# SHUBHAM SINGH

1 of 17

A**bstract: Early diagnostic discovery of disease globally is a significant challenge due to the complexities of disease mechanisms and diverse patient symptomology. Developing tools that are capable of disease discovery earlier is difficult, but advancements in machine learning (ML), a subfield of artificial intelligence (AI), are beginning to address these issues. ML provides a framework for researchers, doctors, and patients to discover solutions to these challenges. This review discusses the applications of ML in disease diagnosis, particularly in disease detection. It begins with a bibliometric exploration of disease diagnosis in ML literature, which is tracked in 1,216 publications from databases, including Scopus and Web of Science (WOS). This analysis provides an overview of the most productive authors, countries, institutions, and the most cited articles in the literature. Second, the review will characterize the trends and recent application practices of machine learning-based disease diagnosis (MLBDD), including types of algorithms used, diseases being diagnosed, data/inputs used, applications of ML, including metrics utilized to measure the impact of ML-DBD. In conclusion, the paper will summarize the findings and provide some thoughts of future prospects and opportunities in the area of ML-based disease discovery.**

**Keywords:** artificial neural networks; convolutional neural networks; COVID-19; deep learning; deep neural networks; diabetes; disease diagnosis; heart disease; kidney disease; machine learning; review

* the condition of rare disease
* the disease is omitted mistakenly from the consideration

Machine Learning (ML) has emerged as a strong tool across many domains, from cutting-edge technologies in smartphones, computers, and robots, to healthcare settings involving manageable problems like diagnosing patients or ensuring patient safety. It is particularly gaining traction in the healthcare sector, and already shows significant promise in optimizing how diseases are diagnosed. So far, ML-based disease diagnosis (MLBDD) has been characterized by its cost-effectiveness and time effectiveness relative to traditional diagnosis processes. While traditional diagnostic approaches are costly, time-consuming, and often very human-dependent, ML systems can operate with automation and process expansive datasets without significant human limitations, such as fatigue.ML has the potential to facilitate improvements in the accuracy of diagnosis and increase turnaround time via the use of healthcare data, such as medical images (such as X-rays or MRIs) and tabular datasets (numerical and textual patient information including basics such as age, gender, and medical history). ML is a sub-field of artificial intelligence (AI) that operates by processing data through complex mathematical functions to produce results, which could be through classification or regression. ML can perform complex and difficult tasks that humans usually cannot, such as effectively identifying malignant cells in a microscopic image. Furthermore, the accuracy of ML has improved with advancement in deep learning, which is another type of machine learning.

## Basics and Background

Machine learning (ML) can be described as a process that utilizes statistical and mathematical methods to analyze data samples that generate significant insights that in turn allow machines to learn and make decisions without explicit programming. Arthur Samuel made the first assertion about "learning" in ML in 1959, in regard to games and pattern recognition, demonstrating that computers were capable of learning from experience. This was a landmark event in ML in terms of recognizing "the machine learns."The essence of ML is to allow systems to learn from a dataset and be able to predict some outcome or make a decision based on the task. With recent advancements in technology, especially in computing ability and data storage, the training of ML

models leads to considerably enhanced accuracy that benefits real-world predictions. In a wide variety of cases, tasks which at one time were incredibly time-consuming and required considerable human resources have been significantly automated and completed rapidly with little human interpretation of results.The rapid growth of ML has also benefited from greatly improved capabilities for working with large datasets and processing speed. More available data allow ML models to be accurate in predicting outcomes from learning, allowing for widespread use of the approach. Types of ML and related approaches have been investigated research, whereby numerous articles have offered new methods or enhancements.

Machine learning (ML) algorithms can generally be grouped into three primary types: supervised, unsupervised, and semi-supervised learning [10]. These classifications are based on the method by which the model is trained and how the data is applied.1.

-Supervised Learning: With supervised learning, the model is trained with labeled data, i.e., each training example is associated with a target or outcome. The algorithm learns the relationship between the input data and the output target, and it can then predict the target variable on new, unknown data samples. Common examples of supervised algorithms include

* Linear regression: Used to predict continuous values.
* Logistic regression: Typically used when the task is a binary classification type.
* Support Vector Machine (SVM): Uses classification or regression type algorithms to find a hyperplane that separates classes of data.2.

-Unsupervised Learning: In unsupervised learning, the model is given the data without labels. The objective is to find patterns or class rations in the data without a target. Unsupervised learning is more commonly used for clustering and dimensionality reduction tasks. Common unsupervised algorithms include:

* K-means clustering: Commonly used when the goal is to cluster data points that are similar to each other.
* rincipal Components Analysis (PCA): Used to reduce the dimensionality of the dataset while explaining as much of the variance as possible.3.

-Semi-supervised Learning: Semi-supervised learning is generally

# SHUBHAM SINGH

3 of 17



* 1. *Machine Learning Algorithms*

This section provides a comprehensive review of the most frequently used machine learning algorithms in disease diagnosis.

* + 1. Decision Tree

The decision tree (DT) algorithm is grounded in divide-and-conquer principles. In DT models, the attribute can have several values, so they are known as classification trees; leaves indicate separate classes, while branches indicate the combination of characteristics associated with that class label. Decision trees can also, however, take continuous variables, called regression trees. C4.5 and EC4.5 are the two most well-known and widely used DT algorithms [12]. DT is widely used by the following reference literature

* + 1. Support Vector Machine

Support vector machine (SVM) is a widely used machine learning approach for classification and regression-type problems. SVM was first introduced by Vapnik in the late 20th century [17]. SVMs are also used in a variety of areas apart from disease diagnosis such as facial expression, protein fold, distant homology detection, speech recognition, and text classification. For unlabeled data, supervised ML algorithms cannot work. Using a hyperplane to find the clustering of the data, SVM can classify unlabeled data. In addition, the SVM output is not nonlinearly separable. Selecting a suitable kernel and parameters are two major factors when using SVM in data analysis [11].

* + 1. *K*-Nearest Neighbor:

The K-nearest neighbor (KNN) classification is a nonparametric classification method that was first introduced in 1951 by Evelyn Fix and Joseph Hodges. KNN can be used for classification and regression analysis. The result of KNN classification is class membership. A voting mechanism

classifies an item. The Euclidean distance methods are used to find the distance between two samples of data. The predicted value of KNN in the case of regression analysis is the mean of the KNN [18].

* + 1. Naïve Bayes:

The naïve Bayes (NB) classifier represents a probabilistic classifier that is Bayesian-based. It predicts membership probability for each class based on the input record or data point. The class with the highest probability is the predicted class. The NB classifier predicts likelihoods rather than predictions [11].

* + 1. Logistic Regression

Logistic regression (LR) is a machine learning (ML) technique for classification problems. The LR model is considered probabilistic and predicts output values between 0 and 1. Examples of LR approaches in ML are spam identification in emails, online fraud transactions, and detection of malignant tumors. LR uses a cost function also known as the sigmoid function, which takes every real number such that the output is in the range of 0 and 1 [19].

* + 1. AdaBoost

Adaptive Boosting, also known as AdaBoost, was developed by Yoav Freund and Robert Schapire. AdaBoost is essentially a classifier combining different weak classifiers into a single classifier. AdaBoost works by increasing the weight placed on samples that are difficult to classify while decreasing weight on those samples that have been well classified. It can be used for classification and regression analysis [20].

* 1. *Deep Learning Overview*

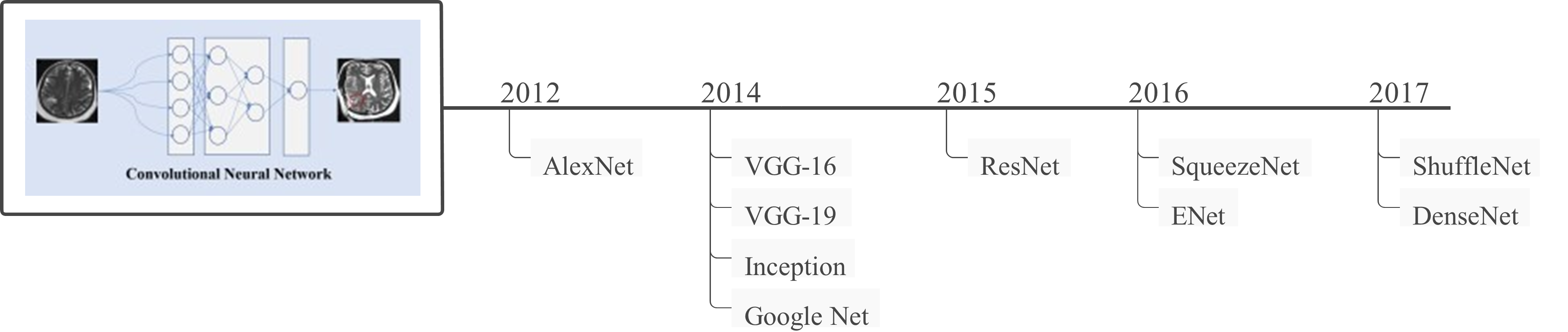
Deep learning (DL) is a branch of machine learning (ML) that aims to extract higher and lower- level information from input (i.e., images, numerical value, categorical values) using multiple layers. Most modern DL models are based on artificial neural networks (ANN), specifically convolutional neural networks (CNN), and may be combined with other DL models such as generative models, deep belief networks, and the Boltzmann machine. DL can be categorized as supervised, semisupervised, or unsupervised. Some well-known architectures for DL are deep neural networks (DNN), reinforcement learning, and recurrent neural networks (RNN) [21].At each level in DL, it learns to transform its input data to the next layers while learning different characteristics of the data. For example, when processing images, raw input could be a pixel matrix, where the first layers may learn to detect the image's edges. The second layer then will construct and encode the nose, eyes, etc. The third layer may detect the face after incorporating all the information captured by the previous two layers [6].DL has tremendous potential to impact medical fields. DL has been heavily used in the fields of radiology and pathology for disease diagnosis.

Convolutional Neural Network

Convolutional neural networks (CNNs) are a subcategory of artificial neural networks (ANNs) that have found various applications in image processing. CNNs have been used significantly for facial recognition, textual analysis, human organ position identification, and biological image detection and recognition [24]. After CNN was introduced in 1989, a different type of CNN's architecture has been developed over 30 years which has been effective in disease diagnosis. A CNN architecture has three components: the input layer, hidden layer, and output layer. Any feedforward network with intermediate levels are referred to as hidden layers and the number of hidden layers depends on the specific architecture type. The first layers that perform convolutions are hidden layers that consist of dot products of convolution kernel with an input matrix, and each convolutional layer generates feature maps that offer inputs to the next layers. Following the hidden layers, there are additional layers, including pooling layers and fully connected layers [21]. Over the years, numerous CNN models have been proposed, and popular CNN Models and the most frequently used CNN models are listed in Figure 2.

