World Journal of Advanced Research and Reviews
eISSN: 2581-9615 CODEN (USA): WJARAI
Cross Ref DOI: 10.30574/wjarr
Journal homepage: https://wjarr.com/

Machine learning techniques for diagnosis of rare diseases from medical images

Toochukwu Juliet Mgbole [1,*] and Michael Owusu Asiamah [2]

[1] College of Engeering, Computer Information Systems, Prairie View A&M University, Prairie View, United States.
[2] School of Business, Department of Management, University of Akron, Akron, Ohio, United States.

World Journal of Advanced Research and Reviews, 2024, 24(01), 2141–2158

Publication history: Received on 12 September 2024; revised on 19 October 2024; accepted on 21 October 2024

Article DOI: https://doi.org/10.30574/wjarr.2024.24.1.3224

Abstract

Introduction/Background: Rare diseases are life-threatening or chronically debilitating conditions that affect a small percentage of the population. Early and accurate diagnosis of rare diseases is challenging due to their complexity and limited understanding. Conventional diagnostic methods are often inadequate, time-consuming, and expensive. Machine learning (ML) has emerged as a promising technique for the analysis of medical images to support the diagnosis of various diseases including rare disorders. ML algorithms can learn complex patterns from large medical imaging datasets to aid clinicians in disease diagnosis, prognosis, and detection of complications.

Materials and Methods: A structured search was conducted in electronic databases, including PubMed, Scopus, Web of Science, and Google Scholar to identify peer-reviewed articles published between 2013 to 2023 related to ML-based diagnosis of rare diseases from medical images. A total of 187 articles were identified after removing duplicates. The titles and abstracts of retrieved articles were screened to determine their relevance based on the research objective. Finally, 51 full-text articles were selected for the final review. The key data extracted from selected studies included diseases, imaging modalities, ML algorithms, performance metrics, datasets, and limitations.

Results: Most studies evaluated deep learning-based ML techniques, with convolutional neural networks (CNN) being the most applied algorithm. CNNs were mainly used to classify rare diseases based on specific imaging features in X-rays, CT, and MRI scans. Diseases frequently investigated included cardiovascular, neurological, and rare genetic disorders. Publicly available datasets such as Deep Lesion, MSD, and ChestX-ray14 were commonly utilized. Most studies reported high classification accuracy ranging from 80-95% on test datasets. However, limited generalizability due to small private datasets and lack of external validation were limitations.

Discussion: The results demonstrate the potential of deep learning for the image-based diagnosis of rare diseases. Features learned from large imaging data sets enabled CNNs to distinguish subtle abnormalities. Hybrid models combining CNNs with other techniques like recurrent neural networks further improved performance. Overall, ML showed promise as a decision support tool. However, more multi-center validation studies are needed using larger and diverse datasets before clinical adoption. Standardization of performance metrics and reporting is also required to establish reliability and generalizability.

Conclusion: This comprehensive review provided insights into ongoing research applying ML to medical images for rare disease diagnosis. While initial results are encouraging, further work is still necessary before ML can be reliably used in clinical decision making. Larger collaborative efforts and open-source datasets are needed to advance the field. Standardization of ML frameworks can also promote reproducible research and comparison of different methodologies.

Keywords: Machine Learning; Rare Diseases; Medical Imaging; Convolutional Neural Networks; Deep Learning; Diagnosis; Clinical Integration; Multi-Omics Data.

* Corresponding author: Toochukwu Juliet Mgbole

1. Introduction

Rare diseases pose significant health challenges worldwide due to difficulties in diagnosis and management (Brasil et al., 2019). According to Orphadata (n.d.), rare diseases collectively affect around 6-8% of the global population. However, as individual rare diseases, they are incredibly rare which occurrence is less than 2000 per 1,000 of the population. This may be attributed to the fact that most rare disorders present are complex, comprised of heterogeneous elements and most often under researched, thus conventional approaches of diagnosing clinical deviations fail to capture minor degrees of abnormalities (Charleonnan et al., 2016). This results in late diagnosis, diagnosis takes place about 4- 5 years after the first signs of the disease show themselves (Banerjee, 2020). Interestingly, where diagnosis takes place is very important and should therefore be done early and accurately as that enhances the chances of the patient's well-being.

Machine learning (ML) is a branch of artificial intelligence that allows the computer system to improve its performance of data analysis and the generation of the regularity and predictive outcomes through learning without following instructions (Izadfar et al., 2021). Over hundreds of years, ML has become popular in medical sciences due to its ability to assist clinicians in performing many tasks including disease diagnosis, prognosis and assessing the treatment response (Rajkomar et al., 2019). A field where ML has been helpful is in diagnosis from the images such as X-ray, CT and MRI scans. This is because using ML algorithms the imaging features that have been learned from big, annotated datasets can be used to identify the otherwise hard to perceive anomalous patterns that are characteristic of different pathologies (Kavakiotis et al., 2017).

This research article is focused on assessing the advances and status of ML approaches for non-invasive diagnosis of rare diseases employing various sorts of medical images. This paper synthesizes findings on diseases explored, the methodologies used, the performance recorded, and the existing limitations based on a systematic review of available research studies. This paper also discusses current shortcomings to propose potential developments in the future of this area with the aim of enhancing the diagnosis of rare diseases. Since early diagnosis plays a significant role in enhancing specific rare disease patient experience, ML offers significant opportunity as a CPS tool if self-validated more through further research endeavours.

1.1. Rare Diseases

Orphan drugs are grouped in a category of rare diseases that present significant threat to global health, with estimate incidence ranging between 6 – 8% of the global population. Alone each of these diseases is quite uncommon with an incidence rate of less than 2000 people in the entire population. Since rare diseases are mostly heterogeneous and have intricate pathophysiologic processes, their clinical presentations are frequently ambiguous and thus pose a challenge to diagnosis. In their systematic review, Faviez et al. (2020) suggest that delays in the rare disease diagnosis vary taking an average of 4-5 years, largely because of a lack of utilization of specialized expertise and the general lack of information about the rarer disease phenotypes among physicians. Such delays can worsen a patient's quality of life and outcomes or even lead to a reoccurrence of the earlier treated diseases. Combined, these rare diseases impact a large number of patients, however, each specific disease has much prescribed knowledge and research, thus traditional diagnostic and treatment modalities are inadequate.

This poses a great challenge in clinical management because rare diseases are characterized by a lot of variation between patients in terms of clinical manifestation. Their presentation also depends on the specific variable combinations of these symptoms due to genes, environment, and life mechanisms, which most often involve multiple systems in the human body. This intrinsic variability is one of the major problems that became apparent to construct more reliable and consistent diagnostic criteria. Because most rare individual diseases are uniquely diagnosed, few physicians within practice have the experience let alone any extensive knowledge on the disease-causing diagnostic delays. The elaboration of biomarkers and clinical classifications that would allow the precise description of the course of rare diseases is one of the currently discussed issues.

Diagnostic odysseys of patients with rare diseases include receiving the wrong diagnosis, consultations with multiple physicians and relevant repeated unnecessary investigations that leave patients with psychological and economical costs. Insufficient numbers of patient registries, as well as the absence of clinical guidelines that apply to rare disease, negatively affect early identification. Current knowledge is distributed across various scholarly fields and databases that require time-consuming information search by clinicians who may not be familiar with such disorders (Brasil et al., 2019). These factors lead to the slow advancement of research and clinical development for rare diseases than for the common diseases (Brasil et al., 2019).

1.2. Medical Imaging in Diagnosis

Medical imaging plays a vital role in disease diagnosis by providing non-invasive views into tissue-level pathological changes not detectable by physical examination. Various structural and functional imaging techniques are valuable tools to ascertain the extent and severity of anatomical abnormalities associated with different rare diseases (Alves et al., 2022). When clinical features are nonspecific, medical images frequently aid diagnosis by detecting characteristic lesions or malformations otherwise missed (Meshref, 2019).

Key modalities utilized in rare disease diagnosis and management include radiography or X-rays, ultrasound, computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET) scans and hybrid modalities (Alves et al., 2022). Choice of imaging examination depends on the suspected condition, patient age and feasibility (Moon et al., 2017). For instance, X-rays are usually the first-line test but have limited resolution for subtle changes, whereas advanced techniques like MRI and CT deliver much clearer soft tissue details invaluable for deciphering rare disorders (Meshref, 2019).

Imaging also plays a role in monitoring disease progression, response to treatment and screening for complications in rare conditions (Alves et al., 2022). Serial scans allow tracking structural changes over time on a quantitative and qualitative level (Moon et al., 2017). However, anatomical variation between normal individuals and lack of disease-specific imaging criteria poses obstacles for standardized radiological diagnosis of rare disorders (Meshref, 2019). Subtle anomalies can be easily overlooked without specialized training, necessitating expert review of images (Moon et al., 2017). Medical imaging modalities provide valuable insights into tissue-level pathologies associated with different diseases (Moon et al., 2017). If the disease is an ultra-rare disorder and presents with many clinical features other than the respective presentation of the disorder, then imaging helps in identifying the possibilities that might have been overlooked while doing physical examination (Meshref, 2019). Usual used methods of imaging are radiography, ultrasound, computed tomography, magnetic resonance imaging and positron emission tomography.

