ISSN: 1672 - 6553

JOURNAL OF DYNAMICS AND CONTROL

VOLUME 9 ISSUE 4: 13-22

TRENDS ON MACHINE LEARNING TECHNIQUES USED IN MEDICAL FOR DISEASES DETECTION

Shaba Irram, Dr. Mohammad Suaib

Integral University Lucknow, India

https://doi.org/10.71058/jodac.v9i4002

TRENDS ON MACHINE LEARNING TECHNIQUES USED IN MEDICAL FOR DISEASES DETECTION

[1*]**Shaba Irram,** [2]**Dr. Mohammad Suaib**

Integral University Lucknow, India

[1]*helloshabairram@gmail.com,* [2]*suaibcs09@gmail.com*

[*]***Corresponding Author***

Abstract: *The branch of computer science that focuses on creating machines that behave like humans is called artificial intelligence (AI). Medical disease detection is a rapidly expanding field of study in artificial intelligence. Many efforts have been made in recent years to enhance medical disease detection since mistakes and issues with this process can result in grave medical errors and incorrect treatment. In the field of biomedicine, meta-heuristic approaches have been widely used to identify medical conditions and offer improved perception and prognostication accuracy. Consequently, Swarm intelligence has primarily been used to address the various types of optimization problems due to the versatility of numerical experimentation. Nonetheless, despite the widespread use of Swarm intelligence techniques for disease detection, a gap remains in the comparative survey. This paper provides an overview of the different approaches used in medical disease diagnosis. The goal of this study is to provide a thorough literature review of the most recent methods for disease detection knowledge discovery. While the main objective is to provide directions for future enhancement and development in this area, the systematic analysis also reveals research gaps related to Swarm intelligence strategies. This paper provides an organized overview of the conceptual model for advanced research that has been studied thus far in the designated literature. The review adheres to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) principles and synthesizes papers from the 2020-2024 to assess trends, effectiveness, and future directions.*

Keywords: *Machine Learning, Medical disease, Diseases Detection*

1. Introduction

In recent years, using computational intelligence to diagnose medical conditions has gained popularity. The increased availability of healthcare data, along with advances in computer power, has enabled machine learning models to improve diagnostic accuracy, aid in early disease identification, and optimize clinical decision making. Various techniques for diagnosing medical conditions can be categorized as tasks involving intelligent data classification. Two categories can be identified based on the total number of groups that are consistently distributed in the classification techniques. Binary Classification (Two-class task) is the first classification distribution that only distinguishes between the two classes in the data. Data from more than two classes can be distinguished using the second classification method, known as multi-classification (multi-class task). To increase the authenticity (accuracy) of data classification, a number of scientists and researchers working in the medical field have experimented with various methods. In this domain, state-of-the-art algorithms like Tabu Search, Genetic Algorithm (GA), Bat Algorithm (BA), Particle Swarm Optimization (PSO), and data mining tools like Decision Tree and Neural Networks have been applied recently, and the results have been impressive [1].

Furthermore, the classifications of medical datasets are used in disease detection, excluding the other standard classification complexities. As a result, medical professionals or patients need to be aware of the symptoms that were employed in the classification process in addition to observing the evaluated classification findings. Neural Networks (NN) and linear programming models have been proposed as solutions to these kinds of issues. Nevertheless, the classification models' decision-making processes remain opaque, offering no insight into how the desired outcomes were achieved. The problems brought about by black box techniques have also been addressed by hybrid approaches like NN or GA that incorporate fuzzy rules, but they are still unable to identify which input factors are more appropriate than others. Numerous researchers have employed the state-of-the-art PSO algorithm in the literature to solve this problem by integrating it with a number of other strategies, including K-nearest neighbour and random forests [2], among others.

This review categorizes several machine learning approaches used in illness diagnosis, including their benefits and limitations, as well as the problems associated with incorporating them into medical practice.

https://doi.org/10.71058/jodac.v9i4002

2. Methods

2.1 Search Strategy

Several electronic databases, such as PubMed, Scopus, Web of Science, IEEE Xplore, and Google Scholar, were used in a methodical search. Medical and machine learning-related keywords, such as "deep learning for healthcare," "AI in medical diagnosis," "machine learning in disease detection," and "ML for medical imaging," were combined in the search approach.

2.2 Inclusion and Exclusion Criteria: Relevant studies were chosen using the following criteria

Inclusion Criteria:

- Research that used machine learning methods to identify disease.
- Studies that have been published in prestigious conference proceedings or peer-reviewed journal.
- Research that offers machine learning model performance measures (such as accuracy, sensitivity, specificity, and AUC-ROC).
- Research using publicly accessible medical datasets or clinical datasets

Exclusion Criteria:

- Research without methods based on machine learning.
- Opinion, editorial, and review papers.
- Research without quantitative assessment criteria.
- Articles that aren't fully accessible.

2.3 Data Extraction and Synthesis: Information that was extracted included study specifics (author and year), type(s) of disease examined. examples of machine learning algorithms that are utilized include SVM, Random Forest, CNN, RNN, and Transformer models, dataset attributes (source, size, and data preprocessing techniques), performance evaluation metrics (e.g., F1-score, AUC-ROC, recall, accuracy, and precision) and benefits and drawbacks of the strategy. Machine learning methods were compared across several medical applications using a narrative synthesis.