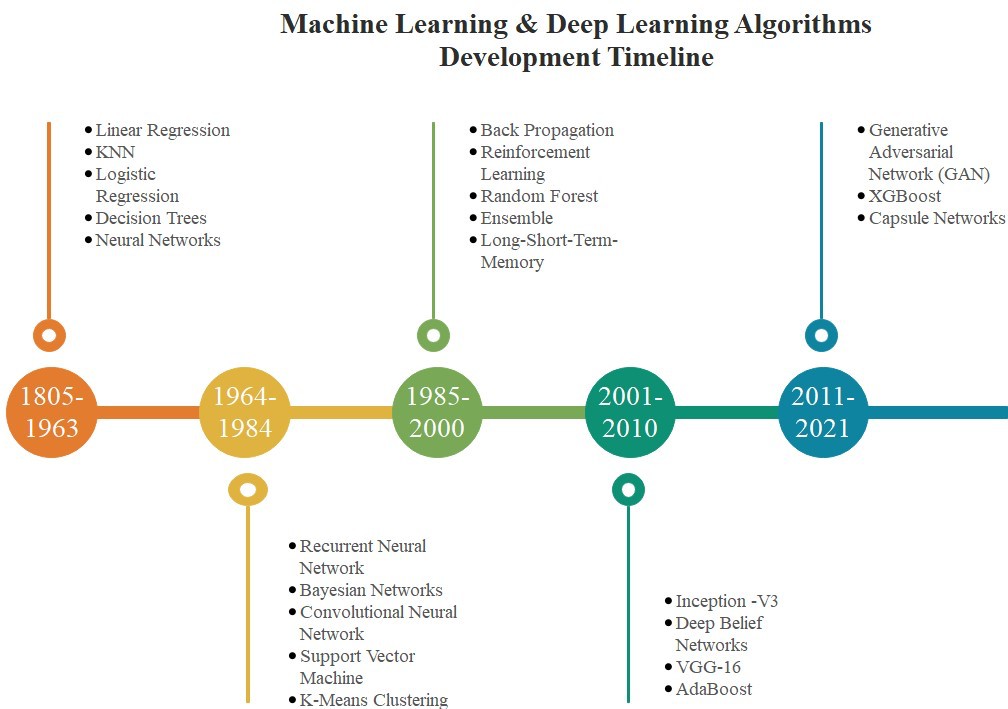
# SHUBHAM SINGH

5 of 17



**Figure 2.** Some of the most well-known CNN models, along with their development time frames.

In general, it may be considered that ML and DL have grown substantially throughout the years. The increased computational capability of computers and the enormous number of data available inspire academics and practitioners to employ ML/DL more efficiently. A schematic overview of machine learning and deep learning algorithms and their development chronology is shown in Figure 3, which may be a helpful resource for future researchers and practitioner.



**Figure 3.** Illustration of machine learning and deep learning algorithms development timeline.

* 1. *Performance Evaluations*

This part discusses the evaluation metrics used in the reference literature. Evaluation metrics used in disease diagnosis are accuracy, precision, recall, and f1 score. Lung cancer, for instance, can be classified as TP or TN if diagnosed correctly, or FP or FN, if diagnosed incorrectly. Below are the most widely used metrics [10]. Accuracy (Acc) : Accuracy represents the total correctly identified instances out of all of the instances. Accuracy can be calculated using following equations:

*Tp* + *TN*

*ACC* = (1)

*Tp* + *TN* + *Fp* + *FN*

***Precision (****Pn****):*** Precision is measured as the proportion of precisely predicted to all expected positive observations.

*Tp*

*Pn* = (2)

*Tp* + *Fp*

***Recall (****Rc****):*** The proportion of overall relevant results that the algorithm properly recognizes is referred to as recall.

*Tp*

*Tn* + *Fp*

(3)

***Sensitivity (****Sn****)*:** Sensitivity denotes only true positive measure considering total instances and can be measured as follows:

*Tp*

*Sn* = (4)

*Tp* + *FN*

***Specificity (****Sp****):*** It identifies how many true negatives are appropriately identified and calculated as follows:

*TN*

*Sp* = (5)

*Tn* + *Fp*

***F-measure:*** The F1 score is the mean of accuracy and recall in a harmonic manner. The highest F score is 1, indicating perfect precision and recall score.

Precision × Recall

*F* − *Measure* = 2 × (6)

Precision + Recall

***Area under curve (AUC):*** The area under the curve represents the models’ behaviors in different situations. The AUC can be calculated as follows:

∑ *Ri*(*Ip*)− *Ip*((*Ip* + 1)/2)

*AUC* = (7)

*Ip* + *In*

where *lp* and *ln* denotes positive and negative data samples and *Ri* is the rating of the *i*th positive samples.

## Article Selection

* 1. *Identification*

The databases Scopus and Web of Science (WOS) are being used to find original research publications. Because of their high quality and peer review index for papers, Scopus and WOS are significant databases when searching for articles, as many students and scholars have used Scopus and WOS for systematic review [25,26]. Using keywords in combination with Boolean operators, the title search was performed in the following way:“disease” AND (“diagnosis” OR “Support vector machine” OR “SVM” OR “KNN” OR“K-nearest neighbor” OR “logistic regression” OR “K-means clustering” OR “random forest” OR “RF” OR “adaboost” OR “XGBoost” , “decision tree” OR “neural network” OR “NN” OR “artificial neural network” OR “ANN" OR “convolutional neural network” OR “CNN” OR “deep neural network” OR “DNN” OR “machine learning" or “adversarial network” or “GAN”).The search started with a total of 16,209 from Scopus, and 2129 from Web of Science (WOS).

* 1. *Screening*

Once the search period was narrowed to 2012–2021 and only peer-reviewed English papers were evaluated, the total number of articles decreased to 9117 for Scopus and 1803 for WOS, respectively.

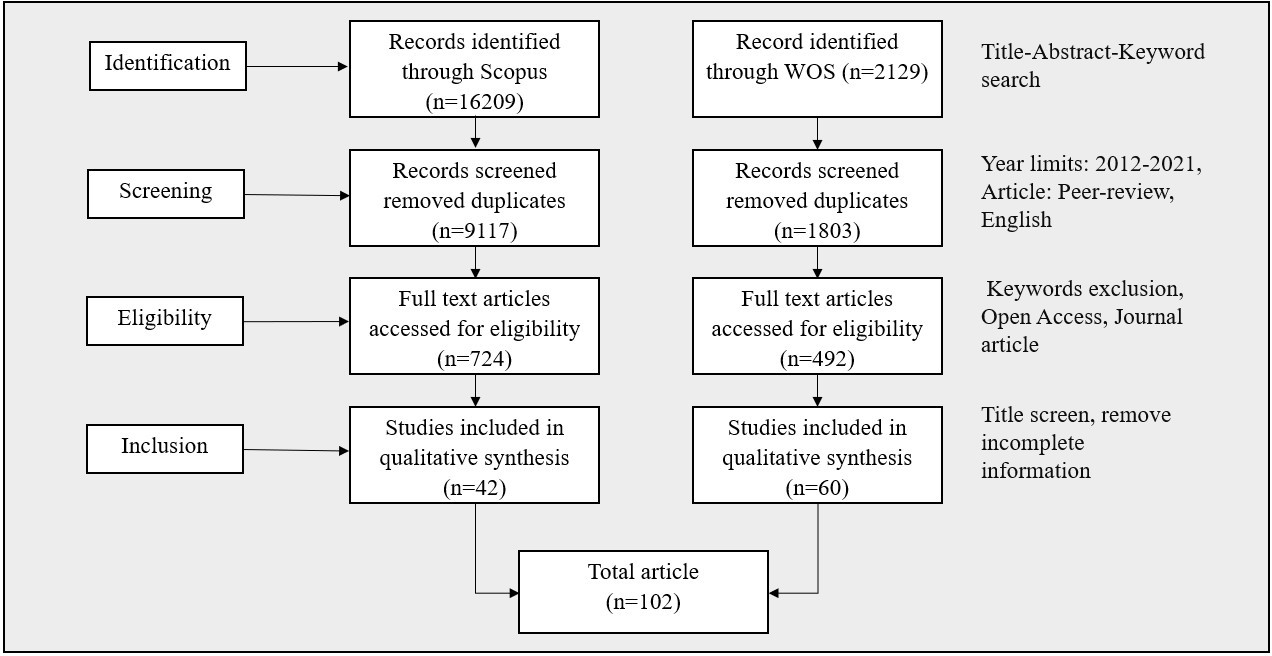
* 1. *Eligibility and Inclusion*

These articles were chosen for further exploration if they are open access and in the journal form. Overall, there were 1216 full-text articles (724 from Scopus database and 492 from WOS).

# SHUBHAM SINGH

7 of 17

Bibliographic analysis was carried out on all 1216 included articles. One investigator (Z.S.) imported information from 1216 articles as excel CSV data for future analysis. We then used the excel duplication functions to identify and remove duplicate articles. Two independent reviewers (M.A. and Z.S.) reviewed the titles and abstracts of 1192 articles. Disagreements were resolved through discussion. We removed articles that did not pertain to machine learning but the disease diagnosis did or vice-versa. After examining the titles and abstracts, we then reviewed the full texts of 102 total papers, with all 102 papers meeting all inclusion criteria. Reasons for exclusion from full text review, included.

Figure 4 shows the flow diagram of the systematic article selection procedure used in this study.

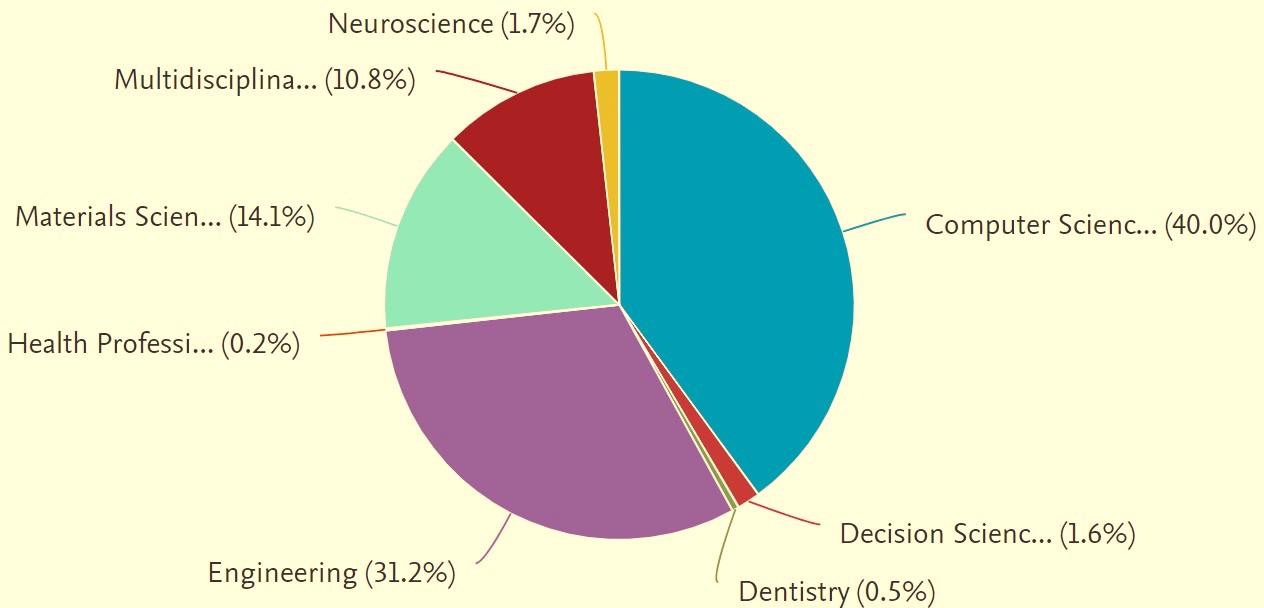
**Figure 4.** MLBDD article selection procedure used in this study.