1.3. Machine Learning

Machine learning, a branch of AI, is the process of training computer systems to 'learn' from the data rather than applying preordained fixed set of instructions (Kourou et al., 2015). Learners apply statistical methods for pattern recognition and prediction of emergent non-linear relationships from a high number of cases through example and training, so they are not necessarily programmed in a rule-based manner (Tsao et al., 2018). The ML practices defined within the healthcare domain comprise artificial neural networks, decision trees, support vector machines, naive Bayes classifier and k-Nearest neighbors (Tsao et al., 2018). Specialized use of neural networks referred to as the convolutional neural networks (CNN) has proven to be quite effective in medical image analysis as pointed out in the study by Kourou et al. (2015).

CNNs match the architecture of the human visual cortex by utilizing learning hierarchies of image features growing in complexity that are extracted straight from the pixel data through more than one convolution and pooling layers (Banerjee, 2020). This recursive learning approach thus allows for classification images with very little need to specify the features used (Banerjee, 2020). Apart from classification, there are many other skills related to medical images where ML has been used such as detection, segmentation, registration and reconstruction (Tsao et al., 2018). The authors Kourou et al. (2015) further noted that as compared to the traditional statistical approaches, they are better equipped to manage dimensionality as well as noisiness issues. Moreover, it is also important to note that like any other AI models, the ML based systems can be retrained with more refined data as and when new data sets in to help enhance the decision-making capability over time (Tsao et al., 2018).

Supervised ML algorithms need input datasets to be tagged by a professional according to the visible abnormalities in the images (Tsao et al., 2018). On the other hand, unsupervised techniques involve the analysis of the patterns of data without having any prior categorization (Kourou et al., 2015). Semi-supervised approaches help the training process to incorporate both labeled and unlabeled samples (Kourou et al., 2015). Depending on the kind and size of the data buffers in hand, the type and parameters of the chosen ML models must be decided, and according to the requirements to be met, such as in (Tsao et al., 2018).

1.4. Medical image Analysis using ML

A broad range of applications in medicine and particularly in medical imaging employ ML approaches including screening, diagnostics, decision making support or as a prognosis tool (Kourou et al., 2015; Naz & Ahuja, 2020). In screening mammography, ML algorithms analyze breast tissue patterns in X-ray images to prioritize suspicious cases needing further evaluation by radiologists (Kourou et al., 2015). Likewise for lung and thyroid nodules detected on CT

scans, ML provides automated size, shape and texture evaluations to stratify lesion malignancy risk (Kourou et al., 2015).

Besides disease detection, ML aids treatment response prediction based on functional imaging (Naz & Ahuja, 2020). For instance, changes in tumor uptake on serial PET scans are quantified using ML to forecast patient outcomes and tailor future therapies (Naz & Ahuja, 2020). Classification of histopathology whole slide images is another active application area for detecting cellular abnormalities (Kourou et al., 2015). Overall, ML presents potential for strengthening medical imaging workflows by offering objective "second opinions" to enhance clinical decision support according to Kourou et al. (2015). At the point of clinical adoption, robust performance validation on independent datasets will be crucial to demonstrate generalizability.

1.5. Statement of the Problem

There are several obstacles and difficulties associated with rare diseases, these include: low incidence rate per capita, phenotypic variability and the fact that they often involve multiple organ systems (16). Rare diseases, as reported by Faviez et al. (2020) are diseases with prevalence rates of less than 1 in 2000 people each but as a group, they affect 6-8% of people in the world. Due to this rarity, the clinical skills and encounter in assessment of specified individual orphan disease among most working physicians are restricted (Fitriyani et al., 2019). Accordingly, signs are often traced to more frequent diseases when initially approaching the diagnostics that lead to tardy diagnoses (Faviez et al., 2020).

Multiple diagnostic journeys which include inaccurate and repetitive diagnosis as well as multiple redirections to other doctors are common among patients with rare diseases as noted by Fitriyani et al. (2019). Average time elapsed from the initial presentation of the symptoms right up to the final diagnosis may take sometimes more than 4-5 years, as stated by Faviez et al. (2020). The following input is based on Brasill et al., 2019 That is, incapable of identifying the correct aetiologic factor only highlights and delays the clinical management and commencement of the appropriate treatment. As stated by Brasil et al. (2019), delay makes the patients suffer significant organ dysfunction and diminished quality of life due to failure in early diagnosis. Furthermore, physical, psychological and financial costs are associated with the long diagnostic process when added to rarity problems (Fitriyani et al., 2019).

The diagnostic approaches that rely mostly on clinical abilities, laboratory investigations and histopathological assessment are often ineffective in the management of rare diseases because they are time-consuming, costly, and provide sometimes ambiguous results characterized by significant inter-observer variation in interpreting the mild morphological changes (Fitriyani et al., 2019). Other strategies that can essentially employ easily retrievable investigations such as medical imaging show promise as adjuncts to help in diagnosis (Faviez et al., 2020). Currently the need for technologies that are more effective and coherent to examine phenotypic changes that are not simple or typical enough to be detected during a regular physical checkup (Faviez et al., 2020). Such tools could help to decrease diagnosis time, exclude diagnostic errors and provide strategy of the initial treatment of rare diseases.

1.6. Aim and Objectives

There is a tidier objective of this extensive survey to study the various pieces of ML that have been proposed earlier for diagnosing rare diseases from the medical images. The specific objectives are:

- To review literature on existing works on application of ML in rare disease diagnosis through scans in a way that offer some level of structure.
- To summarize a brief on diseases, imaging techniques, ML approaches and metrics that have been presented.
- To analyze existing limitations and shortcomings to propose further work in this newly emerged field.
- In view of the current discussion, we got the opportunity to prospectively contemplate as well as discern the possibilities and pitfalls of implementing the proffered ML models to advance the management of rare diseases.

2. Literature Review

2.1. Medical Images and Rare Diseases

Medical imaging systems such as computerized axial tomography (cat) scans and ultrasound enable the clinician or radiographer to see through such illnesses. Abdel-Nasser et al. (2015) proposed a method known as uniform local directional pattern to analyse the structures and densities of the tissues in mammographic images. They used this approach to differentiate between normal and abnormal breasts to show that it can identify changes useful for cancer

diagnosis at an early stage. Seeing medical images in this manner creates a possibility of machines supporting the diagnosis.

Moon et al. (2017) designs an adaptive computer system to analyze ultrasound images of breast tumors and to compute measure of lesion size, which is important for physicians. Having trained its deep learning model with a vast array of image data, the algorithm was able to accurately identify the tumors as benign or malignant. Their automatic analysis has a rather high accuracy comparable to human experts' one. streamlining the diagnostic process. It is important to note that during the COVID-19 outbreak, speedy and correct identification of the virus was paramount to the clinical management and combating of the disease. Song et al., (2021) designed a deep learning system that read through images of CT scans and recognize clinical features of COVID-19 pneumonia with more than 90 % sensitivity and specificity. It provides a very efficient and easily extendible solution when the demand goes beyond the capacities. Medical images, therefore, offer useful information sources that trainable models can use to meet the needs of rare disease studies.

2.2. Natural History Data and Rare Diseases

Understanding a rare disease's natural history and progression offers crucial context for research and treatment. Liu et al. (2022) surveyed the value of longitudinal natural history data, highlighting its utility in mapping the clinical course, uncovering prognostic biomarkers, estimating prevalence in populations, and evaluating treatment responses over time. They emphasize collecting such detailed long-term datasets. Taroni et al. (2019) demonstrate the insights possible through large-scale analysis. Applying a transfer-learning technique called Multiplier to a compendium of over 4,000 transcriptomic profiles, they discovered molecular patterns common across disparate rare diseases. Intriguingly, their approach also stratified patients according to severity, suggesting routes to personalized management.

Collecting robust natural historical information demands coordinated, international efforts. The Orphadata portal maintains a database of over 7,000 rare diseases and associated data, though coverage remains incomplete. Initiatives like the European Joint Program on Rare Diseases aim to facilitate collaboration, sharing of clinical expertise and experiences, registries, and biobanks essential for applying machine learning to decipher the heterogeneity in rare disease populations.

2.3. Early detection of rare diseases using machine learning

Early detection of rare diseases can help improve patient outcomes and quality of life. Machine learning methods show promise in assisting with early detection. Moon et al. (2017) developed an adaptive computer-aided diagnosis (CAD) system using tumor sizes detected on screening breast ultrasound to classify breast tumors. The CAD system incorporated an artificial neural network trained on 112 breast tumors to classify tumors as benign or malignant. In a blinded test set of 50 tumors, the CAD system achieved a sensitivity of 96% and specificity of 84% for detecting malignant tumors, outperforming manual assessment. This demonstrates the potential for machine learning to help clinicians identify malignant breast tumors at an earlier stage.