Ashraf A. et al. [3] used a combination of data from the autism brain imaging data exchange (ABIDE I and ABIDE II) datasets to classify and represent learning tasks of the most powerful deep learning networks, such as convolution neural network (CNN) and transfer learning algorithm. Resting state-fMRI (rs-fMRI) data can be used to develop diagnostic biomarkers for brain dysfunction due to their four-dimensional nature (three spatial dimensions and one temporal dimension). ABIDE is a global scientific collaboration in which 1112 rs-fMRI datasets from 573 typical control (TC) and 539 autism individuals, and 1114 rs-fMRI datasets from 521 autism and 593 typical control individuals, respectively, were collected from 17 different sites. Their proposed optimised version of CNN achieved an accuracy of 81.56%. This performs better.

M. Hussain et al. [4] described a novel approach to overcoming critical challenges in the automated detection of diabetic retinopathy using fundus images. They investigated the process of modelling fundus images by proposing the 'Retina-Based Affine Mapping' mechanism to combat data scarcity. This facilitated the generation of representative augmentations to model occurrences influenced by various internal and external factors during fundus image acquisition, departing from previous works that focused primarily on data scaling rather than increased representation. Furthermore, they proposed a 'Design Flow Mechanism' to streamline custom Convolutional Neural Network architecture development resulting in a highly efficient model with 99.51% validation accuracy using only 1.40 million parameters, outperforming state-of-the-art alternatives like ResNet-18, which has 11.69 million learnable parameters. These contributions advance the field of automated DR detection, with significant advancements in medical image analysis and early disease diagnosis on the horizon.

Almohimeed A. et al. [5] proposed a novel multi-level stacking model for predicting different types of Alzheimer's Disease by combining heterogeneous models and modalities. Cognitive sub-scores (e.g., clinical dementia rating - sum of boxes, Alzheimer's disease assessment scale) from the Neuroimaging Initiative dataset are among the modalities. To train each modality (ADAS, CDR, and FQA) in level 1, They used six base models (Random Forest (RF), Decision Tree (DT), Support Vector Machine (SVM), Logistic Regression (LR), K-nearest Neighbours (KNN), and Native Bayes (NB). Then, for the training set, they constructed stacking training, which combines the outputs of each base model, and for the testing set, they constructed staking testing, which combined the outcomes of each model. In order to create training stacking for the training set and testing stacking for the testing set, Level 2 combined the output of six base models (RF, LR, DT, SVM, KNN, and NB) to create three stacking models for each modality. Meta-learners (RF) are trained using stacking training, and meta-learners (RF) are evaluated using stacking testing. Finally, at level 3, the output prediction of the stacking model from each modality (ADAS, CDR, and FQA) in the training and testing datasets is merged to create a new dataset for staking training and stacking testing. The meta-learner is trained using training stacking, and the testing set is used to evaluate the meta-learner and generate the final

https://doi.org/10.71058/jodac.v9i4002

prediction. Through explainable artificial intelligence (XAI), they also hope to provide model explanations, ensuring efficiency, effectiveness, and trust. To select the most appropriate sub-scores, feature selection optimisation based on Particle Swarm Optimisation is used. The proposed model has a high potential for improving early disease detection. The results show that multi-modal approaches outperform single-modal approaches. Furthermore, when compared to regular ML classifiers and stacking models using full multi-modalities, the proposed multi-level stacking models achieve the highest performance with selected features, with accuracy, precision, recall, and F1-scores of 92.08%, 92.07%, 92.08%, and 92.01% for two classes, and 90.03%, 90.19%, 90.03%, and 90.05% for three classes, respectively.

S. A. Salih et al. [6] used two data sets in this study: X-ray medical images represent the first group, and CT images represent the second. To verify and test them, 250 RGB medical images of Covid-19 are trained on image data using the developed artificial neural networks (ANN) algorithm. Furthermore, they compared two types of medical images, X-ray and CT, to 15 neurons. Their findings revealed that the ANN performed admirably. The first group had 99.8% accuracy, sensitivity, specificity, F1 score, and accuracy, while the second group had 100% accuracy. Furthermore, their approach outperformed previous machine learning models.

K. Khalil et al. [7] proposed, a federated learning (FL)-based diagnostic model for Alzheimer's disease using blood bio-samples. To test and compare the efficacy of their models, they used blood bio-sample data sets downloaded from the ADNI website. They used a hardware acceleration scheme to implement FL model to speed up the training and testing operations based on the massive data collected for early AD detection. VHDL and an Altera 10 GX FPGA are used to implement the hardware accelerator method. The simulation results show that the proposed algorithms achieve early detection accuracy and sensitivity of 89% and 87%, respectively, while taking less time to train than other state-of-the-art algorithms.