## Bibliometric Analysis

The bibliometric study in this section was carried out using reference literature gathered from the Scopus and WOS databases. The bibliometric study examines publications in terms of the subject area, co-occurrence network, year of publication, journal, citations, countries, and authors.

* 1. *Subject Area*

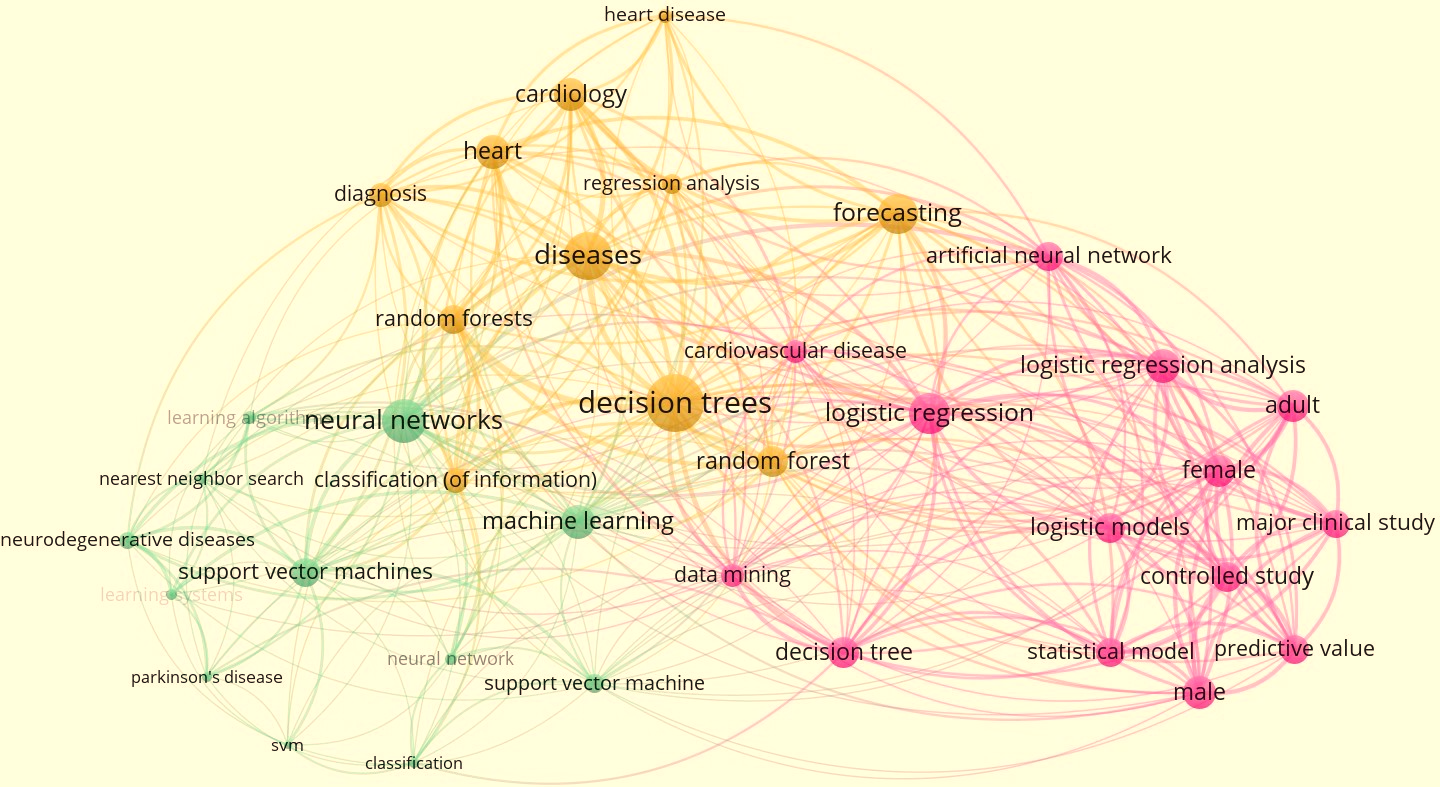
Many research disciplines have uncovered machine learning-based disease diagnostics throughout the years. Figure 5 depicts a schematic representation of machine learningbased disease detection spread across several research fields. According to the graph, computer science (40%) and engineering (31.2%) are two dominating fields that vigorously concentrated on MLBDD.



**Figure 5.** Distribution of articles by subject area.

* 1. *Co-Occurrence Network*

The co-occurrence of keywords provides a general understanding of how researchers use them or how the keywords are associated with each other. Figure 6 displays the keyword co- occurrence network and its connection of the article as was created with VOSviewer software. From the figure, we see some of the larger keyword clusters consist of neural networks (NN), decision trees (DT), machine learning (ML), and logistic regression (LR). Each larger cluster is also linked to other keywords that fall within those larger clusters. For example, the NN cluster contains support vector machine (SVM), Parkinson’s disease, and classification.



**Figure 6.** Bibliometric map representing co-occurrence analysis of keywords in network visualization.

## Machine Learning Techniques for Different Disease Diagnosis

Machine learning (ML) techniques have been employed in disease diagnosis by many authors and professionals. This section describes a range of types of machine learning-based disease diagnosis (MLBDD) that have received a lot of attention because of their significance and severity. For instance, due to COVID-19 being of global significance, many studies have focused on COVID-19 disease detection using ML during the period from 2020 onwards, this also have received higher precedence in our study. Serious diseases like heart disease, kidney disease, breast cancer, diabetes, Parkinson’s, Alzheimer’s and COVID-19 are discussed in brief, while other diseases are addressed in brief with the “other disease” tag.

* 1. *Heart Disease*

Most researchers and practitioners use machine learning (ML) approaches to identify cardiac disease [37,38]. Ansari et al. (2011), for example, offered an automated coronary heart disease diagnosis system based on neurofuzzy integrated systems that yield around 89% accuracy [37]. One of the study’s significant weaknesses is the lack of a clear explanation for how their proposed technique would work in various scenarios such as multiclass classification, big data analysis, and unbalanced class distribution. Furthermore, there is no explanation about the credibility of the model’s accuracy, which has lately been highly encouraged in medical domains, particularly to assist users who are not from the medical domains in understanding the approach.

Rubin et al. (2017) uses deep-convolutional-neural-network-based approaches to detect irregular cardiac sounds. The authors of this study adjusted the loss function to improve the training dataset’s sensitivity and specificity. Their suggested model was tested in the 2016 PhysioNet computing competition. They finished second in the competition,

with a final prediction of 0.95 specificity and 0.73 sensitivity [39].

Aside from that, deep-learning (DL)-based algorithms have lately received attention in detecting cardiac disease. Miao and Miao et al. (2018), for example, offered a DL-based technique to diagnosing cardiotocographic fetal health based on a multiclass morphologic pattern. The created model is used to differentiate and categorize the morphologic pattern of individuals suffering from pregnancy complications. Their preliminary computational findings include accuracy of 88.02%, a precision of 85.01%, and an F-score of 0.85 [40]. During that study, they employed multiple dropout strategies to address overfitting problems, which finally increased training time, which they acknowledged as a tradeoff for higher accuracy.

Although ML applications have been widely employed in heart disease diagnosis, no research has been conducted that addressed the issues associated with unbalanced data with multiclass classification. Furthermore, the model’s explainability during final prediction is lacking in most cases. Table 3 summarizes some of the cited publications that employed ML and DL approaches in the diagnosis of cardiac disease. However, further information about machine-learning-based cardiac disease diagnosis can be found in [5].

**Table 3.** Referenced literature that considered machine-learning-based heart disease diagnosis.

**Study Contributions Algorithm Dataset Data Type Performance Evaluation**

1. Predict coronary heart disease Gaussian NB,

Bernoulli NB, and RF

Cleveland dataset Tabular Accuracy— 85.00%, 85.00%

RF (Accuracy—80.327%, Precision—82%,

and 75.00%

Recall—80%,

1. Predicting heart diseases RF, CNN Cleveland dataset Tabular

F1-score—80%), CNN

(Accuracy—78.688, Precision—80%, Recall— 79%,

F1-score—78%)

1. Heart disease classification SVM Cleveland database Tabular Accuracy—73–91%
2. Heart disease classification Back-propagation

NN, LR

Cleveland dataset Tabular Accuracy (BNN—85.074%,

LR—92.58%)

1. ECG arrhythmia for heart disease detection

Intelligent scoring system for the

SVM and Cuckoo search optimized NN

Cleveland dataset Tabular Accuracy (SVM—94.44%)

Specificity—78.8%, Sensitivity—62.3%, Positive

prediction of cardiac arrest

within 72 h

Automatically identify 5 different categories of heartbeats in ECG signals

SVM Privately ownend Tabular

CNN MIT-BIH Tabular

predictive value—10%, Negative predictive value—98.2% Accuracy—94% (balance data) Accuracy—89.07% (imbalance data)

Accuracy—97.77%

1. Novel heartbeat recognition method is presented

SVM MIT-BIH Tabular

(imbalance data), Accuracy—97.08% (noise-free ECGs)

* 1. *Kidney Disease:*

*Renal disease, commonly referred to as kidney disease, indicates nephropathy or kidney injury. Individuals suffering from renal disease have reduced kidney functional activity, and without timely treatment, this will lead to kidney failure. As stated by the National Kidney Foundation 10% of the world's population has chronic kidney disease (CKD); millions of people die annually due to a lack of management related to CKD [49]. The emergence of ML- and DL-based kidney disease diagnosis and management could deliver hope to those countries who are unable to manage the diagnosing-related tests for kidney disease [49]. For instance, Charleonnan et al. (2016) focused on four various ML algorithms: K-nearest neighbors (KNN), support vector machine (SVM), logistic regression (LR), as well as decision tree classifiers using publicly available datasets and presented an accuracy of 98.1%, 98.3%, 96.55%, and 94.8%, respectively [50]. Aljaaf et al. (2018) undertook a similar study, with the authors testing different ML algorithms including RPART, SVM, LOGR, and MLP using a similar dataset CKD, as the dataset used by [50] and produced an accuracy of MLP at 98.1% in detecting chronic kidney disease [51]. To detect chronic kidney disease Ma et al. (2020) used a collection of datasets which contained data gathered from many data sources [52]. Their heterogeneous modified artificial neural network (HMANN) model achieved the accuracy of 87 - 99%.*

Table 4 summarizes some of the cited publications that employed ML and DL approaches to diagnose kidney disease.