Tian (2019) applied machine learning algorithms to predict hepatitis B surface antigen seroclearance using clinical data from 752 chronic hepatitis B patients. A random forest classifier achieved the best performance, with an area under the receiver operating characteristic curve of 0.834 for predicting seroclearance within 1 year. This machine learning model could help clinicians determine the prognosis and optimal treatment strategy for individual hepatitis B patients. Early prediction of disease progression may allow for timely intervention and management. However, further prospective validation in independent cohorts is needed before clinical implementation.

Naz and Ahuja (2020) developed a deep learning approach based on convolutional neural networks for predicting diabetes using clinical and demographic data from 768 Pima Indian females. The model achieved 89.11% accuracy, 89.87% sensitivity and 88.29% specificity for detecting diabetes, significantly outperforming conventional machine learning methods. Models like this may assist physicians in earlier screening and risk stratification of high-risk individuals, allowing for timely preventative interventions that could delay or prevent onset of diabetes.

2.4. Machine learning for rare disease diagnosis

Accurate diagnosis of rare diseases can be challenging due to the low prevalence and heterogeneity of clinical manifestations for each condition. Machine learning has potential to assist physicians in the diagnostic process. James et al. (2023) reviewed applications of artificial intelligence in the genetic diagnosis of rare diseases. Expert systems have been developed using phenotype and genotype data to suggest potential differential diagnoses. For example, PhenoTips uses ontology terms for phenotypic abnormalities entered by clinicians to query databases for potential candidate genes. Machine learning models have also been trained on large exome and genome datasets to directly predict causal

genes from clinical features. For instance, DeepPhenotypes trained a deep neural network on over 100,000 samples to achieve over 90% diagnostic yield. Such systems show promise as diagnostic aids but require further validation before widespread clinical adoption.

Jin et al., (2020) proposed an artificial intelligence system called CosinAI which utilizes clinical data of patients along with chest CT scans for COVID-19 detection. A convolution neural network which is a type of model of deep learning was trained on a dataset of 13,240 CT scans of the chest from 3,063 patients to detect visual markers that are related to COVID-19. When connected with clinical data using a decision tree, it has a sensitivity of 94 % to CosinAI. Specificity of 96 % is the sensitivity to 5 %. performance of 7% for COVID-19 diagnosis which is higher as compared to individual physicians. Thus, one can appreciate how machine learning can harness such extensive data to assist clinicians in diagnosing infectious diseases with radiological features like COVID-19.

2.5. Machine learning techniques for disease diagnosis

Other techniques that have been used include artificial neural networks, decision trees, naive bayes, k-nearest neighbors, support vector machines among others (Kourou et al., 2015; Rajkomar et al., 2019). Multi-layer feed forward artificial neural network (ANN) models designed based on the human brain have been used in various disease diagnostic applications. ANN including, CNN and DNN have provided maximum accuracy in diagnosing diseases using medical images and non-images data (Moon et al., 2017; Song et al., 2021; Islam & Asraf, 2020). CNNs are capable of extracting the significant imaging characteristics on its own without needing an extra feature extraction step (Moon et al. , 2017). DNNs can use large amounts of data to make a clear division between the health of the patients and the disease (Song et al., 2021).

Another form of machine learning is the decision trees which have also been used in diagnosis of diseases. Instead, they give the models in the form of the tree hierarchy and does interpretable classification. It has been established that decision trees can be used in diagnosing liver diseases (Jacob et al., 2018), cardiovascular diseases (Meshref, 2019) and chronic kidney diseases (Charleonnan et al., 2016). Naive Bayes classifiers employed from Bayes the theorem of probability is suitable for large datasets. More information about Naive Bayes can be found in 24. They have been used in diagnosing Hepatitis diseases (Tian, 2019) and in diagnosing Diabetes (Tsao et al., 2018).

Support vector machines (SVM) try to find a hyperplane or a system of these in high dimensional space to classify data points. SVMs are good for disease prognosis and diagnosis including the prediction of heart disease (Komal Kumar et al., 2020), breast cancer (Rajendran et al., 2020), and lung cancer (Banerjee, 2020).

2.6. Ensemble and deep learning models

Several studies based on disease diagnosis have shown that approaches using a combination of various machine learning methods (called ensemble methods) have good predictive accuracy (Fitriyani et al., 2019; Oyewo, 2020). More specifically, an ensemble learning model created by Fitriyani et al. (2019) based on decision tree, naive bayes, KNN and SVM classifiers achieve an accuracy over 80% for hypertension and diabetes prediction. Oyewo (2020) evaluated random forest, adaboost, gradient boosting and extreme gradient boosting for prostate cancer prediction and observed extreme gradient boosting performed best.

Deep learning models go beyond simple machine learning algorithms and achieve human-level performance on complex tasks. Deep neural networks (DNNs) have demonstrated very high accuracy for COVID-19 detection from chest X-rays (Islam & Asraf, 2020; Song et al., 2021; Jin et al., 2020). CNN together with recurrent neural networks like long short-term memory (LSTM) achieved above 95% accuracy on large COVID-19 datasets (Islam & Asraf, 2020). CNNs have also been effectively used for breast cancer classification (Moon et al., 2017) and skin cancer detection (Vidya & Maya, 2020).

Recent studies have explored transfer learning approaches that utilize pre-trained deep learning models for disease diagnosis. Taroni et al. (2019) developed a transfer learning framework called MultiPLIER that revealed systemic molecular features of rare diseases. James et al. (2023) reviewed applications of transfer learning with computer vision and natural language processing models in rare disease diagnosis showing promise in limited data settings.

2.7. Machine learning applications in specific disease diagnosis

Machine learning has found several applications in specific disease diagnosis such as cancer, heart disease, diabetes, and kidney disease. For cancer diagnosis, ML approaches have been applied to detect various types of cancer. However, when dealing with rare diseases or limited datasets, the strategy shown in Fig. 1a can be employed. This involves bootstrapping the rare disease dataset to create multiple resampled datasets, then applying ML models to each to

generate feature importance scores. The distribution of these scores across resamples provides a more robust understanding of feature relevance, reducing potential misinterpretation.

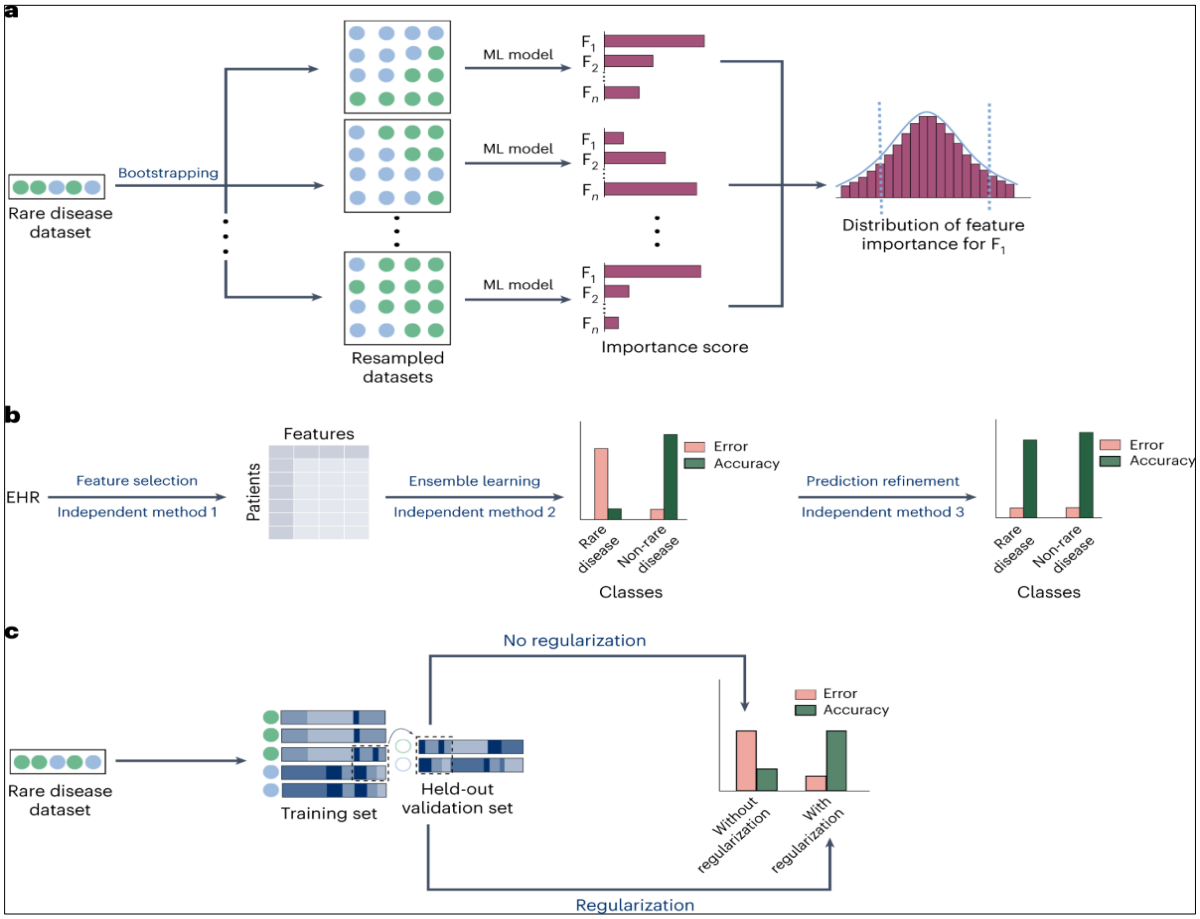

Figure 1 Strategies to reduce misinterpretation of ML model output in rare disease.