S. Khan et al. [8] described an advanced transfer learning-based mechanism that uses AlexNet in conjunction with Inception-V4 to detect a brain haemorrhage automatically. It also compared performance to individual deep learning models. Experiment results on the Computed Tomography scans dataset show that the proposed approach outperforms other machine learning methods in terms of accuracy and F1 score, including a support vector machine (SVM), two-dimensional convolution neural network (2D CNN), and bidirectional long short-term memory (BLSTM). The proposed method achieves the highest accuracy of 94.54%, which is significantly higher than the accuracy of 2D CNN, BLSTM, and SVM, which are 85.07%, 79.2%, and 71.53%, respectively. Furthermore, the highest F1 score for the proposed approach is 0.938, which is significantly higher than the 0.846, 0.781, and 0.693 for 2D CNN, BLSTM, and SVM, respectively. The proposed approach's performance in terms of accuracy, time consumption, and F1 score, as well as its non-data-hungry nature, indicate its potential usefulness for brain hemorrhage detection.

A.Zafar et al. [9] described an improved method for detecting and segmenting lung tumours that is both efficient and precise. The proposed method employs a multimodal approach, utilising both CT and PET scans to improve tumour detection. For effective tumour classification, the methodology incorporates cutting-edge deep learning architectures such as ResNet, DenseNet, and Inception-v3. To integrate data from multiple modalities, both immediate fusion (early fusion) and late fusion techniques are used. Metrics such as precision, F1 score, accuracy, and sensitivity are used to assess the performance of classification models. The experimental results show that the proposed method is effective at accurately segmenting lung tumours. The findings add to existing knowledge in tumour segmentation and medical image analysis by providing valuable insights into the benefits of multimodal fusion and deep learning techniques for lung cancer diagnosis and treatment planning.

N. Pratama et al. [10] analysed and evaluated the accuracy of the deep learning algorithm and its application to detecting or diagnosing pneumonia from chest x-ray images. This paper used a systematic literature review approach on several papers found on Google Scholar via a keyword-based search mechanism related to pneumonia detection and deep learning. The authors compare and rank each method in the papers based on accuracy. The AlexNet model has the highest accuracy at 99.62%. Although deep learning models can be used to detect pneumonia, a doctor's supervision is still required to avoid misdiagnosis.

R. V. Sharan et al. [11] suggested a completely automated method for analysing cough sounds in order to differentiate paediatric pneumonia from other acute respiratory illnesses. Cough sound segmentation, cough sound classification, and cough sound denoising are all part of the suggested approach. While the segmentation algorithm directly detects cough sounds from the denoised audio waveform, the denoising algorithm uses multi-conditional spectral mapping with a multilayer perceptron network. They extracted multiple hand-crafted features and feature embeddings from a pretrained deep learning network from the segmented cough signal. The combined feature set is used to train a multilayer perceptron for the purpose of detecting pneumonia. They assessed proposed method on a dataset of 173 children's cough sounds who were diagnosed with acute respiratory diseases or pneumonia. The denoising algorithm increased the signal-to-noise ratio by 44% on average. Additionally, cough sounds alone can detect childhood

https://doi.org/10.71058/jodac.v9i4002

pneumonia with a sensitivity and specificity of 82% and 71%, respectively, and 91% and 86%, respectively, in cough segmentation. This indicates how it can be used, for example, to quickly diagnose issues using smartphone technology.

Sutradhar A. et al [12] suggested a novel machine learning (ML) based disease prediction system that takes into account three important steps in order to potentially predict it. First, three feature selection methods—Feature Importance (FIS), Information Gain Selections (IGS), and Least Absolute Shrinkage and Selection Operator (LAS)—were used to reduce the dimension of the dataset. Also taken into account when creating a feature set with the same characteristics as High-Risk Factors (HRF) were suggested medical references. Second, the models are applied to the training data set as classifiers. These models include the Three Stage Hybrid Artificial Neural Network (3SHANN) and the Three Stage Hybrid Classifier (3SHC). Third, each prediction was independently explained by applying a Local Interpretable Model-agnostic Explanations (LIME) to the 3SHC using the HRF samples. Then, a Partial Dependency Plot (PDP) was utilised to examine the general behaviours of both gender and age groups. In the end, a comprehensive set of experiments validates the suggested system, and the 3SHC achieves an accuracy (ACC) of 99.29%. This system has the potential to significantly reduce stress and prevent thyroid disease in the healthcare industry.

M. Mahadevi et al. [13] worked at how segmentation, classification, and treatment outcome prediction techniques have advanced recently for the study of pancreatic and lung cancers. EM methods and the Gaussian Mixture model are used for preprocessing. Following the segmentation of the lung and pancreatic tumours, the classification model is applied to ascertain the disease prediction. Classical machine learning techniques have shown promising results in accurately identifying lung tumours from CT images, which could aid radiologists in the early detection of malignancies. To do this, pancreatic tumours must be segmented and their edges must be defined from medical images. Lastly, deep learning holds great promise for studying lung and pancreatic cancers since it provides practical techniques for early detection, cancer segmentation, and outcome prediction. Some issues need to be resolved in order to ensure the dependability, accessibility, and generality of models developed using deep learning in clinical practise.

A.N. et al. [14] allowed for the quick and accurate location and identification of aberrant cells, facilitating prompt intervention and better patient outcomes. Moreover, the suggested framework makes use of the Nvidia Jetson Nano developer kit to assess the effectiveness of YOLO tiny detectors in real-time, with the goal of achieving an integrated deep prostate cancer detection system (PCDS). For the training and testing stages, the study uses a dataset with 3585 instances. The outcomes of the experiment show that YOLO v8s achieves a mean average precision (mAP) of 99%, precision of 99.5%, recall of 99%, and F1-score of 99.2%. These results demonstrate the efficacy of this approach in detecting prostate cancer in real-time, and it holds significant potential for improving medical diagnostics.