**Table 4.** Referenced literature that considered machine-learning-based kidney disease diagnosis.

**Study Contributions Algorithm Dataset Data Type Performance Evaluation**

Analysis of Chronic Kidney

[13]

Disease

Kidney disease detection and

NB, DT, and RF Chronic kidney disease 100 collected image data

Tabular Accuracy—100% (RF)

[53]

dataset

segmentation ANN & kernel KMC

of patients Ultrasound

Image Accuracy—99.61%

Feedforward NN

[54]

Classification of Chronic kidney disease

LR, Feedforward NN and Wide DL

Chronic kidney disease Tabular dataset

(F1-score—99%, Precision—97%, Recall—99%, and AUC—99%)

Accuracy—97.67%,

[55] Chronic kidney disease CNN-SVM Privately own dataset Tabular

Detection and localization of

Sensitivity—97.5%, Specificity—97.83%

[56]

kidneys in patients with autosomal dominant polycystic

CNN Privately own data Image Accuracy—95%

* 1. *Breast Cancer*

Numerous academics in medicine have put forward utilizing machine-learning (ML)-based breast cancer analysis as a possible answer to this early-stage diagnosis question. Miranda and Felipe (2015), for example, proposed fuzzy-logic-based computer-aided diagnosis systems to perform

breast cancer categorization. With fuzzy logic, one advantage is the potential to lessen computational difficulty, while effectively simulating the expert radiologist's reasoning and mode of action. If there are specified parameters in, such as contour, form and density, another method for categorizing cancer again is provided by the algorithm based on that person's individual method [57]. The accuracy was achieved at roughly 83.34%, following Miranda and Felipe (2015)'s suggested model. It should be noted that all images used for the experiment were made up in approximately equal ratios which led to greater accuracy and no biases in the results. Note that since the authors did not weight their interpretation of their results in an explainable manner, it could be difficult to conclude that all accuracy in general denotes true accuracy for both conditions of benign or malignant. Additionally, no confusion matrix is included to evaluate the models' actual prediction for each class.One of the pieces of work (Zheng et al. 2014) included hybrid strategies for diagnosing breast cancer disease utilizing k-means clustering (KMC) and SVM strategies as a model. The accuracy achieved was certainly improved, and the dimensional issues reduced considerably.

**Study Contributions Algorithm Dataset Data Type Performance Evaluation**

[14] Breast cancer NB, BN, RF and DT Classification of breast density and

(C4.5)

BCSC Image ROC—0.937 (BN) Mini-MIAS:

[66]

mass SVM Mini-MIAS, INBreast Image

Accuracy—99%, AUC—

0.9325

IRMA: Sensitivity—99%,

1. Classify vector features as malignant or non-malignant
2. Classification of breast cancers by

tumor size

SVM IRMA, DDSM Image

LR-ANN 156 Privately owned cases Image

Specificity—99% , DDSM: Sensitivity—97%, Specificity—96% Accuracy—81.8%, Sensitivity—85.4%, Specificity—77.8%, AUC—0.855

1. CAD tumor Binary-LR 18 Privately owned cases Image Accuracy—80.39%

Differentiating malignant and NB, LR-AdaBoost 246 Privately owned image Image benign masses

* 1. *Diabetes*

Sensitivity—90%, Specificity—97.5%, AUC—0.98

As reported by the International Diabetes Federation (IDF), there are now over 382 million people worldwide with the disease, and this is projected to worsen to 629 million by 2045 [71]. Many studies also widely presented ML-based systems to identify diabetes patients. For example, Kandhasamy and Balamurali (2015) compared ML classifiers (J48 DT, KNN, RF, and SVM) on diabetes mellitus patient identification. The study used the UCI Diabetes dataset, and found that KNN (K = 1) and RF classifiers had near-complete success [72]. However, a downside of this research was predicated on a simplified dataset, consisting of only eight parameters, which were binary classified about Diabetes. Therefore, having an accuracy at 100% from a less difficult dataset is certainly not surprising. Furthermore, there is no conversation about how the algorithms relate to the final prediction nor how to look at the result from the nontechnical side of the experiment.Yahyaoui et al. (2019) presented a Clinical Decision Support Systems (CDSS) to assist physicians or practitioners in diagnosing Diabetes. To carry out this goal, the researchers used a variety of ML methods such as SVM, RF, and deep convolutional neural network (CNN). In their calculations, RF outperformed all other methods with an accuracy of 83.67%, while DL and SVM had accuracies of 76.81% and 65.38%, respectively [73].Naz and Ahuja (2020) used a

**Table 6.** Referenced literature that considered machine-learning-based diabetic disease diagnosis.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Study** | **Contributions** | **Algorithm** | **Dataset** | **Data Type** | **Performance Evaluation** |
| [76] | Diabetes and hypertension | DPM | Privately owned | Tabular | Accuracy—96.74% |
| [77] | Type 1 diabetes | RF | DIABIM-MUNE | Tabular | AUC—0.80 |

1. Diabetes classification KNN Privately owned4900

samples

Tabular Accuracy—99.9%

[15]

Predict diabetic retinopathy and identify interpretable biomedical features

SVM, DT, ANN, and Privately owned Tabular LR

SVM (Accuracy—79.5%, AUC—0.839)

1. Diabetes classification PSO and MLPNN Privately owned Tabular Accuracy—98.73%
   1. *COVID-19*

The new severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) pandemic COVID-19 has become humankind’s greatest challenge faced in recent times. Although the vaccine had been expedited in the distribution because of the global crisis, the majority of the world did not have access to <br> it for most of the pandemic[88]. Adding to the stress and strain is the fact that the new COVID-19 Omicron strain is highly transmissible and has elements of resistance related to the vaccine. The new gold standard for identifying COVID-19 infection is Real-Time Reverse Transcription-Polymerase Chain Reaction analysis (RT-PCR) [89,90]. During the pandemic, the researcher suggested potential technologies such as chest X-rays and Computed Tomography (CT) including Machine Learning and Artificial Intelligence to assist with identifying individuals who could be potentially infected. For example, Chen et al. (2020) proposed a UNet model of CT images for 51 COVID-19 patients and 82 non-COVID-19 patients, which resulted in an accuracy of 98.5% [91]. Ardakani et al. (2020) evaluated 10 different DL models using a small dataset of 108 patients with COVID-19 and 86 non-COVID-19 patients, resulting in an accuracy of 99% overall [92]. Wang et al. (2020) produced an inception-based model, using a large dataset of 453 CT scan images, with an accuracy of 73.1%. Nevertheless, the model’s network

**Table 8.** Referenced literature that considered machine-learning-based COVID-19 disease diagnosis.

|  |  |  |  |
| --- | --- | --- | --- |
| **Study** | **Contributions** | **Algorithm** | **Dataset** |
| [94] | COVID-19 disease detection | CNN | Mixed dataset |
| [91] | COVID-19 disease detection | CNN | Mixed dataset |
| [101] | COVID-19 disease detection | CNN | Mixed dataset |
| [102] | COVID-19 disease detection | CNN | Cohen’s dataset |

[6] COVID-19 disease detection and image

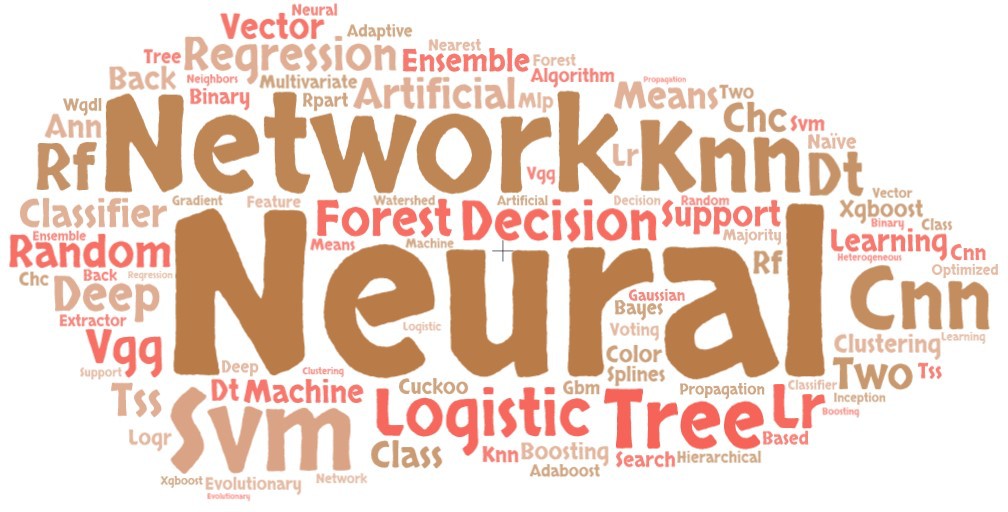
segmentation

CNN Cohen’s dataset

.

## Algorithm and Dataset Analysis

**Most studies acknowledged multiple algorithms in MLBDD strategies. In this context, we use the term multiple algorithms to describe hybrid strategies. For example, Sun et al. (2021) utilized hybrid approaches in predicting coronary healthcare disease using Gaussian Naïve Bayes, Bernoulli Naïve Bayes and Random Forest (RF) algorithms. Bemando et al. (2021) used CNN and SVM in the automation of the diagnosis of Alzheimer’s disease and mild cognitive impairment. Saxena et al. (2019) used KNN and Decision Tree (DT) for diagnosing Heart disease; and Elsalamony (2018) used Neural Networks (NN) and SVM to detect Anaemia disease in human red blood cells. The main advantage of using hybrid strategies is the high level of accuracy compared to using a singular ML model.The literature indicates that the most common individual algorithms for MLBDD models using ML techniques are CNN, SVM, and LR. Specifically, Kalaiselvi et al. (2020) used a CNN based approach for detecting Brain tumors, Dai et al. (2019) developed a device inference application for Skin cancer detection using CNN, Fathi et al. (2020) classified liver disease using SVM, Sing et al. (2019) used SVM to classify Heart disease symptomology , and Basheer et al. (2019) detected Heart disease using Logistic Regression. Figure 10 below represents the most common**

****

**Figure 10.** Word cloud for most frequently used ML algorithms in MLBDD publications.

|  |  |  |  |
| --- | --- | --- | --- |
| **Study** | **Disease** | **Dataset** | **URL** |
| [41–45,134,135] | Heart disease | Cleveland database | https://archive.ics.uci.edu/ml/datasets/heart+disease |

[13,50,51,54] Kidney disease Chronic kidney disease

dataset

https://archive.ics.uci.edu/ml/datasets/Chronic\_ Kidney\_Disease

[71,72,74,136] Diabetics Pima diabetic dataset https://[www.kaggle.com/uciml/pima-indians-diabetesdatabase](http://www.kaggle.com/uciml/pima-indians-diabetesdatabase)

[16,81,85,86] Parkinson disease

Parkinsons Dataset https://archive.ics.uci.edu/ml/datasets/Parkinsons

[58–60] Breast cancer WDBC dataset https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+

Wisconsin+(Diagnostic)

[2,97] COVID-19 Covid-chest X-ray dataset

https://github.com/ieee8023/covid-chestxray-dataset

## Discussion

This study's annotated literature has reinforced the growing importance of machine learning (ML) and deep learning (DL) in disease diagnosis over the past ten years. The review started with particular research questions and used the reference literature to try and answer them. According to extensive research, CNN is one of the newest algorithms and outperforms all other machine learning algorithms because it performs well with both tabular and image data [94,123,128,137]. Because transfer learning outperforms conventional ML techniques and does not necessitate creating a CNN model from scratch, it is also gaining popularity [47,91]. According to the reference literature, SVM, RF, and DT are among the most often used algorithms in MLBDD, aside from CNN.