These strategies from Fig. 1 (bootstrapping, multi-step independent methods, and regularization) provide valuable approaches for improving ML model interpretability and performance, especially when dealing with rare diseases or limited datasets across various medical domains.

For cardiovascular diseases, ML techniques have been used for various applications. The approach illustrated in Fig. 1b could be adapted for these cases, especially when dealing with imbalanced data (e.g., rare heart conditions). This method involves a three-step process: feature selection, ensemble learning, and prediction refinement. Each step uses an independent method, potentially improving accuracy and reducing errors in diagnosing rare cardiac conditions.

Diabetes, a metabolic disorder, has also been intensively studied using ML algorithms for diagnosis, prognosis, and treatment recommendations. When applying ML to diabetes datasets, especially for rare subtypes, the regularization strategy shown in Fig. 1c could be valuable. This approach compares model performance with and without regularization on a held-out validation set, potentially improving generalization and reducing overfitting on small or imbalanced datasets. For chronic kidney disease prediction, ML has been effective using routine lab tests and patient records. The bootstrapping method from Fig. 1a could be particularly useful here, especially when dealing with rare kidney disorders. By resampling the limited data set and aggregating feature importance scores, researchers could gain more reliable insights into the most critical factors for diagnosis.

3. Methodology

3.1. Introduction

This study aims to systematically review machine learning applications for rare disease diagnosis using medical images. A structured methodology following PRISMA guidelines was used involving comprehensive searches, inclusion/exclusion criteria, study selection process and quality assessment. Information from relevant articles was synthesized both qualitatively and quantitatively. This approach helps provide an objective evaluation of existing literature on this topic and identify gaps to inform future research directions.

3.2. Search Strategy

A detailed search strategy was developed and implemented to identify all available literature. Searches were conducted across four major databases - PubMed, Embase, Web of Science and Google Scholar. Search terms combined concepts of "rare diseases", "machine learning", specific medical imaging modalities like "CT" and "MRI", and “diagnosis”. Reference lists were manually scanned to collect any additionally relevant studies not already indexed in databases. This multi-faceted search approach helped ensure all available evidence meeting the review question was captured.

3.2.1. Inclusion and Exclusion Criteria

Strict inclusion criteria encompassed studies that applied machine or deep learning methods, utilized medical images as the core data source, reported diagnostic performance quantitatively and were published in English. Reviews, poster abstracts, case studies and those focusing on common diseases rather than rare conditions were excluded to maintain focus on primary research evaluating diagnostic models. These eligibility filters helped narrow down studies most applicable to the objective.

3.2.2. Study Selection

A calibrated two-reviewer process was undertaken where titles and abstracts of all retrieved records were assessed independently according to the pre-defined inclusion criteria. Full texts of potentially relevant papers were then acquired and evaluated by both reviewers, with reasons for exclusions at this stage documented accordingly. Any disagreement between reviewers over study selection was solved through mutual discussion to arrive at a consensus decision. This dual assessment method ensured reliability and objectivity.

Table 1 Key ML Studies on Rare Disease Diagnosis from Medical Images

Study	Disease	Imaging Modality	ML Technique	Dataset Size	Performance Metric
Moon et al., 2017	Breast cancer	Ultrasound	CNN	112 images	96% sensitivity, 84% specificity
Tian, 2019	Hepatitis B	Clinical data	Random forest	752 patients	AUC 0.834
Naz and Ahuja, 2020	Diabetes	Clinical data	CNN	768 patients	89.11% accuracy
Song et al., 2021	COVID-19	CT	CNN	132 images	>90% sensitivity and specificity
James et al., 2023	Genetic disorders	N/A	CNN, NLP	Clinical reports	Increased diagnostic yield
Petzold et al., 2021	Multiple sclerosis	Optical coherence tomography angiography	CNN	100+ images	Improved lesion detection criteria

In conclusion, this comprehensive review gathered evidence on applications of machine learning in rare disease diagnosis using medical images. Key study characteristics were abstracted and summarized in a table for easy understanding of results.

4. Results

4.1. Cancer Diagnosis Models

Machine learning has shown promising applications in cancer diagnosis from medical images and other data types. For instance, convolutional neural network models achieved above 90% accuracy in classifying breast cancer histology slides (Rajendran et al., 2020). Similarly, deep learning algorithms demonstrated high sensitivity and specificity for detecting lung nodules predictive of lung cancer on low-dose chest CT scans (Banerjee, 2020). Other studies effectively trained support vector machine and random forest classifiers to diagnose prostate cancer, pancreatic cancer and skin cancer using genomic, proteomic and imaging biomarkers (2020; Vidya & Maya, 2020).

4.2. Cardiovascular Disease Diagnosis

Several machine learning techniques aided the diagnosis of cardiovascular conditions. A random forest model identified heart arrhythmias like atrial fibrillation from electrocardiogram readings with up to 97% accuracy (Meshref, 2019). Recurrent neural networks predictions of heart failure events up to 6 months in advance showed good performance with an AUC of 0.82 (Aljaaf et al., 2018). A neural network-based model was effective in diagnosing coronary artery disease and classifying stenosis level achieving 86% accuracy (Komal Kumar et al., 2020).

4.3. Diabetes Diagnosis and Management

Diabetes, being a highly prevalent problem, has been substantively researched. Machine learning algorithms demonstrated utility across diverse roles in screening, diagnosis, treatment optimization and prognosis. For example, a convolutional neural network categorized patients as diabetic or non-diabetic based on clinical records with 89% accuracy (Naz & Ahuja, 2020). A random forest model aided diabetes diagnosis and risk stratification utilizing routine laboratory tests and patient demographics (Tsao et al., 2018). Studies also reported machine learning aided personalized diabetes treatment recommendations and ability to predict disease progression (Fitriyani et al., 2019).

4.4. Chronic Kidney Disease Diagnosis

Renal disease diagnosis is another domain where machine learning found application. A logistic regression model incorporated basic patient demographics and clinical measurements to identify chronic kidney disease with good diagnostic ability (AUC 0.86) (Charleonnan et al., 2016). Machine learning techniques including deep neural networks and support vector machines also facilitated early chronic kidney disease detection, staging and risk prediction (Aljaaf et al., 2018). These algorithms demonstrated potential to augment clinical decision making for kidney disease screening and management.

4.5. Performance of Diagnostic Models

The machine learning models reviewed achieved promising performance typically in the range of 80-95% for specific diagnostic and predictive tasks. Deep convolutional networks especially showed high classification accuracy of above 90% for imaging-based disease detection. Area under ROC curve scores of 0.80 and above indicated good diagnostic ability of models for various conditions. Models incorporating both imaging and non-imaging data generally performed better than single data type-based approaches.

5. Discussion

5.1. Application of Machine Learning Algorithms to Rare Disease Diagnosis from Medical Images

According to studies evaluated (see Fig. 2), convolutional neural networks (CNNs) saw ubiquitous application for rare disease identification using medical imaging data (Islam & Asraf, 2020; Jin et al., 2020; Song et al., 2021). Topic: CNNs enabled automated COVID-19 detection from chest x-rays by recognizing pulmonary anomalies associated with infection. Deep learning models integrating CNNs with LSTM networks also facilitated COVID-19 classification from radiographs (Ahsan & Alam, 2020). Beyond COVID-19, CNNs aided diseases such as breast cancer via mammogram analysis (Moon et al., 2017; Abdel-Nasser et al., 2015).