Y. Sneha et al. [15] conducted a thorough assessment of machine learning models to detect brain tumours in MRI scans using a dataset of 2034 images. Based on their performance metrics, four alternative models—Convolutional Neural Networks (CNN), Support Vector Machines (SVM), Random Forest (RF), and Logistic Regression (LR)—were carefully investigated. The best model was found to be the Convolutional Neural Network, which had excellent recall, precision, and F1 scores in addition to a high accuracy of 97.5%. This illustrates how brain tumour identification in medical imaging can be improved and automated through the use of deep learning, specifically CNNs. The study also emphasises the robustness and flexibility of the SVM and RF models, which showed remarkable performance metrics and were thus suitable for use in practical healthcare applications. The Logistic Regression model also significantly helps in diagnosis, although it has much lower scores. In conclusion, this research highlights the value of machine learning in the healthcare industry and its potential for early brain tumour identification, which would enhance patient care and treatment results. It highlights how medical image analysis is constantly developing and getting better, which leads to advancements in the identification and treatment of serious medical conditions.

Dhanushyar et al. [16] offered a ground-breaking technique for enhancing precision medicine in thyroid disease diagnostics through the use of machine learning algorithms. The goal of this study is to investigate how machine learning models can be used to increase the accuracy and speed of thyroid problem diagnosis. Prediction models for thyroid problems are built using a variety of machine learning techniques, including random forests, deep learning neural networks, and support vector machines. The performance of each model was evaluated using a variety of metrics, such as area under the curve (AUC) of the receiver operating characteristic (ROC) curve, accuracy, sensitivity, and specificity. Their tests' findings show that machine-learning models perform noticeably better at diagnosing thyroid disorders than traditional methods. These models can detect thyroid abnormalities with high sensitivity and accuracy, which enables faster and more precise thyroid problem diagnosis.

A.Marouf et al. [17] investigated the use of tissue microarray (TMA) and clinical data gathered by Alberta Precision Laboratory pathologists to predict prostate cancer severity using a variety of machine learning techniques. Through the machine learning pipeline that includes imputation and sampling techniques, traditional classifiers like Naïve Bayes, Decision Tree, Support Vector Machine with Radial basis function (RBF), Logistic Regression, and ensemble

https://doi.org/10.71058/jodac.v9i4002

classifiers like Random Forest and Bagging with k-nearest neighbours have been applied. A combined SMOTE-Tomek Links approach is used to address the issue of class imbalance. 99.64% accuracy is the highest that the Random Forest method can achieve.

Y. Guan et al. [18] created and tested a framework that combines a deep learning diagnosis model with block chain technology in this study. An attention-based pyramid semantic segmentation network and a discrete wavelet transformation-processed residual classification network are combined to form the diagnosis model. They also compared the performance of benchmark models to that of our models. With an accuracy of 91.77%, their diagnosis model outperformed benchmarks on the segmentation and classification tasks. They used the InterPlanetary File System protocol to create a secure and private sharing environment for the blockchain system. This framework can automatically grade the severity of UPJO using ultrasound images, ensure secure medical data sharing, aid doctors' diagnostic abilities, relieve patients' burdens, and provide technical support for future federated learning and Internet of Medical Things linkage.

Y. Zhong *et al*. [19] improved malaria diagnosis efficiency by combining smartphones and microscopes. Image acquisition and algorithmic detection of malaria parasites in various thick blood smear (TBS) datasets sourced from various global regions, including low-quality images from Sub-Saharan Africa, are part of this integration. Methods: To distinguish between white blood cells, artefacts, and malaria parasites, this method combines image segmentation and a convolutional neural network (CNN). A portable system combines a microscope and a graphical user interface to aid in the rapid detection of malaria from smartphone images. The CNN model was trained using open-source data from the Chittagong Medical College Hospital in Bangladesh. The validation process demonstrated that the proposed model achieved an accuracy of 97.74% 0.05% and an F1-score of 97.75% 0.04% using microscopic TBS from both the training dataset and an additional dataset from Sub-Saharan Africa.

Assefa M. et al. [20] sought to provide an overview of bacterial pneumonia with a focus on gram-negative aetiology, pathogenesis, risk factors, resistance mechanisms, treatment updates, and vaccine concerns. In summary, the global prevalence of GNB in CAP (Community acquired pneumonia) patients ranged from 49.7% to 83.1%, while it ranged from 76.13% to 95.3% in VAP (Ventilator-associated pneumonia) patients. The most common MDR-GNB (Multi-Drug-Resistant Gram-Negative Bacteria) causes of pneumonia that were reported were A. baumannii, K. pneumoniae, and P. aeruginosa; older and VAP patients were more likely to have A. baumannii isolated from them. Amoxicillin-clavulanic acid, ampicillin, tetracyclines, cephalosporins, and carbapenems were found to be highly resistant in the majority of studies. MDR-GNB colonisation is linked to a number of factors, including prior MDR-GNB infection, advanced age, prior use of broad-spectrum antibiotics, high rates of local antibiotic resistance, extended hospital stays, ICU admission, mechanical ventilation, and immunosuppression. Because S. maltophilia can form biofilms, adhere to the respiratory device's surface, and have both intrinsic and acquired drug resistance mechanisms, it has been reported to be a severe cause of HAP/VAP in patients undergoing mechanical ventilation and hematologic malignancy. In the future, pathogen-specific lymphocytes, gene-based vaccinations, antibiofilm agents, and effective combination therapies that target drug-resistant genes and PDR strains should be developed.