## Research Challenges and Future Agenda

Even though machine learning-based applications have been widely used in the diagnosis of diseases, researchers and practitioners still encounter a number of difficulties when putting them into practice in the healthcare industry. The following is a summary of the main difficulties with using machine learning to diagnose diseases.:

* 1. *Data Related Challenges*

1. Data scarcity: Even though many patients’ data has been recorded by different hospitals and healthcare, due to the data privacy act, real-world data is not often available for global research purposes.
2. Noisy data: Frequently, the clinical data contains noise or missing values; therefore, such kind of data takes a reasonable amount of time to make it trainable.
3. Adversarial attack: Adversarial attack is one of the key issues in the disease dataset. Adversarial attack means the manipulation of training data, testing data, or machine learning model to result in wrong output from ML.
   1. *Disease Diagnosis-Related Challenges*
4. Misclassification: While the machine learning model can be used to develop as a disease diagnosis model, any misclassification for a particular disease might bring

severe damage. For instance, if a patient with stomach cancer is diagnosed as a non-cancer patient, it will have a huge impact.

1. Wrong image segmentation: One of the key challenges with the ML model is that the model often identifies the wrong region as an infected region. For instance, author Ahsan et al. (2020) shows that even though the accuracy is around 100% in detecting COVID-19 and

non-COVID-19 patients, the pre-trained CNN models such as VGG16 and VGG19 often pay attention to the wrong region during the training process [2]. As a result, it also raises the question of the validity of the MLBDD.

1. Confusion: Some of the diseases such as COVID-19, pneumonia, edema in the chest often demonstrate similar symptoms; in these particular cases, many CNN models detect all of the data samples into one class, i.e., COVID-19.
   1. *Algorithm Related Challenges*
2. Supervised vs. unsupervised: Most machine learning models, like linear regression and logistic regression, really shine when they have labeled data to work with. But when it comes to unlabeled data, their performance tends to take a hit. On the flip side, there are popular algorithms that handle unlabeled data quite well, such as K-means clustering, SVM, and KNN. However, even these can struggle when faced with multidimensional data.
3. Blackbox-related challenges: Convolutional neural networks are among the most commonly used machine learning algorithms. Yet, one of the major hurdles with this approach is that it can be quite tricky to understand how the model tweaks its internal settings, like the learning rate and weights. In the healthcare sector, using such a model requires clear explanations to ensure proper implementation.
   1. *Future Directions*

The challenges discussed in the previous section could pave the way for future researchers and practitioners. We've highlighted some potential algorithms and applications that could help tackle the current issues in MLBDD.

1. GAN-based approach: Generative Adversarial Networks (GANs) have become a go-to method in the deep learning arena. This technique allows us to create synthetic data that closely resembles real data, making it a solid choice for addressing data scarcity. Plus, it lessens our reliance on actual data and supports compliance with data privacy regulations.
2. Explainable AI: This is a hot topic right now, as it focuses on making the behavior of algorithms clear during both training and prediction phases. While there are still hurdles to overcome in the realm of explainable AI, enhancing interpretability and transparency is crucial for deploying machine learning models effectively in real-world scenarios.
3. Ensemble-based approach: Thanks to advancements in technology, we can now capture high-resolution and multidimensional data. Traditional machine learning methods may struggle with such high-quality data, but combining multiple machine learning models can be a fantastic strategy for managing this complex, high-dimensional information.

## Conclusions and Potential Remarks

This study took a close look at papers published from 2012 to 2021 that delve into Machine Learning-based Disease Diagnosis (MLBDD). Researchers have shown a keen interest in several diseases, including heart disease, breast cancer, kidney disease, diabetes, Alzheimer’s, and Parkinson’s, all of which are examined through the lens of machine learning and deep learning techniques. The study also touches on various other ML-based approaches to disease diagnosis. Before diving into the specifics, a bibliometric analysis was conducted, considering factors like

## References

subject area, publication year, journal, and country, while also pinpointing the leading contributors in the MLBDD arena. Our bibliometric findings reveal that the use of machine learning in disease diagnosis has surged dramatically since 2017. When it comes to the sheer number of publications over the years, the top three journals are IEEE Access, Scientific Reports, and the International Journal of Advanced Computer Science and Applications. The most-cited works in MLBDD are by Motwani et al. (2017), Gray et al. (2013), and Mohan et al. (2019). In terms of overall output, China, the United States, and India lead the pack as the most productive countries. Notably, Kim J stands out as the most influential author, with around 20 publications between 2012 and 2021, followed closely by Wang Y and Li J in second and third place, respectively. Approximately 40% of the publications hail from computer science, while around 31% come from engineering, showcasing their dominance in the MLBDD field. In the end, we meticulously selected 102 papers for a deeper analysis, with our key findings highlighted in the discussion sections. Our main takeaway is that deep learning has emerged as the go-to method for researchers, thanks to its impressive ability to build robust models. However, it’s worth noting that while deep learning is widely used in MLBDD, many studies fall short in providing clear explanations for their final predictions.

1. McPhee, S.J.; Papadakis, M.A.; Rabow, M.W. (Eds.) *Current Medical Diagnosis & Treatment*; McGraw-Hill Medical: New York, NY, USA, 2010.
2. Ahsan, M.M.; Ahad, M.T.; Soma, F.A.; Paul, S.; Chowdhury, A.; Luna, S.A.; Yazdan, M.M.S.; Rahman, A.; Siddique, Z.; Huebner, P. Detecting SARS-CoV-2 From Chest X-ray Using Artificial Intelligence. *IEEE Access* **2021**, *9*, 35501–35513. [CrossRef] [PubMed]
3. Coon, E.R.; Quinonez, R.A.; Moyer, V.A.; Schroeder, A.R. Overdiagnosis: How our compulsion for diagnosis may be harming children.

*Pediatrics* **2014**, *134*, 1013–1023. [CrossRef] [PubMed]

1. Balogh, E.P.; Miller, B.T.; Ball, J.R. *Improving Diagnosis in Health Care*; National Academic Press: Washington, DC, USA, 2015. [CrossRef]
2. Ahsan, M.M.; Siddique, Z. Machine Learning-Based Heart Disease Diagnosis: A Systematic Literature Review. *arXiv* **2021**, arXiv:2112.06459.
3. Ahsan, M.M.; E Alam, T.; Trafalis, T.; Huebner, P. Deep MLP-CNN model using mixed-data to distinguish between COVID-19 and Non-COVID- 19 patients. *Symmetry* **2020**, *12*, 1526. [CrossRef]
4. Stafford, I.; Kellermann, M.; Mossotto, E.; Beattie, R.; MacArthur, B.; Ennis, S. A systematic review of the applications of artificial intelligence and machine learning in autoimmune diseases. *NPJ Digit. Med.* **2020**, *3*, 1–11. [CrossRef]
5. Ahsan, M.M.; Gupta, K.D.; Islam, M.M.; Sen, S.; Rahman, M.; Shakhawat Hossain, M. COVID-19 symptoms detection based on nasnetmobile with explainable ai using various imaging modalities. *Mach. Learn. Knowl. Extr.* **2020**, *2*, 490–504. [CrossRef] 9. Samuel, A.L. Some studies in machine learning using the game of checkers. *IBM J. Res. Dev.* **1959**, *3*, 210–229. [CrossRef]
6. Brownlee, J. Machine learning mastery with Python. *Mach. Learn. Mastery Pty Ltd.* **2016**, *527*, 100–120.
7. Houssein, E.H.; Emam, M.M.; Ali, A.A.; Suganthan, P.N. Deep and machine learning techniques for medical imaging-based breast cancer: A comprehensive review. *Expert Syst. Appl.* **2021**, *167*, 114161. [CrossRef]
8. Brijain, M.; Patel, R.; Kushik, M.; Rana, K. A survey on decision tree algorithm for classification. *Int. J. Eng. Dev. Res.* **2014**. Available online: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.673.2797> (accessed on 10 December 2021).
9. Walse, R.S.; Kurundkar, G.D.; Khamitkar, S.D.; Muley, A.A.; Bhalchandra, P.U.; Lokhande, S.N. Effective Use of Naïve Bayes,

Decision Tree, and Random Forest Techniques for Analysis of Chronic Kidney Disease. In Proceedings of the International Conference on Information and Communication Technology for Intelligent Systems, Ahmedabad, India, 15–16 May 2020; Springer: Berlin/Heidelberg, Germany, 2020; pp. 237–245.

1. Rajendran, K.; Jayabalan, M.; Thiruchelvam, V. Predicting breast cancer via supervised machine learning methods on class imbalanced data.

*Int. J. Adv. Comput. Sci. Appl.* **2020**, *11*, 54–63. [CrossRef]

1. Tsao, H.Y.; Chan, P.Y.; Su, E.C.Y. Predicting diabetic retinopathy and identifying interpretable biomedical features using machine learning algorithms. *BMC Bioinform.* **2018**, *19*, 111–121. [CrossRef] [PubMed]
2. Nurrohman, A.; Abdullah, S.; Murfi, H. Parkinson’s disease subtype classification: Application of decision tree, logistic regression and logit leaf model. In *AIP Conference Proceedings*; AIP Publishing LLC: Melville, NY, USA, 2020; Volume 2242, p. 030015.
3. Drucker, H.; Wu, D.; Vapnik, V.N. Support vector machines for spam categorization. *IEEE Trans. Neural Netw.* **1999**, *10*, 1048–1054. [CrossRef] [PubMed]
4. Fix, E.; Hodges, J.L. Discriminatory analysis. Nonparametric discrimination: Consistency properties. *Int. Stat. Rev. Int. De Stat.* **1989**, *57*, 238–

247. [CrossRef]

1. Wright, R.E. Logistic regression. In *Reading and Understanding Multivariate Statistics*; American Psychological Association: Washington, DC, USA, **1995**.
2. Schapire, R.E. Explaining adaboost. In *Empirical Inference*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 37–52.
3. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016.
4. Hayashi, Y. The right direction needed to develop white-box deep learning in radiology, pathology, and ophthalmology: A short review. *Front. Robot. AI* **2019**, *6*, 24. [CrossRef]

23. Akkus, Z.; Galimzianova, A.; Hoogi, A.; Rubin, D.L.; Erickson, B.J. Deep learning for brain MRI segmentation: State of the art

and future directions. J. Digit. Imaging 2017, 30, 449–459. [CrossRef]

23. Yap, M.H.; Pons, G.; Martí, J.; Ganau, S.; Sentís, M.; Zwiggelaar, R.; Davison, A.K.; Marti, R. Automated breast ultrasound lesions

detection using convolutional neural networks. IEEE J. Biomed. Health Inform. 2017, 22, 1218–1226. [CrossRef]

24. Malviya, R.K.; Kant, R. Green supply chain management (GSCM): A structured literature review and research implications.

Benchmarking Int. J. 2015, 22, 1360–1394. [CrossRef]

25. Fahimnia, B.; Sarkis, J.; Davarzani, H. Green supply chain management: A review and bibliometric analysis. Int. J. Prod. Econ.