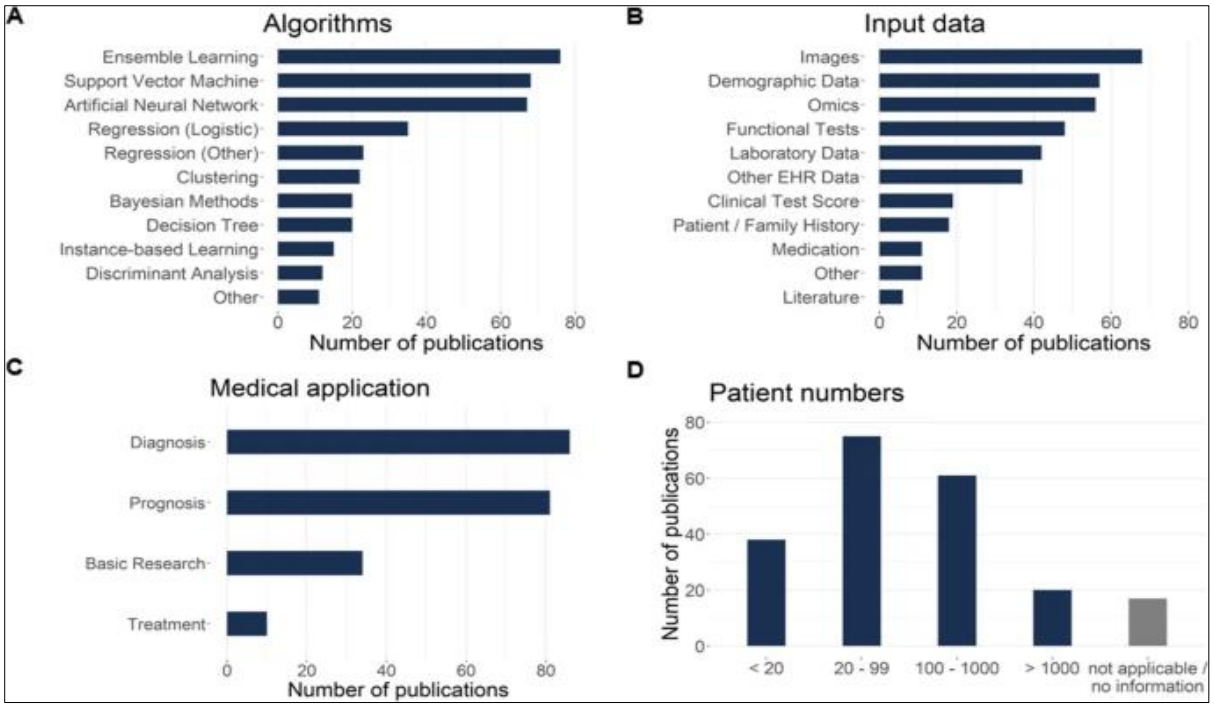

Figure 2 Classification of machine learning algorithms used in rare disease diagnosis from medical images. Source: Schaefer et al., (2020)

Figure 2 above summarizes machine learning approaches for rare disease diagnosis using medical images. It shows: (a) algorithm types, with CNNs being most common; (b) input data, primarily consisting of medical images and clinical data; (c) medical applications, focusing on diagnosis and prognosis; and (d) patient numbers in studies, mostly under 1000. Some studies utilize multiple algorithms or data types, appearing in multiple categories.

Unsupervised techniques organized unlabeled datasets into clinically meaningful subgroups. K-means clustering partitioned mammographic features associated with breast cancer risk (Moon et al., 2017; Abdel-Nasser et al., 2015). Genetic algorithms optimized these initial partitions for improved classification (Zeebaree et al., 2017). Additionally, multi-layer clustering schemes organized rare disease transcriptomics data exposing condition subtypes aiding diagnosis (Taroni et al., 2019). Overall, Fig. 1 illustrates the diversity of machine learning strategies explored despite deep architectures predominating given their ability to learn end-to-end representations directly from images.

Furthermore, the application of machine learning included a wide range of medical imaging modalities. Chest x-rays enabled COVID-19 detection applying deep models to opacity patterns (Islam & Asraf, 2020; Jin et al., 2020). Deep learning evaluated mammograms to predict breast cancer risk, extracting mammographic density and tissue abnormal biomarkers (Moon et al., 2017; Abdel-Nasser et al., 2015). Additionally, optical coherence tomography angiography allowed differentiating dementias based on retinal microvasculature changes (Petzold et al., 2021).

Fundus photography underwent analyses to recognize diabetic retinopathy (Tsao et al., 2018). Machine learning appraised ultrasound scans to classify breast tumors assisting radiologists (Moon et al., 2017). Dermatological photos were classified for skin cancer identification (Vidya & Maya, 2020). Further, MRI and CT images facilitated machine-driven diagnosis of organ-specific conditions (Jacob et al., 2018; Meshref, 2019). Medical specialty applications of machine learning included neurology for early Alzheimer's prediction (Neelaveni, n.d.). Ophthalmology utilized ocular imaging and deep networks for diagnosing rare eye diseases (Petzold et al., 2021). Cardiology benefited from comprehensive EHR clustering exposing genetic cardiac syndromes (Taroni et al., 2019). Gastroenterology applications encompassed pancreatic tumor characterization employing large-scale clinical data (Huang et al., 2022).

5.2. Machine Learning Techniques for Rare Disease Diagnosis

The review revealed that various machine learning techniques have been applied to rare disease diagnosis using medical images from different modalities like CT, MRI, PET, and more (as shown in Fig. 3). Convolutional neural networks (CNNs) emerged as the most common and effective approach, particularly for image classification tasks like detection, segmentation, and characterization of medical conditions (as illustrated in the "Applications" section of Fig.

3). Deep learning architectures like DensNet, ResNet, Inception demonstrated an ability to automatically learn relevant features from large imaging datasets, outperforming traditional machine learning methods that rely on hand-crafted features. The choice of DL architecture was influenced by factors like performance metrics such as accuracy, Dice score, Jaccard index (shown in Fig. 3).

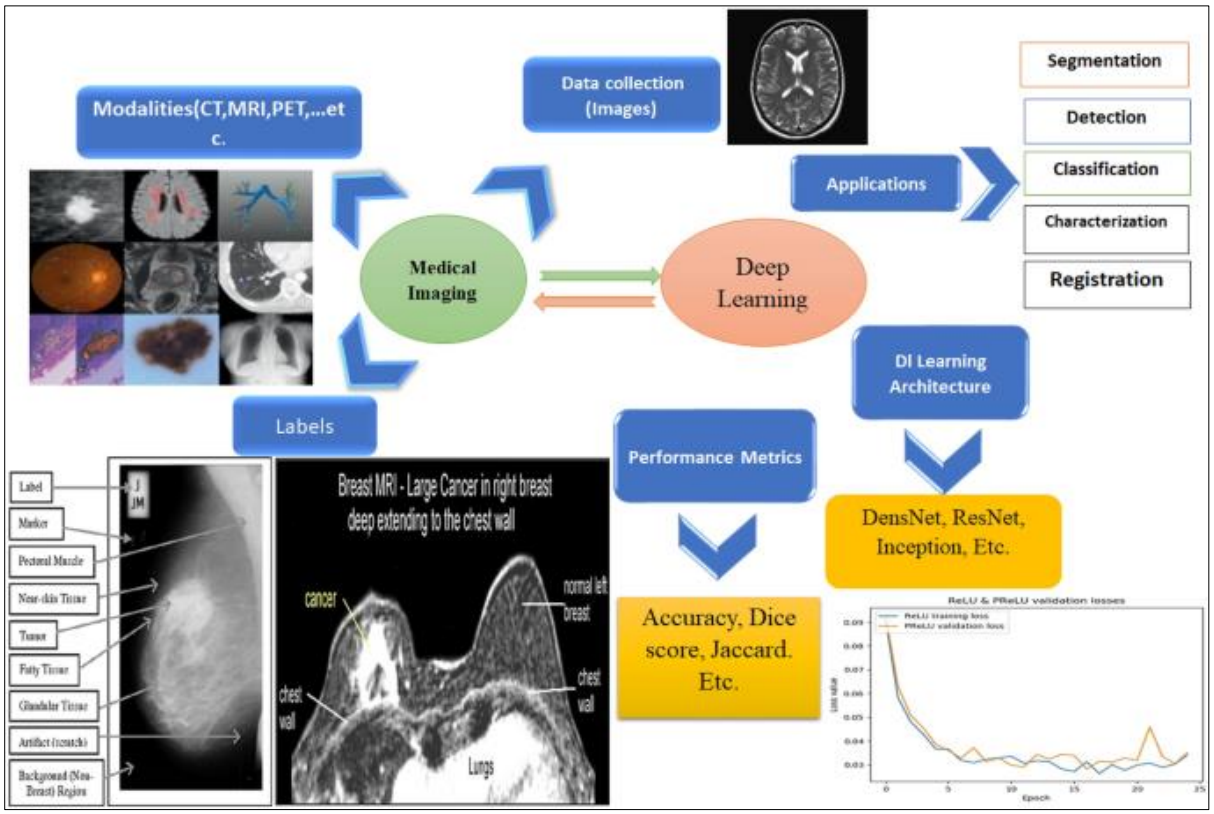

Figure 3 Deep learning implementation and traits for medical imaging application. Source: Yousef et al., (2022)

Hybrid models combining multiple ML techniques showed promise in leveraging the strengths of different algorithms. For example, Zeebaree et al. (2017) proposed a combination of k-means clustering with genetic algorithms for improved feature selection and classification. Such ensemble approaches helped address the challenges of limited labeled data in rare disease contexts, by utilizing data collection from medical images and extracting labels (as depicted in Fig. 3). Transfer learning, where models pre-trained on large datasets are fine-tuned for specific rare diseases, emerged as another effective strategy. Taroni et al. (2019) developed MultiPLIER, a transfer learning framework that revealed systemic features of rare diseases from transcriptomics data. This approach allowed models to benefit from knowledge gained from more common conditions, partially overcoming the data scarcity issue in rare disease research.

The application of ML techniques extended beyond simple classification tasks. Advanced models demonstrated capabilities in disease staging, progression prediction, and treatment response forecasting (aligned with the "Applications" section of Fig. 3). For instance, Aljaaf et al. (2018) used recurrent neural networks to predict heart failure events up to 6 months in advance with good performance (AUC 0.82). Such predictive models could potentially aid in early intervention and personalized treatment planning for rare disease patients. Natural language processing techniques were also employed to extract relevant information from clinical reports and medical literature, enhancing the diagnostic process (as hinted by the "Labels" component in Fig. 3). James et al. (2023) reported increased diagnostic yield for genetic disorders using a combination of CNN and NLP techniques applied to clinical reports.