Anisha I. et al [21] has designed and implemented a CAD system to diagnose lung disorders from chest CT slices. The six main subsystems that make up the CAD framework are the segmentation, preprocessing, feature extraction, feature selection, and classification subsystems; they also include the Region of Interest (ROI) extraction subsystem. The diagnostic accuracy in identifying the various subtypes of emphysema is assessed using a real-time emphysema dataset. The suggested model performed better than the Gradient Descent BPNN classifier, which had an accuracy of 82.41%, with an accuracy of 87.52%. In order to help doctors give patients the best care possible, the work can be expanded to assess the severity of emphysema subtypes. Further research could be done on more clinically significant feature descriptors that are involved in the disease classification process. As an extra line of inquiry into the clinical findings, the relationship between emphysema and other pulmonary diseases can also be examined.

Chandra et al. [22] developed an automatic CAD system for PD detection, leading to more reliable and accurate diagnostic decisions. Furthermore, a second opinion will be provided by this kind of CAD system to help doctors reduce the number of missed diagnoses and their overall workload. In situations where resources are limited, particularly in remote locations where there is a shortage of qualified healthcare professionals with experience in CXR interpretation, the suggested automatic CAD system will be extremely helpful. For mass screening programmes, the suggested system might also be combined with a mobile vehicle equipped with a digital radio imaging device. The suggested method performs satisfactorily, as evidenced by the obtained classification results (using a calibration dataset in a 10-fold cross-validation setup): ACC=95.52%, F1-Score=95.48%, AUC=0.955 for Stage-I and ACC=85.35%, F1-Score =85.20%, AUC=0.853 for Stage-II. Moreover, it demonstrated improved agreement between the radiologist score and the predicted severity score (Pearson's correlation coefficient r=0.9589) and produced encouraging results for disease localization (average JI=0.82, average DC=0.77). Based on the aforementioned results,

https://doi.org/10.71058/jodac.v9i4002

it can be deduced that lung subdivision using superpixels greatly enhanced localization performance and produced a compact disease boundary, both of which contribute to the accurate grading of disease severity.
Sangole et al [23] proposed method for detecting parasites is expanded upon in this work to include the recognition of parasite species and their life stages. The ability to identify the species of Plasmodium is crucial for administering the correct medication. Life-stage detection helps to measure parasite growth in infected blood. This work proposes a Systematically Applied Mean Filter (SAMF) for removing salt and pepper noise from microscopic thin blood film images. It also proposes a more efficient Otsu algorithm which helps in more accurate segmentation of RBCs from microscopic thin blood film images. In this research, a method has been proposed to diagnose malaria based on computer vision. As a pre-processing stage, the proposed SAMF algorithm has been used to remove impulse noise from the corrupted malaria-infected images. The proposed Otsu method has been used to obtain the binary version of images for cropping blood cells from the complete image. A total of 17 texture and colour features were extracted from these cropped cells and these features were used to train Cubic SVM classifier. Thus, a precise malaria diagnosis system has been developed for detecting Plasmodium parasites, identifying their life stages and species using images of thin blood smears. The system has been tested on the available standard malarial blood smear image databases (CDC [1] and MPIDB [2]). For CDC database, the system recorded classification accuracy of 97.40%, 96.64%, and 96.50% respectively, for detection of infected erythrocytes, species identification, and stages determination and took 0.4378sec to predict one sample. Also, the system recorded 95.8% accuracy for detection of infected erythrocytes and took 3.56sec to predict one sample for the MPIDB database. The proposed precise malaria diagnosis system outperforms detection results obtained by previous research works using similar datasets. The proposed system is helpful for the detection of malaria parasites, which is more efficient and faster as compared to the manual microscopic diagnosis process. The proposed approach helps diagnose malaria in an early stage thereby reducing deaths from infectious transfusions, reducing antimalarial medicine misuse, and optimising resource use in malaria-endemic areas.
Chandra et al. [24] used the deep convolutional neural network (Deep-CNN) multimodal model and transfer learning techniques to build and test a machine learning application for pneumonia detection. Deep-CNNs and Transfer Learning Models are used in the proposed methodology, which have proven to perform exceptionally well in image analysis applications. The model improves its ability to identify significant patterns and features suggestive of pneumonia by employing a multimodal approach that incorporates both contextual information and visual data retrieved from chest radiographs. Furthermore, transfer learning strategies are used to exploit pre-trained models, providing the network with access to information gained from large datasets even in the absence of a large amount of labelled data.According to the experimental results, VIYU, the state-of-the-artrnodel, achieved the highest accuracy and recall scores of 98.08% and 98.91%, respectively. We proposed a state-of-the-art VIYU model algorithm, a Deep CNN with transfer learning model, for efficient pneumonia recognition in chest X-ray images in this study. We used a number of feature extractor models that had already been trained across multiple domains. We used the attention strategy to classify pneumonia and normal after joining feature vectors from previously trained models. While the results for the Guangzhou Women and Children's Medical Centre dataset using the VIYU model by integrating all four models were 98.08%, ResNet152, DenseNet121, ResNet18, and SE-attention all demonstrated 95.03%, 95.35%, 94.87%, and 96.63% accuracy for transfer learning using a single model.
Verma S. et al [25] suggested that impulse oscillometry and clinical presentation can distinguish between wheezing episodes and pneumonia in children aged 3-6 who have respiratory issues. The prediction model is very helpful; it has an accuracy of 86.76%, sensitivity of 86.11%, and specificity of 87.50% when it comes to predicting wheezing episodes. Additionally, the model's area under the curve is 0.8715. Although it requires validation, the scoring system may be useful in diagnosing wheezing episodes. Children between the ages of three and six may occasionally refuse to participate in impulse oscillometry. For impulse oscillometry, there are no available local reference values. As a result, the manufacturers' values were applied.
Jawad Ahmad Dar et al [26] developed an effective Fractional Water Cycle Swarm Optimizer-based Deep Residual Network (Fr-WCSO-based DRN) to identify pulmonary problems utilizing respiratory sound signals. Fractional Calculus (FC) and Water Cycle Swarm Optimizer WCSO are combined to create the new Fr-WCSO. Conversely, WCSO combines the Competitive Swarm Optimizer (CSO) and Water Cycle Algorithm (WCA). Important features required for subsequent processing are successfully retrieved from the respiratory input sound signals after they have been pre-processed. Data augmentation is done using the extracted features to reduce overfitting problems and enhance detection performance overall. Following data augmentation, the suggested Fr-WCSO method is used to choose features. With evaluation metrics such as True Positive Rate (TPR), True Negative Rate (TNR), and testing accuracy, the created approach performed better than expected, with respective values of 0.963 (96.3%), 0.932 (93.2%), and 0.948 (94.8%).