2015, 162, 101–114. [CrossRef]

26. Motwani, M.; Dey, D.; Berman, D.S.; Germano, G.; Achenbach, S.; Al-Mallah, M.H.; Andreini, D.; Budoff, M.J.; Cademartiri, F.;

Callister, T.Q.; et al. Machine learning for prediction of all-cause mortality in patients with suspected coronary artery disease: A

5-year multicentre prospective registry analysis. Eur. Heart J. 2017, 38, 500–507. [CrossRef]

27. Gray, K.R.; Aljabar, P.; Heckemann, R.A.; Hammers, A.; Rueckert, D.; Initiative, A.D.N. Random forest-based similarity measures

for multi-modal classification of Alzheimer’s disease. NeuroImage 2013, 65, 167–175. [CrossRef] [PubMed]

28. Mohan, S.; Thirumalai, C.; Srivastava, G. Effective heart disease prediction using hybrid machine learning techniques. IEEE

Access 2019, 7, 81542–81554. [CrossRef]

29. Yadav, S.S.; Jadhav, S.M. Deep convolutional neural network based medical image classification for disease diagnosis. J. Big Data

2019, 6, 1–18. [CrossRef]

30. Zhang, Y.; Dong, Z.; Phillips, P.; Wang, S.; Ji, G.; Yang, J.; Yuan, T.F. Detection of subjects and brain regions related to Alzheimer’s

disease using 3D MRI scans based on eigenbrain and machine learning. Front. Comput. Neurosci. 2015, 9, 66. [CrossRef] [PubMed]

31. Austin, P.C.; Tu, J.V.; Ho, J.E.; Levy, D.; Lee, D.S. Using methods from the data-mining and machine-learning literature for disease

classification and prediction: A case study examining classification of heart failure subtypes. J. Clin. Epidemiol. 2013, 66, 398–407.

[CrossRef]Healthcare 2022, 10, 541 26 of 30

32. 33. 34. 35. 36. 37. 38. 39. 40. 41. 42. 43. 44. 45. 46. 47. 48. 49. 50. 51. 52. 53. 54. 55. 56. Sharmila, A.; Geethanjali, P. DWT based detection of epileptic seizure from EEG signals using naive Bayes and k-NN classifiers.

IEEE Access 2016, 4, 7716–7727. [CrossRef]

Lebedev, A.; Westman, E.; Van Westen, G.; Kramberger, M.; Lundervold, A.; Aarsland, D.; Soininen, H.; Kłoszewska, I.; Mecocci, P.;

Tsolaki, M.; et al. Random Forest ensembles for detection and prediction of Alzheimer’s disease with a good between-cohort

robustness. NeuroImage Clin. 2014, 6, 115–125. [CrossRef]

Luz, E.J.d.S.; Nunes, T.M.; De Albuquerque, V.H.C.; Papa, J.P.; Menotti, D. ECG arrhythmia classification based on optimum-path

forest. Expert Syst. Appl. 2013, 40, 3561–3573. [CrossRef]

Challis, E.; Hurley, P.; Serra, L.; Bozzali, M.; Oliver, S.; Cercignani, M. Gaussian process classification of Alzheimer’s disease and

mild cognitive impairment from resting-state fMRI. NeuroImage 2015, 112, 232–243. [CrossRef]

Ansari, A.Q.; Gupta, N.K. Automated diagnosis of coronary heart disease using neuro-fuzzy integrated system. In Proceedings of

the 2011 World Congress on Information and Communication Technologies, Mumbai, India, 11–14 December 2011; pp. 1379–1384.

Ahsan, M.M.; Mahmud, M.; Saha, P.K.; Gupta, K.D.; Siddique, Z. Effect of data scaling methods on machine Learning algorithms

and model performance. Technologies 2021, 9, 52. [CrossRef]

Rubin, J.; Abreu, R.; Ganguli, A.; Nelaturi, S.; Matei, I.; Sricharan, K. Recognizing abnormal heart sounds using deep learning.

arXiv 2017, arXiv:1707.04642.

Miao, J.H.; Miao, K.H. Cardiotocographic diagnosis of fetal health based on multiclass morphologic pattern predictions using

deep learning classification. Int. J. Adv. Comput. Sci. Appl. 2018, 9, 1–11. [CrossRef]

Bemando, C.; Miranda, E.; Aryuni, M. Machine-Learning-Based Prediction Models of Coronary Heart Disease Using Naïve Bayes

and Random Forest Algorithms. In Proceedings of the 2021 International Conference on Software Engineering & Computer

Systems and 4th International Conference on Computational Science and Information Management (ICSECS-ICOCSIM), Pekan,

Malaysia, 24–28 August 2021; pp. 232–237.

Kumar, R.R.; Polepaka, S. Performance Comparison of Random Forest Classifier and Convolution Neural Network in Predicting

Heart Diseases. In ICCII 2018, Proceedings of the Third International Conference on Computational Intelligence and Informatics; Springer:

Singapore, 2020; pp. 683–691. [CrossRef]

Singh, H.; Navaneeth, N.; Pillai, G. Multisurface Proximal SVM Based Decision Trees For Heart Disease Classification. In

Proceedings of the TENCON 2019–2019 IEEE Region 10 Conference (TENCON), Kerala, India, 17–20 October 2019; pp. 13–18.

Desai, S.D.; Giraddi, S.; Narayankar, P.; Pudakalakatti, N.R.; Sulegaon, S. Back-propagation neural network versus logistic

regression in heart disease classification. In Advanced Computing and Communication Technologies; Springer: Berlin/Heidelberg,

Germany, 2019; pp. 133–144.

Patil, D.D.; Singh, R.; Thakare, V.M.; Gulve, A.K. Analysis of ECG Arrhythmia for Heart Disease Detection using SVM and

Cuckoo Search Optimized Neural Network. Int. J. Eng. Technol. 2018, 7, 27–33. [CrossRef]

Liu, N.; Lin, Z.; Cao, J.; Koh, Z.; Zhang, T.; Huang, G.B.; Ser, W.; Ong, M.E.H. An intelligent scoring system and its application to

cardiac arrest prediction. IEEE Trans. Inf. Technol. Biomed. 2012, 16, 1324–1331. [CrossRef]

Acharya, U.R.; Oh, S.L.; Hagiwara, Y.; Tan, J.H.; Adam, M.; Gertych, A.; San Tan, R. A deep convolutional neural network model

to classify heartbeats. Comput. Biol. Med. 2017, 89, 389–396. [CrossRef]

Yang, W.; Si, Y.; Wang, D.; Guo, B. Automatic recognition of arrhythmia based on principal component analysis network and

linear support vector machine. Comput. Biol. Med. 2018, 101, 22–32. [CrossRef]

Levey, A.S.; Coresh, J. Chronic kidney disease. Lancet 2012, 379, 165–180. [CrossRef]

Charleonnan, A.; Fufaung, T.; Niyomwong, T.; Chokchueypattanakit, W.; Suwannawach, S.; Ninchawee, N. Predictive analytics

for chronic kidney disease using machine learning techniques. In Proceedings of the 2016 Management and Innovation

Technology International Conference, Bang-Saen, Chonburi, Thailand, 12–14 October 2016; pp. MIT-80–MIT-83.

Aljaaf, A.J.; Al-Jumeily, D.; Haglan, H.M.; Alloghani, M.; Baker, T.; Hussain, A.J.; Mustafina, J. Early prediction of chronic kidney

disease using machine learning supported by predictive analytics. In Proceedings of the 2018 IEEE Congress on Evolutionary

Computation (CEC), Rio de Janeiro, Brazil, 8–13 July 2018; pp. 1–9.

Ma, F.; Sun, T.; Liu, L.; Jing, H. Detection and diagnosis of chronic kidney disease using deep learning-based heterogeneous

modified artificial neural network. Future Gener. Comput. Syst. 2020, 111, 17–26. [CrossRef]

Nithya, A.; Appathurai, A.; Venkatadri, N.; Ramji, D.; Palagan, C.A. Kidney disease detection and segmentation using artificial

neural network and multi-kernel k-means clustering for ultrasound images. Measurement 2020, 149, 106952. [CrossRef]

Al Imran, A.; Amin, M.N.; Johora, F.T. Classification of chronic kidney disease using logistic regression, feedforward neural

network and wide & deep learning. In Proceedings of the 2018 International Conference on Innovation in Engineering and

Technology (ICIET), Dhaka, Bangladesh, 27–28 December 2018; pp. 1–6.

Navaneeth, B.; Suchetha, M. A dynamic pooling based convolutional neural network approach to detect chronic kidney disease.

Biomed. Signal Process. Control 2020, 62, 102068. [CrossRef]

Brunetti, A.; Cascarano, G.D.; De Feudis, I.; Moschetta, M.; Gesualdo, L.; Bevilacqua, V. Detection and segmentation of

kidneys from magnetic resonance images in patients with autosomal dominant polycystic kidney disease. In Proceedings of the

International Conference on Intelligent Computing, Nanchang, China, 3–6 August 2019; Springer: Berlin/Heidelberg, Germany,

2019; pp. 639–650.

Miranda, G.H.B.; Felipe, J.C. Computer-aided diagnosis system based on fuzzy logic for breast cancer categorization. Comput.

Biol. Med. 2015, 64, 334–346. [CrossRef]Healthcare 2022, 10, 541 27 of 30

57. 58. 59. 60. 61. 62. 63. 64. 65. 66. 67. 68. 69. 70. 71. 72. 73. 74. 75. 76. 77. 78. 79. 80. 81. 82. Zheng, B.; Yoon, S.W.; Lam, S.S. Breast cancer diagnosis based on feature extraction using a hybrid of K-means and support

vector machine algorithms. Expert Syst. Appl. 2014, 41, 1476–1482. [CrossRef]

Asri, H.; Mousannif, H.; Al Moatassime, H.; Noel, T. Using machine learning algorithms for breast cancer risk prediction and

diagnosis. Procedia Comput. Sci. 2016, 83, 1064–1069. [CrossRef]

Mohammed, S.A.; Darrab, S.; Noaman, S.A.; Saake, G. Analysis of breast cancer detection using different machine learning

techniques. In Proceedings of the International Conference on Data Mining and Big Data, Belgrade, Serbia, 14–20 July 2020;

Springer: Berlin/Heidelberg, Gemany, 2020; pp. 108–117.

Assegie, T.A. An optimized K-Nearest Neighbor based breast cancer detection. J. Robot. Control (JRC) 2021, 2, 115–118. [CrossRef]

Bhattacherjee, A.; Roy, S.; Paul, S.; Roy, P.; Kausar, N.; Dey, N. Classification approach for breast cancer detection using back

propagation neural network: A study. In Deep Learning and Neural Networks: Concepts, Methodologies, Tools, and Applications; IGI

Global: Hershey, PA, USA, 2020; pp. 1410–1421.

Alshayeji, M.H.; Ellethy, H.; Gupta, R. Computer-aided detection of breast cancer on the Wisconsin dataset: An artificial neural

networks approach. Biomed. Signal Process. Control 2022, 71, 103141. [CrossRef]

Sultana, Z.; Khan, M.R.; Jahan, N. Early breast cancer detection utilizing artificial neural network. WSEAS Trans. Biol. Biomed.

