagnosis as shown in Fig. 6 above.

V. CONCLUSION AND FUTURE WORK

This study adopted the use of deep learning techniques in improving the delicate task of chronic diseases diagnosis in the health sector. The most significant contribution of this study is the deployment of the model in building a multiple chronic disease web application diagnostic tool that allows users to create an account, login and aided by the automated system, detect either pneumonia, malaria, skin cancer or breast cancer by uploading the relevant image input. Data for the four diseases considered in this study was sourced from Kaggle. The machine processes in diagnosing these diseases has brought about a boost in the medical field and help in reducing processing time, risk of misdiagnosing and mortality rate. The model achieve accuracy of 86.33%, 96.0%, 95.38% and 88.45% for pneumonia, malaria, breast cancer and skin cancer detection respectively. Future research work could focus on further improving the accuracy of the detection model by incorporating a feature selection framework that will prioritize selection of most valuable features.

REFERENCES

[1] N. Adler, and J. Casey, -Using electronic health records to inform knowledge of social determinants of health.*Annual Review of Public Health*, 2016.

[2] S. Sandhiya, and U. Palani, -An effective disease prediction system using incremental feature selection and temporal convolutional neural network, *Journal of Ambient Intelligence and Humanized.Computing*, vol. 11, pp. 5547 -5560, April 2020.

[3] L. Zwaan, and H. Singh, -The challenges in defining and measuring diagnostic error. *Diagnosis*, vol. 2, pp. 97-103, May 2015.

[4] The Diagnostic Error in Medicine 14th Annual International Conference. *Diagnosis (Berl)*, pp.294-386, April 2022.

[5] World Health Organization. Pneumonia. October 2020. [Online] Available: https://www.who.int/news-room/fact-sheets/detail/pneumonia,

[6] S. Raheel, -Automated pneumonia diagnosis using a customized sequential convolutional neural network. *In Proc. of the 2019 3rd International Conference on Deep Learning Technologies Association for Computing Machinery*,2019, pp. 64-70.

[7] M. B. Darici, Z. Dokur, and T, Olmez -Pneumonia detection and classification using deep learning on chest x-ray images. *International Journal of Intelligent Systems and Applications in Engineering*, vol. 8 (4), pp. 177-183, October 2020.

[8] M. Rahimzadeh, A. A. Attar,-modified deep convolutional neural network for detecting COVID-19 and pneumonia from chest X-ray images based on the concatenation of Xception and ResNet50V2. *Informatics in medicine unlocked*, vol. 19, June 2020.

[9] G. Jee, G. M. Harshvardhan, and M. K. Gourisaria, -Juxtaposing inference capabilities of deep neural models over posteroanterior chest radiographs facilitating COVID-19 detection. *Journal of Interdisciplinary Mathematics*, vol.24(2), pp. 299-325, May 2021.

[10] Z. Yue, L. Ma, and R. Zhang, -Comparison and validation of deep learning models for the diagnosis of pneumonia. *Computational Intelligence and Neuroscience*, vol.1, pp. 1-8, September 2020.

[11] L. Račić, T. Popovic, S. Cakic, and S. Šandi, -Pneumonia detection using deep learning based on convolutional neural network. *International Conference on Information Technology (IT)*, vol. 25, May 2021.

[12] World Health Organization. Global Malaria Strategy. October 2022. [Online] Available: https://www.who.int/campaigns/world-malaria-day/2022

[13] K.Y. Tai, and J. Dhaliwal, -Machine learning model for malaria risk prediction based on mutation location of large-scale genetic variation data. *Journal of Big Data*, vol. 9, pp. 85-91, November 2022.

[14] S. Rajaraman, S. Jaeger, and S.K. Antani, -Performance evaluation of deep neural ensembles toward malaria parasite detection in thin-blood smear images. *Peer Journal*, vol. 28, October 2019.

[15] K. Motwani, A. Kanojiya, C. Gomes, and A. Yadav, -Malaria Detection using image processing and machine learning. *International Journal of Engineering Research & Technology*, vol. 9, September 2021.

[16] O. Adeola *et al.* -Performance of Microscopy Method and Rapid Diagnostic Tests in Malaria Diagnosis amongst Pregnant Women in Lagos, Southwest Nigeria. *Diversity & Equality in Health and Care*, vol. 15(3), pp. 104-109, October 2018.

[17] B. Dinko, *et al.* -Comparison of malaria diagnostic methods in four hospitals in the Volta region of Ghana. *MW Journal*, vol. 7(5), November 2016.

[18] O. Olabode, J. Openibo, and G. I. Olasehinde, -Malaria diagnosis: current approaches and future prospects. *In 3rd International Conference on African Development Issues*, vol. 5, May 2016.

[19] World Cancer Research Fund International. Skin cancer statistics. October, 2022. [Online] Available: https://www.wcrf.org/cancer-trends/skin-cancer-statistics/.

[20] A.Pacheco, A. Ali, and T. Trappenberg, -Skin cancer detection based on deep learning and entropy to detect outlier samples. *Computers in Biology and Medicine*, vol. 116, November 2019.

[21] M. Jafari, E. Nasr-Esfahani, N. Karimi, S. M. R. Soroushmehr, S. Samavi, and K. Najarian, -Extraction of skin lesions from non-dermoscopic images for surgical excision of melanoma. *International Journal of Computer Assisted Radiology and Surgery*, vol. 12, pp. 1021-1030, March 2017.

[22] M. M. Vijayalakshmi, -Melanoma skin cancer detection using image processing and machine learning. *International Journal of Trend in Scientific Research and Development*, Vol. 3(4), pp.780-784, December, 2019.

[23] S. Bano, and A. Srivastava, -Skin cancer detection using classification framework of neural network. *International Journal of Modern Communication Technologies and Research*, vol.6 (4), August 2018.

[24] U. B. Ansari, and T. Sarode, -Skin cancer detection using image processing. *Int Res Journ of Eng Technol*, vol.4(4), pp. 2875-2881, March 2017.

[25] S. Jain, V. Jagtap, and N. Pisea, -Computer aided melanoma skin cancer detection using image processing. *International Conference on Intelligent Computing, Communication & Convergence Procedia Computer Science*, vol.1, pp. 735-740, July 2015.

[26] American College of Surgeons Clinical Congress. Breast Cancer. October 2021, [Online] Available: https://www.stopbreastcancer.org/information-center/facts-figures, Oct. 2022.

[27] G. Mohamed, K. Sherine, and E. Gad, -Breast cancer detection using automated breast ultrasound in mammographically dense breasts. *The Medical Journal of Cairo University*, 2020, 8(8):1715-1723.

[28] D. Omondiagbe, S. Veeramani, and A. Sidhu, -Machine learning classification techniques for breast cancer diagnosis. I*OP Conference Series: Materials Science and Engineering*, vol. 495, March 2019.

[29] S. M. Badawy, A. A.Hefnawy, H. E..Zidan, and M. T. GadAllah, -Breast cancer detection with mammogram segmentation: a qualitative study. *International Journal of Advanced Computer Science and Applications,* vol. 8(10), August 2017.

[30] N. Tariq, Breast cancer detection using artificial neural network. *Journal of Molecular Biomarkers and Diagnosis*, vol. 9(1), pp. 1-6, July 2017.

[31] M. Kavitha, and S. Subbaiah -A study on feature selection for chronic disease prediction. *International Journal of Emerging Technologies and Innovative Research*, vol. 6(3), pp.351-356, May 2019.