https://doi.org/10.71058/jodac.v9i4002

Table 1: List of Diseases detected with accuracy and methods used.

S. NO.	Author	Methods	Disease	Accuracy
1	**Ashraf A. et al. [3]**	**Optimised version of CNN, Transfer learning algorithm**	**autism**	**81.56%.**
2	**M. Hussain et al. [4]**	**Convolutional Neural Network**	**diabetic retinopathy**	**99.51%**
3	**Almohimeed A. et al. [5]**	**Particle Swarm Optimisation**	**Alzheimer's Disease**	**92.08%**
4	**S. A. Salih et al. [6]**	**Artificial neural networks (ANN)**	**Covid-19**	**99.8%**
5	**K. Khalil et al. [7]**	**Hardware accelerator method , VHDL and Altera 10 GX FPGA**	**Alzheimer's disease**	**89%**
6	**S. Khan et al. [8]**	**Advanced transfer learning-based mechanism using AlexNet combined with Inception-V4**	**Brain haemorrhage**	**94.54%**
7	**N. Pratama et al. [10]**	**Deep learning algorithm**	**pneumonia**	**99.62%**
8	**Sutradhar A. et al [12]**	**Hybrid Artificial Neural Network (3SHANN) and the Three Stage Hybrid Classifier (3SHC)**	**stress and thyroid disease**	**99.29%**
9	**Y. Sneha et al. [15]**	**Convolutional Neural Network,**	**detect brain tumours**	**97.5%**
10	**A.Marouf et al. [17]**	**Naïve Bayes, Decision Tree, Support Vector Machine with Radial basis function (RBF), Logistic Regression, and ensemble classifiers**	**prostate cancer severity**	**99.64%**
11	**Y. Guan et al. [18]**	**Deep learning diagnosis model with block chain technology**	**Ureteropelvic junction obstruction**	**91.77%**
12	**Y. Zhong *et al.* [19]**	**Image segmentation and a convolutional neural network (CNN)**	**malaria**	**97.74%**
13	**Anisha I. et al [21]**	**CAD system**	**lung disorders**	**87.52%.**
14	**Chandra et al. [22]**	**CAD system**	**PD**	**95.52%**
15	**Sangole et al [23]**	**Otsu algorithm**	**detecting parasites**	**97.40%**
16	**Chandra et al. [24]**	**Deep convolutional neural network (Deep-CNN) multimodal model and transfer learning techniques**	**pneumonia detection**	**98.08%**

https://doi.org/10.71058/jodac.v9i4002

In the above table1 we can see that various machine learning techniques are used to diagnose medical diseases like autism, Brain haemorrhage, pneumonia detection, Alzheimer's Disease, lung disorders etc. and these techniques had attend different accuracies in terms of performance and we can see that swarm intelligence had improved the performance of the machine learning algorithm.

2.4 Risk of Bias and Quality Assessment: Every study was assessed using the following criteria, assessing whether the study population was representative and devoid of selective inclusion. Measuring if the machine learning models used were validated on various datasets. Assessing if research included statistical significance and all required performance measures. Verifying any affiliations or funding sources that could introduce bias.The final analysis did not include research that were judged untrustworthy because of severe bias.