2021, 18, 32–42. [CrossRef]

Ghosh, P.; Azam, S.; Hasib, K.M.; Karim, A.; Jonkman, M.; Anwar, A. A performance based study on deep learning algorithms in

the effective prediction of breast cancer. In Proceedings of the 2021 International Joint Conference on Neural Networks (IJCNN),

Online, 18–22 July 2021; pp. 1–8.

Abdel-Nasser, M.; Rashwan, H.A.; Puig, D.; Moreno, A. Analysis of tissue abnormality and breast density in mammographic

images using a uniform local directional pattern. Expert Syst. Appl. 2015, 42, 9499–9511. [CrossRef]

Sharma, S.; Khanna, P. Computer-aided diagnosis of malignant mammograms using Zernike moments and SVM. J. Digit. Imaging

2015, 28, 77–90. [CrossRef]

Moon, W.K.; Chen, I.L.; Chang, J.M.; Shin, S.U.; Lo, C.M.; Chang, R.F. The adaptive computer-aided diagnosis system based on

tumor sizes for the classification of breast tumors detected at screening ultrasound. Ultrasonics 2017, 76, 70–77. [CrossRef]

Lo, C.M.; Chan, S.W.; Yang, Y.W.; Chang, Y.C.; Huang, C.S.; Jou, Y.S.; Chang, R.F. Feasibility testing: Three-dimensional tumor

mapping in different orientations of automated breast ultrasound. Ultrasound Med. Biol. 2016, 42, 1201–1210. [CrossRef]

Venkatesh, S.S.; Levenback, B.J.; Sultan, L.R.; Bouzghar, G.; Sehgal, C.M. Going beyond a first reader: A machine learning

methodology for optimizing cost and performance in breast ultrasound diagnosis. Ultrasound Med. Biol. 2015, 41, 3148–3162.

[CrossRef] [PubMed]

Naz, H.; Ahuja, S. Deep learning approach for diabetes prediction using PIMA Indian dataset. J. Diabetes Metab. Disord. 2020,

19, 391–403. [CrossRef] [PubMed]

Kandhasamy, J.P.; Balamurali, S. Performance analysis of classifier models to predict diabetes mellitus. Procedia Comput. Sci. 2015,

47, 45–51. [CrossRef]

Yahyaoui, A.; Jamil, A.; Rasheed, J.; Yesiltepe, M. A decision support system for diabetes prediction using machine learning

and deep learning techniques. In Proceedings of the 2019 1st International Informatics and Software Engineering Conference

(UBMYK), Ankara, Turkey, 6–7 November 2019; pp. 1–4.

Ashiquzzaman, A.; Tushar, A.K.; Islam, M.; Shon, D.; Im, K.; Park, J.H.; Lim, D.S.; Kim, J. Reduction of overfitting in diabetes

prediction using deep learning neural network. arXiv 2017, arXiv:1707.08386.

Alhassan, Z.; McGough, A.S.; Alshammari, R.; Daghstani, T.; Budgen, D.; Al Moubayed, N. Type-2 diabetes mellitus diagnosis

from time series clinical data using deep learning models. In Proceedings of the 2019 International Conference on Artificial

Neural Networks, Rhodes, Greece, 4–7 October 2018; Springer: Berlin/Heidelberg, Germany, 2018; pp. 468–478.

Fitriyani, N.L.; Syafrudin, M.; Alfian, G.; Rhee, J. Development of disease prediction model based on ensemble learning approach

for diabetes and hypertension. IEEE Access 2019, 7, 144777–144789. [CrossRef]

Fernández-Edreira, D.; Liñares-Blanco, J.; Fernandez-Lozano, C. Machine Learning analysis of the human infant gut microbiome

identifies influential species in type 1 diabetes. Expert Syst. Appl. 2021, 185, 115648. [CrossRef]

Ali, A.; Alrubei, M.A.; Hassan, L.F.M.; Al-Ja’afari, M.A.; Abdulwahed, S.H. Diabetes Diagnosis based on KNN. IIUM Eng. J. 2020,

21, 175–181. [CrossRef]

Qtea, H.; Awad, M. Using Hybrid Model of Particle Swarm Optimization and Multi-Layer Perceptron Neural Networks for

Classification of Diabete. Int. J. Intell. Eng. Syst. 2021, 14, 11–22.

Grover, S.; Bhartia, S.; Yadav, A.; Seeja, K.R. Predicting severity of Parkinson’s disease using deep learning. Procedia Comput. Sci.

2018, 132, 1788–1794. [CrossRef]

Sriram, T.V.; Rao, M.V.; Narayana, G.S.; Kaladhar, D.; Vital, T.P.R. Intelligent Parkinson disease prediction using machine learning

algorithms. Int. J. Eng. Innov. Technol. (IJEIT) 2013, 3, 1568–1572.

Esmaeilzadeh, S.; Yang, Y.; Adeli, E. End-to-end Parkinson disease diagnosis using brain mr-images by 3d-cnn. arXiv 2018,

arXiv:1806.05233.

Warjurkar, S.; Ridhorkar, S. A Study on Brain Tumor and Parkinson’s Disease Diagnosis and Detection using Deep Learning. In

Proceedings of the 3rd International Conference on Integrated Intelligent Computing Communication & Security (ICIIC 2021),

Online, 27–28 August 2021; Atlantis Press: Amsterdam, The Netherlands, 2021; pp. 356–364.Healthcare 2022, 10, 541 28 of 30

83. Sherly Puspha Annabel, L.; Sreenidhi, S.; Vishali, N. A Novel Diagnosis System for Parkinson’s Disease Using K-means Clustering

and Decision Tree. In Communication and Intelligent Systems; Springer: Berlin/Heidelberg, Germany, 2021; pp. 607–615.

84. Asmae, O.; Abdelhadi, R.; Bouchaib, C.; Sara, S.; Tajeddine, K. Parkinson’s disease identification using KNN and ANN Algorithms

based on Voice Disorder. In Proceedings of the 2020 1st International Conference on Innovative Research in Applied Science,

Engineering and Technology (IRASET), Meknes, Morocco, 19–20 March 2020; IEEE: Manhattan, NY, USA, 2020; pp. 1–6.

85. Gürüler, H. A novel diagnosis system for Parkinson’s disease using complex-valued artificial neural network with k-means

clustering feature weighting method. Neural Comput. Appl. 2017, 28, 1657–1666. [CrossRef]

86. Shetty, S.; Rao, Y. SVM based machine learning approach to identify Parkinson’s disease using gait analysis. In Proceedings of

the 2016 International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, 26–27 August 2016; IEEE:

Manhattan, NY, USA, 2016; Volume 2, pp. 1–5.

87. Ahsan, M.M.; Nazim, R.; Siddique, Z.; Huebner, P. Detection of COVID-19 patients from ct scan and chest X-ray data using

modified mobilenetv2 and lime. Healthcare 2021, 9, 1099. [CrossRef]

88. Haghanifar, A.; Majdabadi, M.M.; Choi, Y.; Deivalakshmi, S.; Ko, S. Covid-cxnet: Detecting COVID-19 in frontal chest X-ray

images using deep learning. arXiv 2020, arXiv:2006.13807.

89. Tahamtan, A.; Ardebili, A. Real-time RT-PCR in COVID-19 detection: Issues affecting the results. Expert Rev. Mol. Diagn. 2020,

20, 453–454. [CrossRef] [PubMed]

90. Chen, J.; Wu, L.; Zhang, J.; Zhang, L.; Gong, D.; Zhao, Y.; Chen, Q.; Huang, S.; Yang, M.; Yang, X.; et al. Deep learning-based

model for detecting 2019 novel coronavirus pneumonia on high-resolution computed tomography. Sci. Rep. 2020, 10, 1–11.

[CrossRef] [PubMed]

91. Ardakani, A.A.; Kanafi, A.R.; Acharya, U.R.; Khadem, N.; Mohammadi, A. Application of deep learning technique to manage

COVID-19 in routine clinical practice using CT images: Results of 10 convolutional neural networks. Comput. Biol. Med. 2020,

121, 103795. [CrossRef]

92. Wang, L.; Lin, Z.Q.; Wong, A. Covid-net: A tailored deep convolutional neural network design for detection of COVID-19 cases

from chest X-ray images. Sci. Rep. 2020, 10, 1–12. [CrossRef]

93. Li, L.; Qin, L.; Xu, Z.; Yin, Y.; Wang, X.; Kong, B.; Bai, J.; Lu, Y.; Fang, Z.; Song, Q.; et al. Artificial intelligence distinguishes

COVID-19 from community acquired pneumonia on chest CT. Radiology 2020. [CrossRef]

94. Hemdan, E.E.D.; Shouman, M.A.; Karar, M.E. Covidx-net: A framework of deep learning classifiers to diagnose COVID-19 in

X-ray images. arXiv 2020, arXiv:2003.11055.

95. Sethy, P.K.; Behera, S.K. Detection of Coronavirus Disease (COVID-19) Based on Deep Features and Support Vector Machine.

2020. Available online: https://pdfs.semanticscholar.org/9da0/35f1d7372cfe52167ff301bc12d5f415caf1.pdf (accessed on 10

December 2021).

96. Narin, A.; Kaya, C.; Pamuk, Z. Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional

neural networks. Pattern Anal. Appl. 2021, 24, 1207–1220. [CrossRef] [PubMed]

97. Brunese, L.; Mercaldo, F.; Reginelli, A.; Santone, A. Explainable deep learning for pulmonary disease and coronavirus COVID-19

detection from X-rays. Comput. Methods Programs Biomed. 2020, 196, 105608. [CrossRef] [PubMed]

98. Ghoshal, B.; Tucker, A. Estimating uncertainty and interpretability in deep learning for coronavirus (COVID-19) detection. arXiv

2020, arXiv:2003.10769.

99. Apostolopoulos, I.D.; Mpesiana, T.A. COVID-19: Automatic detection from X-ray images utilizing transfer learning with

convolutional neural networks. Phys. Eng. Sci. Med. 2020, 43, 635–640. [CrossRef] [PubMed]

100. Song, Y.; Zheng, S.; Li, L.; Zhang, X.; Zhang, X.; Huang, Z.; Chen, J.; Wang, R.; Zhao, H.; Chong, Y.; et al. Deep learning

enables accurate diagnosis of novel coronavirus (COVID-19) with CT images. IEEE/ACM Trans. Comput. Biol. Bioinform. 2021,

18, 2775–2780. [CrossRef] [PubMed]

101. Jin, C.; Chen, W.; Cao, Y.; Xu, Z.; Tan, Z.; Zhang, X.; Deng, L.; Zheng, C.; Zhou, J.; Shi, H.; et al. Development and evaluation of an

artificial intelligence system for COVID-19 diagnosis. Nat. Commun. 2020, 11, 5088. [CrossRef]

102. 103. Graham, N.; Warner, J. Alzheimer’s Disease and Other Dementias; Family Doctor Publications Limited: Northampton, UK, 2009.

Neelaveni, J.; Devasana, M.G. Alzheimer disease prediction using machine learning algorithms. In Proceedings of the 2020 6th

International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 6–7 March 2020;

IEEE: Manhattan, NY, USA, 2020; pp. 101–104.