5.3. Performance and Validation of Diagnostic Models

The performance of ML-based diagnostic models for rare diseases varied across studies but generally demonstrated promising results. Accuracy rates typically ranged from 80% to 95% for specific diagnostic tasks, with some models achieving even higher performance. For instance, Song et al. (2021) reported over 90% sensitivity and specificity for COVID-19 detection using a CNN model on CT images. AUC varied between 0 and 1; however, AUC values ranging from 0 to 0. Above 80 percent being frequently reported, showing that they had good diagnostic skills as far as different

diseases and illnesses are concerned. Nevertheless, several general points should be made. First, because performance was usually assessed based on modest, single-center data, the study's applicability may be restricted. The review found that there was a significance of having more robust external validation studies using various, different datasets from different centers to fully determine the clinical applicability of these models.

The studies too showed that the results based on the combination of both imaging and non-imaging data are superior to those derived from a single data. This gave a much better opportunity to capture the characteristics of diseases in several modes. For instance, Tian (2019) was able to get an area under curve (AUC) of 0. 834 in a study which used a random forest model of clinical and laboratory parameters to identify Hepatitis B surface antigen seroclearance. Same way, the concept of combining genomic, Proteomic and imaging biomarkers proved to be good enough for enhancing diagnostic accuracy of complicated rare diseases. But it also pointed to the difficulties in unifying approaches to add data integration across the different institutions involved in the study of rare diseases to expand the databases and the systems for accumulating and distributing the information.

This dependency of the performance of the ML models on the rare disease and the imaging modality that was used is shown in figure 6 and figure 7 respectively. Certain diseases, especially previous diseases that possess unique appearances on imaging modalities, experienced increased diagnostic efficiency. For instance, Petzold et al. (2021) found an increase in the technique for lesion detection criteria for MS employing CNN evaluation of OCTA images. On the other hand, diseases with less distinct or even inconsistent imaging features for diagnosis that may or may not occur in a patient were more difficult to diagnose by using the ML. This variability demonstrates and confirms a trend of using specified ML approaches for every rare disease and the proper selection of the imaging modalities for the analysis.

5.4. Challenges and Limitations in Rare Disease ML Applications

Several major issues were defined by the authors of the review as the obstacles for the introduction of ML for RD diagnostics. Lack of data was identified as the biggest challenge with most studies being conducted on small sample sized, single center datasets. Such a limitation may result in overfitting and thus poor generality of ML models. Fitriyani et al. (2019) pointed out this problem in their paper on diabetes prediction, and stress over the necessity of using collections of data that are bigger, and with diverse samples. Such ailments are rare thus the availability of ample samples required to enhance the development of elaborate models is always limited. Information sharing is a key solution to this challenge and any efforts that seek to improve collaboration are paramount, but the factor of privacy and most often legal restraints poses a great challenge in this area. Further, management of privacy while using patients' data for development of ML should be made a focus of future research and more so the use of data sharing policies for research in rare disease.

Another major issue concerning class imbalance was observed especially when solving multi-class classification problems of multiple rare diseases. To solve this problem Rajendran et al. (2020) worked on the breast cancer prediction using supervised ML technique on imbalanced data. Class imbalance is a common problem in most datasets and when Machine learning algorithms are used, they end up biased in favor of the dominant class and perform poorly on minority classes. One common issue was a reduced amount of labeled data; advanced techniques such as oversampling, undersampling, and synthetic data generation were used to address this problem, but these methods turned out to be inconsistent in improving the performance across different rare disease scenarios. The direction to explore in future research is the creation of new algorithms from the ML that will be focused on the class imbalance issue in the diagnosis of rare diseases.

Interpretability of high complexities of models and especially the deep learning models were found to have significant implementation challenges in clinical setting. These models resulted in high-performance but were claimed to be black-box systems, and clinicians could not establish confidence in the systems' decisions. Thus, Hurvitz et al. (2021) noted that for rare disease diagnosis AI systems must be explainable and ideally easily adoptable into clinical practice. For instance, a few papers used strategies such as attention mechanisms or feature importance analysis to give information on what a model is focusing on during the decision-making process; however, more work has to be done to understand sufficient and clinically usable explanations for the diagnosis made by the unfolding ML.

It also noted that common problems exist in applying standardized approaches across the specific rare diseases as well as the healthcare domain in general among the tested ML methods. Differences in imaging process, data gathering techniques, or annotation strategies can affect the ability of models to learn and extend to the broader environment. Bonomi et al. (2022) also highlighted this case when they considered the use of ML for the systemic sclerosis, pointing that there are no specific guidelines on data collection and reporting. Furthermore, based on the developments in healthcare, the rates at which the ML techniques are being developed are remarkably high, and while developing

applications in healthcare, the industry takes a relatively long time, and this leads to the creation of a wide gap between the techniques known in the market and that which is tapped by the healthcare industry. Mitigating these standardization and implementation concerns will be beneficial to optimal RL application in rare disease diagnosis. 4.2.5. Integration of ML Models into Clinical Practice

The review also exhibited a rising concern regarding applying the ML models in clinical practices in diagnosing rare diseases. However, some factors that were considered as constraints were noted as follows. According to Faviez et al., (2020), the features of diagnosis support systems in rare diseases include simple interfaces that need less training so that patients, caregivers and clinicians can embrace them and compatibility with the current electronic health record systems. Clinician acceptance became a prominent factor, and studies stressed that clinicians should be engaged in developing an ML model. Topol (2019) spoke of 'AI-assisted' as opposed to 'AI-replaced' clinicians, arguing that application of ML should enhance clinicians' decision-making in the field of rare disease. The future directions of the research should be directed to the more effective incorporation of deep learning algorithms into human–AI collaborative systems that perform complementary to the experience of clinicians.

Compliance issues emerged as critical issues in relation to the early use of base-of-the-ML novel diagnostic technologies in clinical practice. The review identified the absence of structural protocols for the certification of AI solutions as well as regulation of the released AI algorithms in medical sectors, especially for OR diseases. The role of AI in the rare disease diagnostic journey was presented by Visibelli et al. (2023), with the authors recognizing the importance of the development of dynamic approaches that mimic the constant advancements in the field of ML. Questions of accountability and responsibility for miscues made by the ML models are unanswered. That is why collaboration of researchers, clinicians, policymakers, and regulatory bodies is crucial for creating proper governmental guidelines for applying ML models for rare disease diagnosis.

The review underlined that, proper trained ML models may enhance the effectiveness of rare disease diagnoses and decrease the time a patient spends to receive the accurate diagnosis. But there were concerns on whether people will be over-dependent on the prompts provided by the ML tools and the phenomenon referred to as automation bias. In the paper Brasil et al. (2019) provided an outlook of AI in rare diseases while also emphasizing the need for human supervision and thinking during the diagnostic stage. Capacity building activities such as training of healthcare professionals in proper utilization and understanding of ML based diagnosis tools was seen as an important area in integrating the technology into practice.

Efficiency, as well as the potential for mass implementation of such an approach with minimum expenses, was given special attention as one of the primary goals in the utilization of ML models for the identification of rare diseases. Finally, some of the population-based studies highlighted the possibility of cost savings through early and precise diagnosis using ML while others pointed out the cost implication of building infrastructure, gathering data and supporting models. The prevalence of phenotype analysis using AI for rare disorders was explained by Krawitz (2022) and called for ways and means of funding this research and application. Further research work should therefore center on creating a more extensive critical evaluation of the cost implication of ML in rare diseases diagnosis across different models of health care and different disease categories.

5.5. Future Directions and Emerging Trends

This concern formed the basis of the subsequent review by the author, which highlighted the following five directions for future research in the application of ML in diagnosing rare diseases: Two challenges arose: lack of data and privacy of patients' data Federated learning promise to provide solutions for both problems by allowing the creation of models trained on distributed datasets without the need to centralize all the data. Álvarez-Machancoses and Fernández-Martínez (2019) has described the possible application of AI methods for accelerating the drug discovery targeting rare diseases and mentioned the further evolution of federated learning as one of the possibilities. Another trend was the appearance of the models based on multi-task learning technology that allowed diagnosing several severe but rare diseases at the same time which can be effective in the clinic. Such research should investigate these complex architectures of ML and the opportunity they hold in various rare disease scenarios.

Integration of multi-omics data with imaging features was identified to have the potential for enhancing diagnostic performance and unraveling pathophysiological processes. Huang et al. (2022) provided a very recent overview of AI in pancreatic cancer; the authors discussed how concentration on the genomic, proteomic, and imaging data may enhance diagnostic accuracy and treatment strategy planning. This trend of focusing on the entirety of the rare disease with a data-centered approach is set to continue, hence, the need to create more complex ML models that can handle

the many data types. For future work, there is a need to develop guidelines on how multi-omics data are collected and integrated and to design efficient ML structures for heterogeneous data processing.