3. Results

Machine learning methods, diseases, and performance indicators are used to characterize the systematic review's findings.In all, 26 studies satisfied the requirements for inclusion. The studies' primary sources of data were medical imaging databases, electronic health records (EHRs), and publicly accessible repositories including the NIH and MIMIC-III datasets. The accuracy values reported by the examined studies varied from 70% to 99%, contingent on the model complexity, feature selection, and dataset size.Using histopathological and imaging data, machine learning demonstrated great accuracy in identifying skin, autism, Brain haemorrhage, pneumonia detection, Alzheimer's Disease, lung disorders and breast cancers. Using clinical factors and ECG data, algorithms were able to accurately predict the risk of heart disease. We may observe that the machine learning algorithm's performance was enhanced by swarm intelligence and some of the neurological diseases like autism and Alzheimer's Disease are still required to be diagnose more accurately.

4.Discussion

With increased accuracy and efficiency over conventional diagnostic techniques, machine learning has shown great promise in enhancing illness detection and diagnosis. Better patient outcomes and early detection result from the automated analysis of massive datasets made possible by its incorporation into clinical practice. The availability of data, the interpretability of the model, and ethical issues are still obstacles, though. Numerous research have demonstrated how important it is to remove biases in machine learning algorithms in order to guarantee impartial and equitable clinical judgments. For machine learning to be widely used in healthcare, regulatory frameworks and standardized validation procedures are required.To enhance generalization across various populations, future research should concentrate on creating interpretable AI models, federated learning approaches, and strong cross-validation procedures. Machine learning has the potential to further revolutionize medical diagnostics and promote precision medicine by overcoming these obstacles.

5. Conclusion

This study is shown a systematic review of existing studies on various techniques used to diagnose the medical diseases for health care. Researchers have been suggested various techniques of AI for medical disease diagnosis in health care, and it is observed that swarm intelligence had improved the performance of the machine learning algorithm. Although, Swarm intelligence still requires an extreme inspection to enhance its performance. This paper provides an in-depth analysis of various medical conditions that have been used in various ways to solve medical condition detection problems in the medical field. We made an effort to provide a thorough overview of a range of medical conditions and examined each method in isolation. Finding the gaps in the literature is necessary to carry out the systematic survey. Further work should be accomplished in the future if the research field grows at the appropriate rate. The results of this systematic survey show that a large number of researchers used various techniques to screen for diabetes, brain, heart, lungs diseases etc. More precise diagnosis is still needed for several neurological conditions, such as autism and Alzheimer's disease. As a future direction, we can diagnose diseases using both the various hybridised versions of swarm intelligence and the enhanced Neural Networks with swarm intelligence versions. We thought that this survey would bring these issues to the forefront and that the extensive research would offer a foundational understanding of how different mutation strategies improve the functionality in the healthcare industry.

https://doi.org/10.71058/jodac.v9i4002

Conflicts of Interest
The authors declare that they have no conflicts of interest.

References

[1] M. Junaid, W. H. Bangyal, and J. Ahmad, “A novel bat algorithm using sobol sequence for the initialization of population,” in 2020 IEEE 23rd International Multitopic Conference (INMIC), pp. 1–6, Bahawalpur, Pakistan, 2020.

[2] W. H. Bangyal, A. Hameed, W. Alosaimi, and H. Alyami, “A new initialization approach in particle swarm optimizationfor global optimization problems,” Computational Intelligence and Neuroscience, vol. 2021, Article ID 6628889, 17 pages,2021.

[3] Ashraf A. , Q. Zhao, W. H. Bangyal and M. Iqbal, "Analysis of Brain Imaging Data for the Detection of Early Age Autism Spectrum Disorder Using Transfer Learning Approaches for Internet of Things," in IEEE Transactions on Consumer Electronics, doi: 10.1109/TCE.2023.3328479

[4] M. Hussain, "Exudate Detection: Integrating Retinal-Based Affine Mapping and Design Flow Mechanism to Develop Lightweight Architectures," in IEEE Access, doi: 10.1109/ACCESS.2023.3328386.

[5]A.Almohimeed et al., "Explainable Artificial Intelligence of Multi-level Stacking Ensemble for Detection of Alzheimer’s Disease based on Particle Swarm Optimization and the Sub-scores of Cognitive Biomarkers," in IEEE Access, doi: 10.1109/ACCESS.2023.3328331.

[6] S. A. Salih, S. K. Gharghan, J. F. Mahdi and I. J. Kadhim, "Neural Network based Covid-19 Diagnosis using X-ray and CT Scan Images," 2023 IEEE 13th International Conference on System Engineering and Technology (ICSET), Shah Alam, Malaysia, 2023, pp. 78-83, doi: 10.1109/ICSET59111.2023.10295132.

[7] K. Khalil, M. M. R. Khan Mamun, A. Sherif, M. Elsersy, A. A. -A. Imam and K. Abouzaid, "A Dementia Diagnosis Technique Based on AI and Hardware Acceleration," 2023 IEEE International Conference on Artificial Intelligence, Blockchain, and Internet of Things (AIBThings), Mount Pleasant, MI, USA, 2023, pp. 1-5, doi: 10.1109/AIBThings58340.2023.10292463.