104. Collij, L.E.; Heeman, F.; Kuijer, J.P.; Ossenkoppele, R.; Benedictus, M.R.; Möller, C.; Verfaillie, S.C.; Sanz-Arigita, E.J.; van Berckel,

B.N.; van der Flier, W.M.; et al. Application of machine learning to arterial spin labeling in mild cognitive impairment and

Alzheimer disease. Radiology 2016, 281, 865–875. [CrossRef]

105. Vidushi, A.R.; Shrivastava, A.K. Diagnosis of Alzheimer disease using machine learning approaches. Int. J. Adv. Sci. Technol.

2019, 29, 7062–7073.

106. Ahmed, S.; Kim, B.C.; Lee, K.H.; Jung, H.Y.; Initiative, A.D.N. Ensemble of ROI-based convolutional neural network classifiers

for staging the Alzheimer disease spectrum from magnetic resonance imaging. PLoS ONE 2020, 15, e0242712. [CrossRef]

107. Nawaz, H.; Maqsood, M.; Afzal, S.; Aadil, F.; Mehmood, I.; Rho, S. A deep feature-based real-time system for Alzheimer disease

stage detection. Multimed. Tools Appl. 2020, 1–19. [CrossRef]Healthcare 2022, 10, 541 29 of 30

108. Haft-Javaherian, M.; Fang, L.; Muse, V.; Schaffer, C.B.; Nishimura, N.; Sabuncu, M.R. Deep convolutional neural networks for

segmenting 3D in vivo multiphoton images of vasculature in Alzheimer disease mouse models. PLoS ONE 2019, 14, e0213539.

[CrossRef] [PubMed]

109. Aderghal, K.; Benois-Pineau, J.; Afdel, K. Classification of sMRI for Alzheimer’s disease diagnosis with CNN: Single Siamese

networks with 2D+? Approach and fusion on ADNI. In Proceedings of the 2017 ACM on International Conference on Multimedia

Retrieval, Bucharest, Romania, 6–9 June 2017; pp. 494–498.

110. Sun, M.; Huang, Z.; Guo, C. Automatic Diagnosis of Alzheimer’s Disease and Mild Cognitive Impairment Based on CNN+ SVM

Networks with End-to-end Training. In Proceedings of the 2021 13th International Conference on Advanced Computational

Intelligence (ICACI), Wanzhou, Chongqing, China, 14–16 May 2021; IEEE: Manhattan, NY, USA, 2021; pp. 279–285.

111. Kuang, J.; Zhang, P.; Cai, T.; Zou, Z.; Li, L.; Wang, N.; Wu, L. Prediction of transition from mild cognitive impairment to

Alzheimer’s disease based on a logistic regression–artificial neural network–decision tree model. Geriatr. Gerontol. Int. 2021,

21, 43–47. [CrossRef] [PubMed]

112. Manzak, D.; Çetinel, G.; Manzak, A. Automated Classification of Alzheimer’s Disease using Deep Neural Network (DNN) by

Random Forest Feature Elimination. In Proceedings of the 2019 14th International Conference on Computer Science & Education

(ICCSE), Bandung, Indonesia, 14–15 October 2019; IEEE: Manhattan, NY, USA, 2019; pp. 1050–1053.

113. Mao, Y.; He, Y.; Liu, L.; Chen, X. Disease classification based on eye movement features with decision tree and random forest.

Front. Neurosci. 2020, 14, 798. [CrossRef]

114. Nosseir, A.; Shawky, M.A. Automatic classifier for skin disease using k-NN and SVM. In Proceedings of the 2019 8th International

Conference on Software and Information Engineering, Cairo, Egypt, 9–12 April 2019; pp. 259–262.

115. Khan, M.A.; Ashraf, I.; Alhaisoni, M.; Damaševiˇcius, R.; Scherer, R.; Rehman, A.; Bukhari, S.A.C. Multimodal brain tumor

classification using deep learning and robust feature selection: A machine learning application for radiologists. Diagnostics 2020,

10, 565. [CrossRef] [PubMed]

116. Amin, J.; Sharif, M.; Raza, M.; Yasmin, M. Detection of brain tumor based on features fusion and machine learning. J. Ambient

Intell. Human. Comput. 2018, 1–17. [CrossRef]

117. Dai, X.; Spasi´c, I.; Meyer, B.; Chapman, S.; Andres, F. Machine learning on mobile: An on-device inference app for skin cancer

detection. In Proceedings of the 2019 Fourth International Conference on Fog and Mobile Edge Computing (FMEC), Rome, Italy,

10–13 June 2019; IEEE: Manhattan, NY, USA, 2019; pp. 301–305.

118. Daghrir, J.; Tlig, L.; Bouchouicha, M.; Sayadi, M. Melanoma skin cancer detection using deep learning and classical machine

learning techniques: A hybrid approach. In Proceedings of the 2020 5th International Conference on Advanced Technologies for

Signal and Image Processing (ATSIP), Sfax, Tunisia, 2–5 September 2020; IEEE: Manhattan, NY, USA, 2020; pp. 1–5.

119. Dhaliwal, J.; Erdman, L.; Drysdale, E.; Rinawi, F.; Muir, J.; Walters, T.D.; Siddiqui, I.; Griffiths, A.M.; Church, P.C. Accurate

Classification of Pediatric Colonic Inflammatory Bowel Disease Subtype Using a Random Forest Machine Learning Classifier.

J. Pediatr. Gastroenterol. Nutr. 2021, 72, 262–269. [CrossRef]

120. Fathi, M.; Nemati, M.; Mohammadi, S.M.; Abbasi-Kesbi, R. A machine learning approach based on SVM for classification of liver

diseases. Biomed. Eng. Appl. Basis Commun. 2020, 32, 2050018. [CrossRef]

121. Wang, A.; An, N.; Xia, Y.; Li, L.; Chen, G. A logistic regression and artificial neural network-based approach for chronic disease

prediction: A case study of hypertension. In Proceedings of the 2014 IEEE International Conference on Internet of Things

(iThings), and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing

(CPSCom), Taipei, Taiwan, 1–3 September 2014; IEEE: Manhattan, NY, USA, 2014; pp. 45–52.

122. Kalaiselvi, T.; Padmapriya, S.; Sriramakrishnan, P.; Somasundaram, K. Deriving tumor detection models using convolutional

neural networks from MRI of human brain scans. Int. J. Inf. Technol. 2020, 1–6. [CrossRef]

123. Usman, K.; Rajpoot, K. Brain tumor classification from multi-modality MRI using wavelets and machine learning. Pattern Anal.

Appl. 2017, 20, 871–881. [CrossRef]

124. Waheed, Z.; Waheed, A.; Zafar, M.; Riaz, F. An efficient machine learning approach for the detection of melanoma using

dermoscopic images. In Proceedings of the 2017 International Conference on Communication, Computing and Digital Systems

(C-CODE), Islamabad, Pakistan, 8–9 March 2017; IEEE: Manhattan, NY, USA, 2017; pp. 316–319.

125. Kamboj, A. A color-based approach for melanoma skin cancer detection. In Proceedings of the 2018 First International Conference

on Secure Cyber Computing and Communication (ICSCCC), Jalandhar, India, 15–17 December 2018; IEEE: Manhattan, NY, USA,

2018; pp. 508–513.

126. Magalhaes, C.; Tavares, J.M.R.; Mendes, J.; Vardasca, R. Comparison of machine learning strategies for infrared thermography of

skin cancer. Biomed. Signal Process. Control 2021, 69, 102872. [CrossRef]

127. Chen, M.; Zhang, B.; Topatana, W.; Cao, J.; Zhu, H.; Juengpanich, S.; Mao, Q.; Yu, H.; Cai, X. Classification and mutation

prediction based on histopathology H&E images in liver cancer using deep learning. NPJ Precis. Oncol. 2020, 4, 14.

128. Das, A.; Acharya, U.R.; Panda, S.S.; Sabut, S. Deep learning based liver cancer detection using watershed transform and Gaussian

mixture model techniques. Cogn. Syst. Res. 2019, 54, 165–175. [CrossRef]

129. Wang, Y.; Ji, C.; Wang, Y.; Ji, M.; Yang, J.J.; Zhou, C.M. Predicting postoperative liver cancer death outcomes with machine

learning. Curr. Med Res. Opin. 2021, 37, 629–634. [CrossRef] [PubMed]Healthcare 2022, 10, 541 30 of 30

130. Saxena, R.; Johri, A.; Deep, V.; Sharma, P. Heart diseases prediction system using CHC-TSS Evolutionary, KNN, and decision tree

classification algorithm. In Emerging Technologies in Data Mining and Information Security; Springer: Berlin/Heidelberg, Germany,

2019; pp. 809–819.

131. Elsalamony, H.A. Detection of anaemia disease in human red blood cells using cell signature, neural networks and SVM.

Multimed. Tools Appl. 2018, 77, 15047–15074. [CrossRef]

132. Basheer, S.; Mathew, R.M.; Devi, M.S. Ensembling Coalesce of Logistic Regression Classifier for Heart Disease Prediction using

Machine Learning. Int. J. Innov. Technol. Explor. Eng. 2019, 8, 127–133.

133. Bharti, R.; Khamparia, A.; Shabaz, M.; Dhiman, G.; Pande, S.; Singh, P. Prediction of heart disease using a combination of machine

learning and deep learning. Comput. Intell. Neurosci. 2021, 2021, 8387680. [CrossRef] [PubMed]

134. Saw, M.; Saxena, T.; Kaithwas, S.; Yadav, R.; Lal, N. Estimation of Prediction for Getting Heart Disease Using Logistic Regression

Model of Machine Learning. In Proceedings of the 2020 International Conference on Computer Communication and Informatics

(ICCCI), Da Nang, Vietnam, 30 November–3 December 2020; IEEE: Manhattan, NY, USA, 2020; pp. 1–6.

135. Gill, N.S.; Mittal, P. A computational hybrid model with two level classification using SVM and neural network for predicting the

diabetes disease. J. Theor. Appl. Inf. Technol 2016, 87, 1–10.

136. Sun, G.; Hakozaki, Y.; Abe, S.; Vinh, N.Q.; Matsui, T. A novel infection screening method using a neural network and k-means

clustering algorithm which can be applied for screening of unknown or unexpected infectious diseases. J. Infect. 2012, 65, 591–592.

[CrossRef]

137. Yang, G.; Pang, Z.; Deen, M.J.; Dong, M.; Zhang, Y.T.; Lovell, N.; Rahmani, A.M. Homecare robotic systems for healthcare 4.0:

Visions and enabling technologies. IEEE J. Biomed. Health Inform. 2020, 24, 2535–2549. [CrossRef]

138. Ngiam, K.Y.; Khor, W. Big data and machine learning algorithms for health-care delivery. Lancet Oncol. 2019, 20, e262–e273.

[CrossRef]

139. Zhang, P.; Schmidt, D.C.; White, J.; Lenz, G. Blockchain technology use cases in healthcare. In Advances in Computers; Elsevier:

Amsterdam, The Netherlands, 2018; Volume 111, pp. 1–41.

140. Engelhardt, M.A. Hitching healthcare to the chain: An introduction to blockchain technology in the healthcare sector. Technol.

Innov. Manag. Rev. 2017, 7. [CrossRef]