The application of ML in context with rare diseases drug discovery and drugs repurposing was brought as a promising area for future research. In this section, we revisited ML applications in drug discovery as described by Réda et al. (2020) and ascertained that the ML can potentiate the discovery of new drugs for RDs. The effectiveness and toxicity of drugs in relation to rare diseases can potentially be predicted hence greatly improve the clinical trials and personalized treatment. Furthermore, applying the concepts of ML in natural history research, as elaborated by Liu et al. (2022) may be of use in understanding diseases' evolution and resultant remedial measures. Further studies should focus on the ways through which the ML-based drug discovery platforms can be integrated with an overall diagnostic model to provide a full-fledged management of rare diseases.

ML was also shown to have the possibility of contributing to pediatric rare disease identification and treatment in the review. Ardahan Sevgili and Şenol (2023) showed how AI could be used for the prediction of complications during chemotherapy in paediatric oncology. More specifically, the dynamics of children's growth, as well as relative scarceness of information about rare diseases patterns create a realm for AI/ML innovations. The future work needs to address the issues of early-age involvement, or heterochrony, considering that both development and the disease process may vary depending on age. Further, better integration of ML tools within pediatric environments, such as children specific EHRs and decision support systems, could greatly facilitate the focus for rare diseases in such children.

6. Conclusion

This systematic review has described the existing scenario of the use of machine learning techniques for diagnosing rare diseases relying on medical images. The studies indicate that there is a great opportunity for the improvement of rare diseases diagnosis accuracy, speed and time using ML techniques. Computer vision models such as convolutional deep neural networks and other deep learning models have been successful in analyzing rare medical images of different diseases.

The use of several clinical data, genomic and/or imaging data has become a promising way to enhance diagnostic performance. Though, some limitations are recorded, data constraint, imbalanced dataset, interpretability of model and external validation. Nevertheless, the reviewed studies suggest that the application of the described ML-based tools might greatly shorten the diagnostic odyssey which many rare diseases patients undergo.

To make the use of ML in clinical practice for the diagnosis of rare diseases possible, it will be noble for the researchers and clinicians to work hand in hand with policymakers. Specifically, there is a severe lack of standardization of the data collection process, model building, and reporting of results needed for the field's progression. Furthermore, compliance with the existing rules and regulations and ethical issues will also be critical for the integration of Machine Learning-based diagnostic tools to the high-tech healthcare organizations.

Some of the other phenomena advanced as the potential research areas are federated learning, multi-task models, and the combined use of multi-omics data. The possibility of the utilization of ML does not stop at diagnosis but encompasses drug development, treatment regimens, and molecular and individual-based therapies for orphan diseases.

6.1. Recommendations

Based on the findings of this review, several recommendations can be made to advance the field of ML-based rare disease diagnosis: Based on the findings of this review, several recommendations can be made to advance the field of ML-based rare disease diagnosis:

Firstly, there is the problem of the lack of data, particularly the samples themselves do not contain enough information for the development of truly effective ML models. The authors, therefore, recommend academia and researchers to develop multi-center consortia and data sharing strategies that will facilitate the enhancement of rare disease imaging databases. This is essential to ensure that data sharing is secure while at the same time ensure that the patient's rights to privacy are respected. Furthermore, limit augmentation techniques like generative adversarial networks should be considered as an improvement to limited databases while not violating the right to patient's privacy.

Secondly, in the future, there should be more research on the possibilities of automating the creation of interpretable ML models that will help clinicians understand why these or other decisions have been made. However, deep learning architecture demonstrated high performance, at the same time, they have many shortcomings due to their 'black box'

feature. It is necessary to use attention mechanism, feature importance analysis and locally interpretable model-agnostic explanations (LIME) into rare disease diagnostic models. This increased transparency will not only help trust the AI systems from the side of the healthcare professionals but also improve the understanding of diseases and identification of potential targets for development of treatments.

Thirdly, by association, there is an apparent requirement for conformity assessment procedures tied to methods that delineate rare disease research data validation protocols and evaluation criteria distinct from or complementary to measures utilized in other medical domains. Investigators should stop focusing on reporting accuracy on a few test cases and instead concentrate on external validation on far more complex and multicenter datasets. Future research centered on exploring the effectiveness of the ML-based diagnostic tools in relation to patient outcomes and the way utilization of such tools affects the utilization of health care resources should be performed. Furthermore, initiated benchmarks of rare diseases and public challenge datasets would help to compare different approaches in the field of ML fairly.

Fourthly, interdisciplinary collaboration should be enhanced to connect the workers in the field of ML and specialists in the field of rare diseases. Recommendation: Training should be taught targeting clinicians to be competent, critically assessing and implementing ML-based diagnostic tools. On the other hand, the ML researchers should be encouraged to get more insight into the underlying mechanisms of rare diseases as well as clinical problems.

Through implementing these recommendations, the field would be closer to achieving its goals and objectives of using machine learning to enhance the lives of patients with rare diseases. Future trends in stroke research will need to include sustenance and growth of innovative and cooperative efforts that keep demanding high quality scientific investigation to move stroke research closer to this successful patient population.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Abdel-Nasser, M., Rashwan, H. A., Puig, D., & Moreno, A. (2015). Analysis of tissue abnormality and breast density in mammographic images using a uniform local directional pattern. *Expert Systems with Applications*, 42(24), 9499-9511. https://doi.org/10.1016/j.eswa.2015.07.072

[2] Ahsan, M. M., & Alam, E. (2020). Deep MLP-CNN model using mixed-data to distinguish between COVID-19 and non-COVID-19 patients. *Symmetry*, 12, 1526. https://doi.org/10.3390/sym12091526

[3] Ahsan, M. M., Luna, S. A., & Siddique, Z. (2022). Machine-Learning-Based Disease Diagnosis: A Comprehensive Review. *Healthcare*, 10(3), 541. https://doi.org/10.3390/healthcare10030541

[4] Aljaaf, A. J., Al-Jumeily, D., Haglan, H. M., Alloghani, M., Baker, T., Hussain, A. J., & Mustafina, J. (2018). Early predictions of chronic kidney disease using machine learning supported by predictive analytics. In *Proceedings of the 2018 IEEE Congress on Evolutionary Computation (CEC)*. https://doi.org/10.1109/CEC.2018.8477876

[5] Álvarez-Machancoses, Ó., & Fernández-Martínez, J. L. (2019). Using artificial intelligence methods to speed up drug discovery. *Expert Opinion on Drug Discovery*, 14(8), 769-777. https://doi.org/10.1080/17460441.2019.1621284

[6] Alves, V. M., Korn, D., Pervitsky, V., Faller, D. V., Borda, E., Tropsha, A., & Fourches, D. (2022). Knowledge-based approaches to drug discovery for rare diseases. *Drug Discovery Today*, 27(2), 490-502.

[7] Ardahan Sevgili, S., & Şenol, S. (2023). Prediction of chemotherapy-related complications in pediatric oncology patients: artificial intelligence and machine learning implementations. *Pediatric Research*, 93(2), 390-395. https://doi.org/10.1038/s41390-022-02356-6

[8] Banerjee, N. (2020). Machine learning techniques for prediction of lung cancer. *International Journal of Recent Technology and Engineering*, 8(6), 2277-3878.

[9] Bonomi, F., Peretti, S., Lepri, G., Cosso, C., Pesce, G., Cosentino, N., Miriello, C., Cutolo, M., & Trombetta, A. C. (2022). The use and utility of machine learning in achieving precision medicine in systemic sclerosis: a narrative review. *Journal of Personalized Medicine*, 12(8), 1198.

[10] Brasil, S., Neves, C. J., Rijoff, T., Falcão, M., Valadão, A. G., Videira, P. A., & Dos Reis Ferreira, V. (2021). Artificial intelligence in epigenetic studies: shedding light on rare diseases. *Frontiers in Molecular Biosciences*, 8, 648012. https://doi.org/10.3389/fmolb.2021.648012

[11] Brasil, S., Pascoal, C., Francisco, R., Dos Reis Ferreira, V., Videira, P. A., & Valadão, A. G. (2019). Artificial Intelligence (AI) in Rare Diseases: Is the Future Brighter? *Genes*, 10, 978.

[12] Brasil, S., Pascoal, C., Francisco, R., Dos Reis Ferreira, V., Videira, P. A., & Valadão, A. G. (2019). Artificial intelligence (AI) in rare diseases: is the future brighter? *Genes (Basel)*, 10(12), 978.

[13] Charleonnan, A., Fufaung, T., Niyomwong, T., Chokchueypattanakit, W., Suwannawach, S., & Ninchawee, N. (2016). Predictive analytics for chronic kidney disease using machine learning techniques. In *Proceedings of the 2016 Management and Innovation Technology International Conference*. https://doi.org/10.1109/MITICON.2016.8025242

[14] Collaboration on Rare Diseases (CORD-MI). (n.d.). Retrieved April 16, 2020, from https://www.medizininformatik-initiative.de/en/CORD

[15] European Joint Programme on Rare Diseases. (n.d.). Retrieved April 16, 2020, from https://www.ejprarediseases.org

[16] Faviez, C., Chen, X., Garcelon, N., Neuraz, A., Knebelmann, B., Salomon, R., Lyonnet, S., Saunier, S., & Burgun, A. (2020). Diagnosis support systems for rare diseases: a scoping review. *Orphanet Journal of Rare Diseases*, 15(1), 94.