[8] S. Khan, H. Ali and Z. Shah, "Brain Hemorrhage Detection Using Improved AlexNet with Inception-v4," 2023 IEEE International Conference on Artificial Intelligence, Blockchain, and Internet of Things (AIBThings), Mount Pleasant, MI, USA, 2023, pp. 1-5, doi: 10.1109/AIBThings58340.2023.10292461.

[9] A.Zafar, S. Muneeb, M. Amir, A. Jamil and A. A. Hameed, "A Multi-modal Approach to Lung Tumor Detection using Deep Learning," 2023 IEEE International Conference on Artificial Intelligence, Blockchain, and Internet of Things (AIBThings), Mount Pleasant, MI, USA, 2023, pp. 1-6, doi: 10.1109/AIBThings58340.2023.10291022.

[10] N. Pratama, Moses and A. A. S. Gunawan, "A Systematic Literature Review: Diagnosing Pneumonia Through X-Ray Images Using Deep Learning Approach," 2023 International Seminar on Application for Technology of Information and Communication (iSemantic), Semarang, Indonesia, 2023, pp. 1-6, doi: 10.1109/iSemantic59612.2023.10295364.

[11] R. V. Sharan, K. Qian and Y. Yamamoto, "Automated Cough Sound Analysis for Detecting Childhood Pneumonia," in IEEE Journal of Biomedical and Health Informatics, doi: 10.1109/JBHI.2023.3327292.

[12] Sutradhar A. et al., "An Intelligent Thyroid Diagnosis System Utilising Multiple Ensemble and Explainable Algorithms with Medical Supported Attributes," in IEEE Transactions on Artificial Intelligence, doi: 10.1109/TAI.2023.3327981.

[13] M. Mahadevi, M. Marimuthu, I. G. A. Poornima, K. Vijayalakshmi, D. Vidyabharathi and G. S. Priya, "Deep Learning for the Investigation of Lung and Pancreatic Tumors," 2023 7th International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Kirtipur, Nepal, 2023, pp. 669-674, doi: 10.1109/I-SMAC58438.2023.10290585.

https://doi.org/10.71058/jodac.v9i4002

[14] A.N and Y. K. G. S, "Real-time Prostate Cancer Detection via YOLO-Tiny Variants," 2023 7th International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Kirtipur, Nepal, 2023, pp. 657-662

[15] Y. Sneha, Y. M. Roopa, P. Sawant, M. V. Rao, D. L. Padmaja and R. V. S. Lalitha, "Advancements in Brain Tumor Detection using Machine Learning Applications from MRI Image Analysis," 2023 7th International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Kirtipur, Nepal, 2023, pp. 809-814, doi: 10.1109/I-SMAC58438.2023.10290231.

[16] Dhanushyar, C. K, S. S and T. K, "Enhancing Precision Medicine with Machine Learning in Thyroid Diagnosis," 2023 7th International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Kirtipur, Nepal, 2023, pp. 675-680, doi: 10.1109/I-SMAC58438.2023.10290323.

[17] A.Marouf, R. Alhajj, J. G. Rokne, S. Ghose and T. A. Bismar, "Exploring Gene Expression and Clinical Data for Identifying Prostate Cancer Severity Levels using Machine Learning Methods," 2023 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), Regina, SK, Canada, 2023, pp. 186-191, doi: 10.1109/CCECE58730.2023.10288946.

[18] Y. Guan, P. Wen, J. Li, J. Zhang and X. Xie, "Deep Learning Blockchain Integration Framework for Ureteropelvic Junction Obstruction Diagnosis Using Ultrasound Images," in Tsinghua Science and Technology, vol. 29, no. 1, pp. 1-12, February 2024, doi: 10.26599/TST.2022.9010016.

[19] Y. Zhong *et al*., "Efficient Malaria Parasite Detection from Diverse Images of Thick Blood Smears for Cross-Regional Model Accuracy," in *IEEE Open Journal of Engineering in Medicine and Biology*, doi: 10.1109/OJEMB.2023.3328435.

[20] Assefa M. Multi-drug resistant gram-negative bacterial pneumonia: etiology, risk factors, and drug resistance patterns. Pneumonia (Nathan). 2022 May 5;14(1):4. doi: 10.1186/s41479-022-00096-z. PMID: 35509063; PMCID: PMC9069761.

[21] http://hdl.handle.net/10603/427428

[22] http://hdl.handle.net/10603/458305

[23] http://hdl.handle.net/10603/486701

[24] http://hdl.handle.net/10603/458305

[25] Verma S, Kulkarni S, Verma N, Singh BK, Jaiswal J, Predictive regression equation and nomogram of peak expiratory flow rate in healthy school going children of Kolhapur, Maharashtra, India. Sri Lanka Journal of Child Health,2021

[26] Jawad Ahmad Dar, Kamal Kr Srivastava, Alok Mishra, Lung anomaly detection from respiratory sound database (sound signals), Computers in Biology and Medicine, Volume 164, 2023, 107311, ISSN 0010-4825, https://doi.org/10.1016/j.compbiomed.2023.107311.