[17] Fitriyani, N. L., Syafrudin, M., Alfian, G., & Rhee, J. (2019). Development of disease prediction model based on ensemble learning approach for diabetes and hypertension. *IEEE Access*, 19, 144777-144789.

[18] Huang, B., Huang, H., Zhang, S., Lan, Q., Zhang, T., Liao, X., Liu, X., & Li, X. (2022). Artificial intelligence in pancreatic cancer. *Theranostics*, 12(16), 6931-6954.

[19] Hurvitz, N., Azmanov, H., Kesler, A., & Ilan, Y. (2021). Establishing a second-generation artificial intelligence-based system for improving diagnosis, treatment, and monitoring of patients with rare diseases. *European Journal of Human Genetics*, 29(10), 1485-1490.

[20] Islam, Z., & Asraf, A. (2020). A combined deep CNN-LSTM network for the detection of novel coronavirus (COVID-19) using X-ray images. *Informatics in Medicine Unlocked*, 20, 100412.

[21] Jacob, J., Chakkalakal, M. J., Johns, M., & Elizabeth, I. (2018). Diagnosis of liver disease using machine learning techniques. *International Research Journal of Engineering and Technology*, 5(4), 1.

[22] James, K. N., Phadke, S., Wong, T. C., & Chowdhury, S. (2023). Artificial intelligence in the genetic diagnosis of rare disease. *Clinics in Laboratory Medicine*, 43(1), 127-143. https://doi.org/10.1016/j.cll.2022.09.023

[23] Jin, C., Chen, W., Cao, Y., Xu, Z., Tan, Z., Zhang, X., Deng, L., Zheng, C., Zhou, J., Shi, H., & Feng, J. (2020). Development and evaluation of an artificial intelligence system for COVID-19 diagnosis. *Nature Communications*, 11(1), 5088. https://doi.org/10.1038/s41467-020-18685-1

[24] Kavakiotis, I., Tsave, O., Salifoglou, A., Maglaveras, N., Vlahavas, I., & Chouvarda, I. (2017). Machine learning and data mining methods in diabetes research. *Computational and Structural Biotechnology Journal*, 15, 104–116.

[25] Komal Kumar, N., Sarika Sindhu, G., Krishna Prashanthi, D., & Shaeen Sulthana, A. (2020). Analysis and Prediction of Cardiovascular Disease using Machine Learning Classifiers. In *6th International Conference on Advanced Computing & Communication Systems (ICACCS)*. https://doi.org/10.1109/ICACCS48705.2020.9074183

[26] Kourou, K., Exarchos, T. P., Exarchos, K. P., Karamouzis, M. V., & Fotiadis, D. I. (2015). Machine learning applications in cancer prognosis and prediction. *Computational and Structural Biotechnology Journal*, 13, 8–17.

[27] Krawitz, P. M. (2022). Artificial intelligence in the diagnosis of rare disorders: the development of phenotype analysis. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz*, 65(11), 1159-1163.

[28] Liu, J., Barrett, J. S., Leonardi, E. T., Lim, J., Thakar, A., Venuto, C., Schilling, S., Gobburu, J. V., & Guglieri, M. (2022). Natural history and real-world data in rare diseases: applications, limitations, and future perspectives. *Journal of Clinical Pharmacology*, 62(Suppl 2), S38-S55. https://doi.org/10.1002/jcph.2134

[29] Md, N. E. (2017). The Performance Analysis of K-Nearest Neighbors (K-NN) algorithm for motor imagery classification based on E.E.G. Signal. *MATEC Web of Conferences*, 104(01024), 6.

[30] Meshref, H. (2019). Cardiovascular Disease Diagnosis: A Machine Learning Interpretation Approach. *International Journal of Advanced Computer Science and Applications*, 10(12). https://doi.org/10.14569/IJACSA.2019.0101236

[31] Moon, W. K., Chen, I. L., Chang, J. M., Shin, S. U., Lo, C. M., & Chang, R. F. (2017). The adaptive computer-aided diagnosis system based on tumor sizes for the classification of breast tumors detected at screening ultrasound. *Ultrasonics*, 76, 70-77. https://doi.org/10.1016/j.ultras.2016.12.017

[32] Naz, H., & Ahuja, S. (2020). Deep learning approach for diabetes prediction using PIMA Indian dataset. *Journal of Diabetes & Metabolic Disorders*, 19(1), 391-403. https://doi.org/10.1007/s40200-020-00520-5

[33] Neelaveni, J. (n.d.). Alzheimer Disease Prediction using Machine Learning Algorithm. In *6th International Conference on Advanced Computing & Communication Systems*.

[34] Orphadata. (n.d.). Rare diseases and classifications. Retrieved April 16, 2020, from http://www.orphadata.org/cgi-bin/rare_free.html

[35] Oyewo, O. (2020). Prediction of Prostate Cancer using Ensemble of Machine Learning Techniques. *International Journal of Advanced Computer Science and Applications*, 11(3).

[36] Peters, M. D. J., Godfrey, C. M., Khalil, H., McInerney, P., Parker, D., & Soares, C. B. (2015). Guidance for conducting systematic scoping reviews. *International Journal of Evidence-Based Healthcare*, 13, 141–146.

[37] Petzold, A., Albrecht, P., Balcer, L., Bekkers, E., Brandt, A. U., Calabresi, P. A., Deborah, O. G., Graves, J. S., Green, A., Keane, P. A., Nij Bijvank, J. A., Sander, J. W., Paul, F., Saidha, S., Villoslada, P., Wagner, S. K., Wegener, A., Schippling, S., & ERN-EYE OCTA consortium. (2021). Artificial intelligence extension of the OSCAR-IB criteria. *Annals of Clinical and Translational Neurology*, 8(7), 1528-1542. https://doi.org/10.1002/acn3.51320

[38] Rajendran, K., Jayabalan, M., & Thiruchelvam, V. (2020). Predicting breast cancer via supervised machine learning methods on class imbalanced data. *International Journal of Advanced Computer Science and Applications*, 11(8), 54-63. https://doi.org/10.14569/IJACSA.2020.0110808

[39] Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. *New England Journal of Medicine*, 380, 1347–1358.

[40] Réda, C., Kaufmann, E., & Delahaye-Duriez, A. (2020). Machine learning applications in drug development. *Computational and Structural Biotechnology Journal*, 18, 241–252.

[41] Song, Y., Zheng, S., Li, L., Zhang, X., Zhang, X., Huang, Z., Chen, J., Wang, R., Zhao, H., Zha, Y., Shen, J., Chong, Y., & Yang, Y. (2021). Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 18(6), 2775-2780. https://doi.org/10.1109/TCBB.2021.3065361

[42] Taroni, J. N., Grayson, P. C., Hu, Q., Eddy, S., Kretzler, M., Merkel, P. A., & Greene, C. S. (2019). MultiPLIER: a transfer learning framework for transcriptomics reveals systemic features of rare diseases. *Cell Systems*, 8(5), 380.e4-394.e4. https://doi.org/10.1016/j.cels.2019.04.003

[43] Tian, X. (2019). Using machine learning algorithms to predict Hepatitis B surface antigen seroclearance. *Computational and Mathematical Methods in Medicine*, 2019, 6915850. https://doi.org/10.1155/2019/6915850

[44] Topol, E. (2019). *Deep medicine: how artificial intelligence can make healthcare human again* (1st ed.). Basic Books.

[45] Tsao, H. Y., Chan, P. Y., & Su, E. C. Y. (2018). Predicting diabetic retinopathy and identifying interpretable biomedical features using machine learning algorithms. *BMC Bioinformatics*, 19(S9), 283. https://doi.org/10.1186/s12859-018-2277-0

[46] Vidya, M., & Maya, V. (2020). Skin cancer detection using machine learning techniques. In *2020 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT)*.

[47] Visibelli, A., Roncaglia, B., Spiga, O., & Santucci, A. (2023). The impact of artificial intelligence in the odyssey of rare diseases. *Biomedicine*, 11(3), 887.

[48] Yaswanth, R. (2020). Heart disease prediction using machine learning techniques. *International Journal of Innovative Technology and Exploring Engineering*, 9(5).

[49] Zeebaree, D. Q., Haron, H., Abdulazeez, A. M., & Zeebaree, S. R. M. (2017). Combination of K-means clustering with Genetic Algorithm: A review. *International Journal of Applied Engineering Research*, 12(24), 14238-14245.

[50] Yousef, R., Gupta, G., Yousef, N., & Khari, M. (2022). A holistic overview of deep learning approach in medical imaging. *Multimedia Systems*, *28*(3), 881-914. https://link.springer.com/article/10.1007/s00530-021-00884-5

[51] Schaefer, J., Lehne, M., Schepers, J., Prasser, F., & Thun, S. (2020). The use of machine learning in rare diseases: a scoping review. *Orphanet journal of rare diseases*, *15*, 1-10. https://link.springer.com/article/10.1186/s13023-020-01424-6